|  |  |
| --- | --- |
| **Joint Video Experts Team (JVET)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29**  33rd Meeting, by teleconference, 17–26 January 2024 | Document: JVET-AG\_notes\_d7 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Meeting Report of the 33rd Meeting of the Joint Video Experts Team (JVET), by teleconference, 17–26 January 2024** | | |
| *Status:* | Report document from the chair of JVET | | |
| *Purpose:* | Report | | |
| *Author(s) or Contact(s):* | **Jens-Rainer Ohm** Institute of Communication Engineering RWTH Aachen Melatener Straße 23 D-52074 Aachen | Tel: Email: | +49 241 80 27671 [ohm@ient.rwth-aachen.de](mailto:ohm@ient.rwth-aachen.de) |
| *Source:* | Chair of JVET | | |

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

# Summary

The Joint Video Experts Team (JVET) of ITU-T WP3/16 and ISO/IEC JTC 1/‌SC 29 held its thirty-third meeting during 17–26 January 2024 as an online-only meeting. For ISO/IEC purposes, JVET is alternatively designated ISO/IEC JTC 1/‌SC 29/‌WG 5, and this was the fourteenth meeting as WG 5. The JVET meeting was held under the chairmanship of Dr Jens-Rainer Ohm (RWTH Aachen/Germany). For rapid access to particular topics in this report, a subject categorization is found (with hyperlinks) in section 2.16 of this document. It is further noted that work items which had originally been conducted by the Joint Collaborative Team on Video Coding (JCT-VC) were continued in JVET as a single joint team, and explorations towards possible future need of standardization in the area of video coding are also conducted by JVET, as negotiated by the parent bodies.

The JVET meeting began at approximately 0500 UTC on Wednesday 17 January 2024. Meeting sessions were held on all days except the weekend days of Saturday and Sunday 20 and 21 January 2024, until the meeting was closed at approximately XXXX hours UTC on Friday 26 January 2024. Approximately XXX people attended the JVET meeting, and approximately XXX input documents (not counting crosschecks, reports, and summary documents), 16 AHG reports, 2 EE summary reports, X BoG reports, and X incoming liaison document(s) were discussed. The meeting took place in coordination with a meeting of various SC29 Working Groups and Advisory Groups – where WG 5 is representing the Joint Video Coding Team(s) and their activities from the perspective of the SC 29 parent body, under whose auspices this JVET meeting was held. The subject matter of the JVET meeting activities consisted of work on further development and maintenance of the twin-text video coding technology standards *Advanced Video Coding* (AVC), *High Efficiency Video Coding* (HEVC), *Versatile Video Coding* (VVC)*, Coding-independent Code Points (Video)* (CICP), and *Versatile Supplemental Enhancement Information Messages for Coded Video Bitstreams* (VSEI), as well as related technical reports, reference software and conformance testing packages. Further important goals were reviewing the results of the Exploration Experiment (EE) on Neural Network-based Video Coding, of the EE on Enhanced Compression beyond VVC capability, of other technical input on novel aspects of video coding technology, and to plan next steps for investigation of candidate technology towards further standard development.

As a primary goal, the JVET meeting reviewed the work that had been performed in the interim period since the thirty-second JVET meeting in producing the following documents:

a) JVET documents

* [JVET-AF1004](https://jvet-experts.org/doc_end_user/current_document.php?id=12567) Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP
* [JVET-AF1006](https://jvet-experts.org/doc_end_user/current_document.php?id=12568) New profiles, colour decriptors, and SEI messages for HEVC (draft 2), also issued as WG 5 DAM N 244
* [JVET-AF1016](https://jvet-experts.org/doc_end_user/current_document.php?id=12568) AVC with extensions and corrections (draft 2), also issued as WG 5 DIS of AVC 11th edition N 241
* [JVET-AF2002](https://jvet-experts.org/doc_end_user/current_document.php?id=12574) Algorithm description for Versatile Video Coding and Test Model 21 (VTM 21), also issued as WG 5 N 245
* [JVET-AF2016](https://jvet-experts.org/doc_end_user/current_document.php?id=12576) Common test conditions and evaluation procedures for neural network-based video coding technology
* [JVET-AF2017](https://jvet-experts.org/doc_end_user/current_document.php?id=12576) Common test conditions and evaluation procedures for enhanced compression tool testing
* [JVET-AF2019](https://jvet-experts.org/doc_end_user/current_document.php?id=12563) Description of algorithms and software in neural network-based video coding (NNVC) version 5, also issued as WG 5 N 248
* [JVET-AF2021](https://jvet-experts.org/doc_end_user/current_document.php?id=12578) Verification test plan for VVC multilayer coding (update 2)
* [JVET-AF2023](https://jvet-experts.org/doc_end_user/current_document.php?id=12560) Exploration experiment on neural network-based video coding (EE1), also issued as WG 5 N 247
* [JVET-AF2024](https://jvet-experts.org/doc_end_user/current_document.php?id=12562) Exploration experiment on enhanced compression beyond VVC capability (EE2), also issued as WG 5 N 249
* [JVET-AF2025](https://jvet-experts.org/doc_end_user/current_document.php?id=12580) Algorithm description of Enhanced Compression Model 11 (ECM 11), also issued as WG 5 N 250
* [JVET-AF2027](https://jvet-experts.org/doc_end_user/current_document.php?id=12582) SEI processing order and processing order nesting SEI messages in VVC (Draft 6), also issued as WG 5 preliminary WD N 246
* [JVET-AF2031](https://jvet-experts.org/doc_end_user/current_document.php?id=12584) Common test conditions for optimization of encoders and receiving systems for machine analysis of coded video content
* [JVET-AF2032](https://jvet-experts.org/doc_end_user/current_document.php?id=12584) Technologies under consideration for future extensions of VSEI (draft 2), also issued as WG 5 N 242
* [JVET-AF2033](https://jvet-experts.org/doc_end_user/current_document.php?id=12584) Report of verification test on VVC multi-layer coding: Content layering, also issued as AG 5 N 105

b) documents produced as WG 5 documents only:

* WG 5 N 240 Disposition of comments received on ISO/IEC CD 14496-10:202x
* WG 5 N 243 Disposition of comments received on ISO/IEC 23008-2:202x (5th ed.) CDAM 1
* WG 5 N 251 Liaison statement to ISO/IEC JTC 1/SC 29/WG 1 (JPEG) on JPEG AI and explorations on video coding
* WG 5 N 252 Liaison response to 3GPP on feasibility study on film grain synthesis

As main results, the JVET produced XX output documents from the current meeting (update after meeting)

* [JVET-AF1004](https://jvet-experts.org/doc_end_user/current_document.php?id=12567) Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP
* [JVET-AF1006](https://jvet-experts.org/doc_end_user/current_document.php?id=12568) New profiles, colour decriptors, and SEI messages for HEVC (draft 2), also issued as WG 5 DAM N 244
* [JVET-AF1016](https://jvet-experts.org/doc_end_user/current_document.php?id=12568) AVC with extensions and corrections (draft 2), also issued as WG 5 DIS of AVC 11th edition N 241
* [JVET-AF2002](https://jvet-experts.org/doc_end_user/current_document.php?id=12574) Algorithm description for Versatile Video Coding and Test Model 21 (VTM 21), also issued as WG 5 N 245
* [JVET-AF2016](https://jvet-experts.org/doc_end_user/current_document.php?id=12576) Common test conditions and evaluation procedures for neural network-based video coding technology
* [JVET-AF2017](https://jvet-experts.org/doc_end_user/current_document.php?id=12576) Common test conditions and evaluation procedures for enhanced compression tool testing
* [JVET-AF2019](https://jvet-experts.org/doc_end_user/current_document.php?id=12563) Description of algorithms and software in neural network-based video coding (NNVC) version 5, also issued as WG 5 N 248
* [JVET-AF2021](https://jvet-experts.org/doc_end_user/current_document.php?id=12578) Verification test plan for VVC multilayer coding (update 2)
* [JVET-AF2023](https://jvet-experts.org/doc_end_user/current_document.php?id=12560) Exploration experiment on neural network-based video coding (EE1), also issued as WG 5 N 247
* [JVET-AF2024](https://jvet-experts.org/doc_end_user/current_document.php?id=12562) Exploration experiment on enhanced compression beyond VVC capability (EE2), also issued as WG 5 N 249
* [JVET-AF2025](https://jvet-experts.org/doc_end_user/current_document.php?id=12580) Algorithm description of Enhanced Compression Model 11 (ECM 11), also issued as WG 5 N 250
* [JVET-AF2027](https://jvet-experts.org/doc_end_user/current_document.php?id=12582) SEI processing order and processing order nesting SEI messages in VVC (Draft 6), also issued as WG 5 preliminary WD N 246
* [JVET-AF2031](https://jvet-experts.org/doc_end_user/current_document.php?id=12584) Common test conditions for optimization of encoders and receiving systems for machine analysis of coded video content
* [JVET-AF2032](https://jvet-experts.org/doc_end_user/current_document.php?id=12584) Technologies under consideration for future extensions of VSEI (draft 2), also issued as WG 5 N242
* [JVET-AF2033](https://jvet-experts.org/doc_end_user/current_document.php?id=12584) Report of verification test on VVC multi-layer coding: Content layering

The following 4 documents were produced as WG 5 documents only, without a corresponding JVET output document or direct repetition of their content in this meeting report:

* WG 5 N 240 Disposition of comments received on ISO/IEC CD 14496-10:202x
* WG 5 N 243 Disposition of comments received on ISO/IEC 23008-2:202x (5th ed.) CDAM 1
* WG 5 N 251 Liaison statement to ISO/IEC JTC 1/SC 29/WG 1 (JPEG) on JPEG AI and explorations on video coding
* WG 5 N 252 Liaison response to 3GPP on feasibility study on film grain synthesis

For the organization and planning of its future work, the JVET established 16 “ad hoc groups” (AHGs) to progress the work on particular subject areas. At this meeting, 2 Exploration Experiments (EE) were defined. The next eight JVET meetings were planned for 17 – 24 April 2024 under ITU-T SG16 auspices in Rennes, FR; during 12 – 19 July 2024 under ISO/IEC JTC 1/‌SC 29 auspices in Sapporo, JP; during 1 – 8 November 2024 under ISO/IEC JTC 1/‌SC 29 auspices, in Antalya, TR; during 14 – 22 January 2025 under ITU-T SG16 auspices in Geneva, CH; during 26 March – 4 April 2025 under ISO/IEC JTC 1/‌SC 29 auspices, to be conducted as teleconference meeting; during 26 June – 4 July 2025 under ISO/IEC JTC 1/‌SC 29 auspices in Daejeon, KR; during October 2025 under ITU-T SG 16 auspices, date and location t.b.d.; and during January 2026 under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.

The document distribution site <https://jvet-experts.org/> was used for distribution of all documents. It was noted that the previous sites <http://phenix.int-evry.fr/jvet/>, <http://phenix.int-evry.fr/jct/>, and <http://phenix.int-evry.fr/jct3v/> are still accessible, but were converted to read-only.

The reflector to be used for discussions by the JVET and all its AHGs is the JVET reflector:  
[jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de) hosted at RWTH Aachen University. For subscription to this list, see <https://lists.rwth-aachen.de/postorius/lists/jvet.lists.rwth-aachen.de/>.

# Administrative topics

## Organization

The ITU-T/ISO/IEC Joint Video Experts Team (JVET) is a group of video coding experts from the ITU-T Study Group 16 Visual Coding Experts Group (VCEG) and ISO/IEC JTC 1/‌SC 29/‌WG 5. The parent bodies of the JVET are ITU-T WP3/16 and ISO/IEC JTC 1/‌SC 29.

The Joint Video Experts Team (JVET) of ITU-T WP3/16 and ISO/IEC JTC 1/‌SC 29 held its thirty-third meeting during 17–26 January 2023 as an online-only meeting. For ISO/IEC purposes, JVET is alternatively designated ISO/IEC JTC 1/‌SC 29/‌WG 5, and this was the fourteenth meeting as WG 5. The JVET meeting was held under the chairmanship of Dr Jens-Rainer Ohm (RWTH Aachen/Germany).

It is further noted that the unabbreviated name of JVET was formerly known as “Joint Video *Exploration* Team”, but the parent bodies modified it when entering the phase of formal development of the *Versatile Video Coding* (VVC) and *Versatile Supplemental Enhancement Information Messages for Coded Video Bitstreams* (VSEI) standards, as well as associated conformance test sets, reference software, verification testing, and non-normative guidance information. Furthermore, starting from the twentieth meeting, work items which had originally been conducted by the Joint Collaborative Team on Video Coding (JCT-VC) were continued to be conducted in JVET as a single joint team, as negotiated by the parent bodies. This particularly consists of work on:

* *High Efficiency Video Coding* (HEVC) and its extensions, the development of associated conformance test sets, reference software, verification testing, and non-normative guidance information,
* Specification of *Coding-independent Code Points (Video)* (CICP), and associated technical report(s),
* Maintenance and enhancement work on the *Advanced Video Coding* (AVC) standard, associated conformance test sets and reference software.

Furthermore, explorations towards possible future need of standardization in the area of video coding are also conducted by JVET. Currently, the following topics are under investigation:

* Exploration on Neural Network-based Video Coding
* Exploration on Enhanced Compression beyond VVC capability

This report contains three important annexes, as follows:

* Annex A contains a list of the documents of the JVET meeting
* Annex B contains a list of the meeting participants, consisting of two parts, (B1) in-person attendees as recorded by a sign-in sheet circulated in meeting rooms, (B2) remote attendees as recorded by the teleconferencing tool used for the meeting
* Annex C contains the meeting recommendations of ISO/IEC JTC 1/‌SC 29/‌WG 5 for purposes of results reporting to ISO/IEC.

## Meeting logistics

Information regarding logistics arrangements for the meeting had been provided via the email reflector [jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de) and at <http://wftp3.itu.int/av-arch/jvet-site/2024_01_AG_Virtual/>.

## Primary goals

As a primary goal, the JVET meeting reviewed the work that was performed in the interim period since the thirty-first JVET meeting in producing the following output documents:

a) JVET documents

* [JVET-AF1004](https://jvet-experts.org/doc_end_user/current_document.php?id=12567) Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP
* [JVET-AF1006](https://jvet-experts.org/doc_end_user/current_document.php?id=12568) New profiles, colour decriptors, and SEI messages for HEVC (draft 2), also issued as WG 5 DAM N 244
* [JVET-AF1016](https://jvet-experts.org/doc_end_user/current_document.php?id=12568) AVC with extensions and corrections (draft 2), also issued as WG 5 DIS of AVC 11th edition N 241
* [JVET-AF2002](https://jvet-experts.org/doc_end_user/current_document.php?id=12574) Algorithm description for Versatile Video Coding and Test Model 21 (VTM 21), also issued as WG 5 N 245
* [JVET-AF2016](https://jvet-experts.org/doc_end_user/current_document.php?id=12576) Common test conditions and evaluation procedures for neural network-based video coding technology
* [JVET-AF2017](https://jvet-experts.org/doc_end_user/current_document.php?id=12576) Common test conditions and evaluation procedures for enhanced compression tool testing
* [JVET-AF2019](https://jvet-experts.org/doc_end_user/current_document.php?id=12563) Description of algorithms and software in neural network-based video coding (NNVC) version 5, also issued as WG 5 N 248
* [JVET-AF2021](https://jvet-experts.org/doc_end_user/current_document.php?id=12578) Verification test plan for VVC multilayer coding (update 2)
* [JVET-AF2023](https://jvet-experts.org/doc_end_user/current_document.php?id=12560) Exploration experiment on neural network-based video coding (EE1), also issued as WG 5 N 247
* [JVET-AF2024](https://jvet-experts.org/doc_end_user/current_document.php?id=12562) Exploration experiment on enhanced compression beyond VVC capability (EE2), also issued as WG 5 N 249
* [JVET-AF2025](https://jvet-experts.org/doc_end_user/current_document.php?id=12580) Algorithm description of Enhanced Compression Model 11 (ECM 11), also issued as WG 5 N 250
* [JVET-AF2027](https://jvet-experts.org/doc_end_user/current_document.php?id=12582) SEI processing order and processing order nesting SEI messages in VVC (Draft 6), also issued as WG 5 preliminary WD N 246
* [JVET-AF2031](https://jvet-experts.org/doc_end_user/current_document.php?id=12584) Common test conditions for optimization of encoders and receiving systems for machine analysis of coded video content
* [JVET-AF2032](https://jvet-experts.org/doc_end_user/current_document.php?id=12584) Technologies under consideration for future extensions of VSEI (draft 2), also issued as WG 5 N242
* [JVET-AF2033](https://jvet-experts.org/doc_end_user/current_document.php?id=12584) Report of verification test on VVC multi-layer coding: Content layering, also issued as AG 5 N 105

b) documents produced as WG 5 documents only:

* WG 5 N 240 Disposition of comments received on ISO/IEC CD 14496-10:202x
* WG 5 N 243 Disposition of comments received on ISO/IEC 23008-2:202x (5th ed.) CDAM 1
* WG 5 N 251 Liaison statement to ISO/IEC JTC 1/SC 29/WG 1 (JPEG) on JPEG AI and explorations on video coding
* WG 5 N 252 Liaison response to 3GPP on feasibility study on film grain synthesis

Further important goals were reviewing the results of the EE on Neural Network-based Video Coding, of the EE on Enhanced Compression beyond VVC capability, of other technical input on novel aspects of video coding technology, and planning next steps for investigation of candidate technology towards further standard development.

## Documents and document handling considerations

### General

The document distribution site <https://jvet-experts.org/> was used for distribution of all documents. It was noted that the previous site <http://phenix.int-evry.fr/jvet/> was still accessible, but had been converted to read-only.

Document registration timestamps, initial upload timestamps, and final upload timestamps are listed in Annex A of this report.

The document registration and upload times and dates listed in Annex A and in headings for documents in this report are in Paris/Geneva time. Dates mentioned for purposes of describing events at the meeting follow the CEST timezone (local time in Geneva), except as otherwise noted.

Highlighting of recorded decisions in this report is practised as follows:

* Decisions made by the group that might affect the normative content of a future standard are identified in this report by prefixing the description of the decision with the string “Decision:”.
* Decisions that affect one of the various software packages but have no normative effect on text are marked by the string “Decision (SW):”.
* Decisions that fix a “bug” in one of the test model descriptions such as VTM, HM, etc. (an error, oversight, or messiness) or in the associated software package are marked by the string “Decision (BF):”.
* Decisions that are merely editorial without effect on the technical content of a draft standard are marked by the string "Decision (Ed.):". Such editorial decisions are merely suggestions to the editor, who has the discretion to determine the final action taken if their judgment differs.
* Other parenthetical comments may be used for describing the impact or motivation of a decision. Some decisions are recorded with the word “agreed” rather than “Decision:”, especially for work items under study, non-normative, editorial and planning matters.

This meeting report is based primarily on notes taken by the JVET chair. The preliminary notes were also circulated publicly by ftp and http during the meeting on a daily basis. It should be understood by the reader that 1) some notes may appear in abbreviated form, 2) summaries of the content of contributions are often based on abstracts provided by contributing proponents without an intent to imply endorsement of the views expressed therein, and 3) the depth of discussion of the content of the various contributions in this report is not uniform. Generally, the report is written to include as much information about the contributions and discussions as is feasible (in the interest of aiding study), although this approach may not result in the most polished output report. Expressions such as “X.XX%” indicate that the desired results were not available at the time the information was recorded.

### Late and incomplete document considerations

The formal deadline for registering and uploading non-administrative contributions had been announced as Wednesday, 10 January 2024. Any documents uploaded after 1159 hours Paris/Geneva time on Thursday 11 January 2024 were considered “officially late”, with a grace period of 12 hours (to accommodate those living in different time zones of the world). The deadline does not apply to AHG reports and other such reports which can only be produced after the availability of other input documents.

All contribution documents with registration numbers higher than JVET-AG0215 were registered after the “officially late” deadline (and therefore were also uploaded late). However, some documents in the “late” range might include break-out activity reports that were generated during the meeting, or documents which were requested to be produced for the purpose of improving specification text, and are therefore better considered as report documents rather than as late contributions.

In many cases, contributions were also revised after the initial version was uploaded. The contribution document archive website retains publicly accessible prior versions in such cases. The timing of late document availability for contributions is generally noted in the section discussing each contribution in this report.

One suggestion to assist with the issue of late submissions has been to require the submitters of late contributions and late revisions to describe the characteristics of the late or revised (or missing) material at the beginning of discussion of the contribution. This has been agreed to be a helpful approach to be followed at the meeting.

The following technical design proposal contributions were registered and/or uploaded late:

* JVET-AG0XXX (a proposal on …), uploaded 01-XX,
* … .

It may be observed that some of the above-listed contributions were submissions made in response to issues that arose in discussions during the meeting or from the study of other contributions, and thus could not have been submitted by the ordinary deadline.

The following other documents not proposing normative technical content, but with some need for consideration, were registered and/or uploaded late:

* JVET-AG0XXX (a document presenting …), uploaded 01-XX,
* … .

All cross-verification reports at this meeting (except for JVET-AG0119) were registered late, and/or uploaded late. In the interest of brevity, these are not specifically identified here. Initial upload times for each document are recorded in Annex A of this report.

The following contribution registrations were noted that were later cancelled, withdrawn, never provided, were cross-checks of a withdrawn contribution, or were registered in error: JVET-AG0050, JVET-AG0181, JVET-AG0189, JVET-AG0190, JVET-AG0218, JVET-AG0238, JVET-AG0239, ….

The following cross-verification reports were still missing three weeks after the end of the meeting: JVET-AG0XXX, …. This was marked as withdrawn by the JVET chair, assuming the registration had become obsolete.

“Placeholder” contribution documents that were basically empty of content, or lacking any results showing benefit for the proposed technology, and obviously uploaded with an intent to provide a more complete submission as a revision, had been agreed to be considered unacceptable and to be rejected in the document management system until a more complete version was available (which would then be counted as a late contribution if the update was after the document deadline). At the current meeting, this situation did apply to documents JVET-AG0XXX and …, which were also categorized as late in the list above, based on the time of the first reasonable document upload; this sentence is kept for future use.

Contributions that had significant problems with uploaded versions were not observed at this meeting.

It was remarked that documents that are substantially revised after the initial upload can also be a problem, as this becomes confusing, interferes with study, and puts an extra burden on synchronization of the discussion. This can especially be a problem in cases where the initial upload is clearly incomplete, and in cases where it is difficult to figure out what parts were changed in a revision. For document contributions, revision marking is very helpful to indicate what has been changed. Also, the “comments” field on the web site can be used to indicate what is different in a revision, although participants tend to seldom notice what is recorded there.

As a general policy, missing documents were not to be presented, and late documents (and substantial revisions) could only be presented when there was a consensus to consider them and there was sufficient time available for their review. Again, an exception is applied for AHG reports, CE and HLS topic summaries, and other such reports which can only be produced after the availability of other input documents. There were no objections raised by the group regarding presentation of late contributions, although there may have been some expression of annoyance and remarks on the difficulty of dealing with late contributions and late revisions.

A few contributions may have had some problems relating to IPR declarations in the initial uploaded versions (missing declarations, declarations saying they were from the wrong companies, etc.). Any such issues were corrected by later uploaded versions in a reasonably timely fashion in all cases (to the extent of the awareness of the responsible coordinators).

Some other errors may have also noticed in other initial document uploads (wrong document numbers or meeting dates or meeting locations in headers, etc.) which were generally sorted out in a reasonably timely fashion. The document web site contains an archive of each upload.

### Outputs of the preceding meeting

All output documents of the previous meeting, particularly the meeting report JVET-AF1000, the Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP JVET-AF1004, the New profiles, colour descriptors, and SEI messages for HEVC (draft 2) JVET-AF1006, the AVC specification with extensions and corrections (draft 2) JVET-AF1016, the Algorithm description for Versatile Video Coding and Test Model 21 (VTM 21) JVET-AF2002, the Common test conditions and evaluation procedures for neural network-based video coding technology JVET-AF2016, the Common test conditions and evaluation procedures for enhanced compression tool testing JVET-AF2017, the Description of algorithms and software in neural network-based video coding (NNVC) version 5 JVET-AF2019, the Verification test plan for VVC multilayer coding (update 2) JVET-AF2021, the Description of the EE on Neural Network-based Video Coding JVET-AF2023, the Description of the EE on Enhanced Compression beyond VVC capability JVET-AF2024, the Algorithm description of Enhanced Compression Model 11 (ECM 11) JVET-AF2025, the SEI processing order SEI message in VVC (Draft 6) JVET-AF2027, the Common test conditions for optimization of encoders and receiving systems for machine analysis of coded video content [JVET-AF2031](https://jvet-experts.org/doc_end_user/current_document.php?id=12584), the Technologies under consideration for future extensions of VSEI (draft 2) JVET-AF2032, and the Report of verification test on VVC multi-layer coding: Content layering JVEZT-AF2033, had been completed and were approved. In a few cases, the corresponding WG 5 N-numbered documents had not yet been uploaded, and this was requested to be done as soon as possible. The software implementations of VTM versions 22.2 and 23.0, ECM versions 10.1 and 11.0, and NNVC (versions 7.0 and 7.1) were also approved.

Only minor editorial issues were found in the meeting report JVET-AF1000; no need to produce an update was identified (see section 2.14 for details).

The available output documents of the previous meeting and the software had been made available in a reasonably timely fashion.

## Attendance

The list of participants in the JVET meeting can be found in Annex B of this report.

The meeting was open to those qualified to participate either in ITU-T WP3/16 or ISO/IEC JTC 1/‌SC 29/‌WG 5 (including experts who had been personally invited as permitted by ITU-T or ISO/IEC policies).

Participants had been reminded of the need to be properly qualified to attend. Those seeking further information regarding qualifications to attend future meetings may contact the responsible coordinators.

It was further announced that it is necessary to register for the meeting through the ISO Meetings website for ISO/IEC experts or through the Q6/16 rapporteur for ITU-T experts. The password for meeting access had been sent to registered participants via these channels. Links to the Zoom sessions (without the necessary password) were available in the posted meeting logistics information and the calendar of meeting sessions in the JVET web site.

The following rules were established for those participating remotely via Zoom teleconference meeting:

* Use the “hand-raising” function to enter yourself in the queue to speak (unless otherwise instructed by the session chair). If you are dialed in by phone, request your queue position verbally. The online queue will be interleaved with the room queue, though it may not always be guaranteed that the sequence perfectly follows the sequence by which hand raising occurred.
* Stay muted unless you have something to say. People are muted by default when they join and need to unmute themselves to speak. The chair may mute anyone who is disrupting the proceedings (e.g. by forgetting they have a live microphone while chatting with their family or by causing bad noise or echo).
* Identify who you are and your affiliation when you begin speaking. The same applies for speakers in the room to let online participants know who is speaking.
* Use your full name and company/organization and country affiliation in your joining information, since the participation list of Zoom would also be used to compile the online part of attendance records.
* Turn on the chat window and watch for chair communication and side commentary there as well as by audio.
* Generally do not use video for the teleconferencing calls in order to avoid overloading internet connections; enable only voice and screen sharing.
* Extensive use of screen sharing is encouraged, to enable participants to view the presented material and the meeting notes. At times, multiple sources of screen sharing may be enabled, so it may be necessary for participants to understand that this is happening and to understand how to select which shared screen they want to see.

## Agenda

The agenda for the meeting, for the further development and maintenance of the twin-text video coding technology standards *Advanced Video Coding* (AVC), *High Efficiency Video Coding* (HEVC), *Versatile Video Coding* (VVC)*, Coding-independent Code Points (Video)* (CICP), and *Versatile Supplemental Enhancement Information Messages for Coded Video Bitstreams* (VSEI), as well as related technical reports, software and conformance packages, was as follows:

* Opening remarks and review of meeting logistics and communication practices
* Roll call of participants
* Adoption of the agenda
* Code of conduct policy reminder
* IPR policy reminder and declarations
* Contribution document allocation
* Review of results of the previous meeting
* Review of target dates
* Reports of ad hoc group (AHG) activities
* Report of exploration experiments on neural-network-based video coding
* Report of exploration experiments on enhanced compression beyond VVC capability
* Consideration of contributions on high-level syntax
* Consideration of contributions and communications on project guidance
* Consideration of video coding technology contributions
* Consideration of contributions on conformance and reference software development
* Consideration of contributions on coding-independent code points for video signal type identification
* Consideration of contributions on film grain synthesis technology
* Consideration of contributions on optimization of encoders and receiving systems for machine analysis of coded video content
* Consideration of contributions on errata relating to standards in the domain of JVET
* Consideration of contributions on technical reports relating to standards and exploration study activities in the domain of JVET
* Consideration of contributions providing non-normative guidance relating to standards and exploration study activities in the domain of JVET
* Consideration of information contributions
* Consideration of future work items
* Coordination of visual quality testing
* Liaisons, coordination activities with other organizations
* Review of project editor and liaison assignments
* Approval of output documents and associated editing periods
* Future planning: Determination of next steps, discussion of working methods, communication practices, establishment of coordinated experiments (if any), establishment of AHGs, future meeting planning, other planning issues
* Other business as appropriate for consideration
* Closing

The agenda was approved as suggested.

The plans for the times of meeting sessions were established as follows, in UTC (which for this meeting was 1 hour behind the time in Geneva and Paris; 8 hours ahead of the time in Los Angeles, etc.). No session was scheduled to last longer than 2 hrs.

* 0500–0700 1st “morning” session [break after 2 hours]
* 0720–0920 2nd “morning” session
* [“overday” break – nearly 12 hours]
* 2100–2300 1st “night” session [break after 2 hours]
* 2320–0120+1 2nd “night” session

It was also pointed out that the session times had been changed from meeting to meeting, such that different time zones of the world might be treated approximately equally fairly either in one meeting or another. For the current meeting, the same UTC session times were used as in the 26th JVET meeting (which had been the ninth meeting conducted as an online-only meeting, whereas the current one is the twelfth).

## ISO and IEC Code of Conduct reminders

Participants were reminded of the ISO and IEC Codes of Conduct, found at (check for new version)

<https://www.iso.org/publication/PUB100011.html>.

<https://www.iecapc.jp/F/IEC_Code_of_Conduct.pdf>

These include points relating to:

* Complying with legal and statutory obligations
* Performing and acting in good faith, consistent with the purpose, policies and principles of the organization
* Behaving ethically
* Promoting and enabling all voices to be heard
* Engaging constructively in ISO and IEC activities
* Declaring actual and potential conflicts of interest and managing them appropriately
* Protecting confidential information
* Protecting ISO and IEC assets
* Avoiding and preventing any form of bribery or corruption
* Escalating and resolving disputes and upholding agreed resolutions

## IPR policy reminder

Participants were reminded of the IPR policy established by the parent organizations of the JVET and were referred to the parent body websites for further information. The IPR policy was summarized for the participants.

The ITU-T/ITU-R/ISO/IEC common patent policy shall apply. Participants were particularly reminded that contributions proposing normative technical content shall contain a non-binding informal notice of whether the submitter may have patent rights that would be necessary for implementation of the resulting standard. The notice shall indicate the category of anticipated licensing terms according to the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form.

This obligation is supplemental to, and does not replace, any existing obligations of parties to submit formal IPR declarations to ITU-T/ITU-R/ISO/IEC.

Participants were also reminded of the need to formally report patent rights to the top-level parent bodies (using the common reporting form found on the database listed below) and to make verbal and/or document IPR reports within the JVET necessary in the event that they are aware of unreported patents that are essential to implementation of a standard or of a draft standard under development.

Some relevant links for organizational and IPR policy information are provided below:

* <http://www.itu.int/ITU-T/ipr/index.html> (common patent policy for ITU-T, ITU-R, ISO, and IEC, and guidelines and forms for formal reporting to the parent bodies)
* <http://ftp3.itu.int/av-arch/jvet-site> (JVET contribution templates)
* <http://www.itu.int/ITU-T/dbase/patent/index.html> (ITU-T IPR database)

The responsible coordinators invited participants to make any necessary verbal reports of previously-unreported IPR in technology that might be considered as prospective candidate for inclusion in future standards, and opened the floor for such reports: No such verbal reports were made.

## Software copyright disclaimer header reminder

It was noted that the VTM and ECM software implementation packages use the same software copyright license header as the HEVC reference software, where the latter had been agreed at the 5th meeting of the JCT-VC and approved by both parent bodies at their collocated meetings at that time. This license header language is based on the BSD license with a preceding sentence declaring that other contributor or third party rights, including patent rights, are not granted by the license, as recorded in [N 10791](http://phenix.it-sudparis.eu/mpeg/doc_end_user/current_document.php?id=27881&id_meeting=16) of the 89th meeting of ISO/IEC JTC 1/‌SC 29/‌WG 11 of June/July 2009. Both ITU and ISO/IEC will be identified in the <OWNER> and <ORGANIZATION> tags in the header. This software header is also used in the process of designing the VTM and ECM software, and for evaluating proposals for technology to be potentially included in these designs. This software or parts thereof might be published by ITU-T and ISO/IEC as an example implementation of a future video coding standard and for use as the basis of products to promote adoption of such technology.

Different copyright statements shall not be committed to the committee software repository (in the absence of subsequent review and approval of any such actions). As noted previously, it must be further understood that any initially-adopted such copyright header statement language could further change in response to new information and guidance on the subject in the future.

These considerations apply to the 360Lib video conversion software and HDRTools as well. The SADL and NNVC packages for neural network-based video coding use the same licensing terms.

Software packages that had been developed in prior work of the JVT, the JCT-VC and the JCT-3V have similar considerations and are maintained according to the past practice in that work.

## Communication practices

The documents for the meeting can be found at <https://jvet-experts.org/>. It was noted that the previous site <http://phenix.int-evry.fr/jvet/> is still accessible, but was converted to read-only. It was reminded to send a notice to the chairs in cases of changes to document titles, authors, etc.

JVET email lists are managed through the site <https://lists.rwth-aachen.de/postorius/lists/jvet.lists.rwth-aachen.de/>, and to send email to the reflector, the email address is [jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de). Only members of the reflector can send email to the list. However, membership of the reflector is not limited to qualified JVET participants.

It was emphasized that reflector subscriptions and email sent to the reflector must use real names when subscribing and sending messages and subscribers must respond to inquiries regarding the nature of their interest in the work. The current number of subscribers on the JVET email list was 1241 (as of 16 January 2024). All discussions (including those on AVC, HEVC, VVC, CICP, etc.) shall be conducted on the JVET reflector rather than any of the old reflectors (including JVT, JCT-VC, and JCT-3V) which are retained for archiving purposes.

For distribution of test sequences, a password-protected ftp site had been set up at RWTH Aachen University, with a mirror site at FhG-HHI. Accredited members of JVET may contact the responsible JVET coordinators to obtain the password information (but the site is not open for use by others).

It is further emphasized that the document JVET-AD1012 (check if recommendation to keep email addresses updated is included) gives valuable hints about communication practices as well as other IT resources used in JVET, such as software, conformance, and test materials.

## Terminology

* **ACT**: Adaptive colour transform
* **AFF**: Adaptive frame-field
* **AI**: All-intra
* **AIF**: Adaptive interpolation filtering
* **ALF**: Adaptive loop filter
* **AMP**: Asymmetric motion partitioning – a motion prediction partitioning for which the sub-regions of a region are not equal in size (in HEVC, being N/2x2N and 3N/2x2N or 2NxN/2 and 2Nx3N/2 with 2N equal to 16 or 32 for the luma component)
* **AMVP**: Adaptive motion vector prediction
* **AMT or MTS**: Adaptive multi-core transform, or multiple transform selection
* **AMVR**: (Locally) adaptive motion vector resolution
* **APS**: Adaptation parameter set
* **ARC**: Adaptive resolution conversion (synonymous with DRC, and a form of RPR)
* **ARMC**: Adaptive re-ordering of merge candidates
* **ARSS**: Adaptive reference sample smoothing
* **ATM**: AVC-based multiview and 3D test model
* **ATMVP** or “subblock-based temporal merging candidates”: Alternative temporal motion vector prediction
* **AU**: Access unit
* **AUD**: Access unit delimiter
* **AVC**: Advanced video coding – the video coding standard formally published as ITU-T Recommendation H.264 and ISO/IEC 14496-10
* **BA**: Block adaptive
* **BC**: See CPR or IBC
* **BCW**: Biprediction with CU based weighting
* **BD**: Bjøntegaard-delta – a method for measuring percentage bit rate savings at equal PSNR or decibels of PSNR benefit at equal bit rate (e.g., as described in document VCEG-M33 of April 2001)
* **BDOF**: Bi-directional optical flow (formerly known as **BIO**)
* **BDPCM**: Block-wise DPCM
* **BL**: Base layer
* **BMS**: Benchmark set (no longer used), a former preliminary compilation of coding tools on top of VTM, which provide somewhat better compression performance, but are not deemed mature for standardzation
* **BoG**: Break-out group
* **BR**: Bit rate
* **BT**: Binary tree
* **BV**: Block vector (used for intra BC prediction)
* **CABAC**: Context-adaptive binary arithmetic coding
* **CBF**: Coded block flag(s)
* **CC**: May refer to context-coded, common (test) conditions, or cross-component
* **CCALF**: Cross-component ALF
* **CCLM**: Cross-component linear model
* **CCCM**: Cross-component convolutional model
* **CCP**: Cross-component prediction
* **CCSAO**:Cross-component SAO
* **CE**: Core Experiment – a coordinated experiment conducted toward assessment of coding technology
* **CG**: Coefficient group
* **CGS**: Colour gamut scalability (historically, coarse-grained scalability)
* **CIIP**: Combined inter/intra prediction
* **CIPF**: CABAC initialization from the previous frame
* **CL-RAS**: Cross-layer random-access skip
* **CPB**: Coded picture buffer
* **CPMV**: Control-point motion vector
* **CPMVP**: Control-point motion vector prediction (used in affine motion model)
* **CPR**: Current-picture referencing, also known as IBC – a technique by which sample values are predicted from other samples in the same picture by means of a displacement vector called a block vector, in a manner conceptually similar to motion-compensated prediction
* **CST**: Chroma separate tree
* **CTC**: Common test conditions
* **CVS**: Coded video sequence
* **DCI**: Decoder capability information
* **DCT**: Discrete cosine transform (sometimes used loosely to refer to other transforms with conceptually similar characteristics)
* **DCTIF**: DCT-derived interpolation filter
* **DF**: Deblocking filter
* **DIMD**: Decoder intra mode derivation
* **DMVR**: Decoder motion vector refinement
* **DoCR**: Disposition of comments report
* **DPB**: Decoded picture buffer
* **DPCM**: Differential pulse-code modulation
* **DPS**: Decoding parameter sets
* **DRC**: Dynamic resolution conversion (synonymous with ARC, and a form of RPR)
* **DT**: Decoding time
* **DQ**: Dependent quantization
* **ECS**: Entropy coding synchronization (typically synonymous with WPP)
* **EMT**: Explicit multiple-core transform
* **EOTF**: Electro-optical transfer function – a function that converts a representation value to a quantity of output light (e.g., light emitted by a display
* **EPB**: Emulation prevention byte (as in the emulation\_prevention\_byte syntax element)
* **ECM**: Enhanced compression model – a software codebase for future video coding exploration
* **ECV**: Extended Colour Volume (up to WCG)
* **EL**: Enhancement layer
* **EOS**: End of (coded video) sequence
* **ET**: Encoding time
* **FRUC**: Frame rate up conversion (pattern matched motion vector derivation)
* **GCI**: General constraints information
* **GDR**: Gradual decoding refresh
* **GLM**: Gradient linear model
* **GOP**: Group of pictures (somewhat ambiguous)
* **GPM**: Geometry partitioning mode
* **GRA**: Gradual random access
* **HBD**: High bit depth
* **HDR**: High dynamic range
* **HEVC**: High Efficiency Video Coding – the video coding standard developed and extended by the JCT-VC, formalized by ITU-T as Rec. ITU-T H.265 and by ISO/IEC as ISO/IEC 23008-2
* **HLS**: High-level syntax
* **HM**: HEVC Test Model – a video coding design containing selected coding tools that conforms to the HEVC standard design (possibly with under-development extensions) – now also used especially in reference to the (non-normative) encoder algorithms (see WD and TM)
* **HMVP**: History based motion vector prediction
* **HOP**: High-complexity operating point for neural network-based filter.
* **HRD**: Hypothetical reference decoder
* **HTM**: HEVC-based multiview and 3D test model (developed by JCT-3V)
* **HyGT**: Hyper-cube Givens transform (a type of NSST)
* **IBC** (also **Intra BC**): Intra block copy, also known as CPR – a technique by which sample values are predicted from other samples in the same picture by means of a displacement vector called a block vector, in a manner conceptually similar to motion-compensated prediction
* **IBDI**: Internal bit-depth increase – a technique by which lower bit-depth (8 bits per sample) source video is encoded using higher bit-depth signal processing, ordinarily including higher bit-depth reference picture storage (ordinarily 12 bits per sample)
* **IBF**: Intra boundary filtering
* **ILP**: Inter-layer prediction (in scalable coding)
* **ILRP**: Inter-layer reference picture
* **IPCM**: Intra pulse-code modulation (similar in spirit to IPCM in AVC and HEVC)
* **IRAP**: Intra random access picture
* **ISP**: Intra subblock partitioning
* **JCCR**: Joint coding of chroma residuals
* **JCT-3V**: Joint collaborative team on 3D video (for AVC and HEVC)
* **JCT-VC**: Joint collaborative team on video coding (for HEVC)
* **JEM**: Joint exploration model – a software codebase previously used for video coding exploration
* **JM**: Joint model – the primary software codebase that has been developed for the AVC standard
* **JSVM**: Joint scalable video model – another software codebase that has been developed for the AVC standard, which includes support for scalable video coding extensions
* **JVET**: Joint video experts team (initially for VVC, later expanded)
* **JVT**: Joint video team (for AVC)
* **KLT**: Karhunen-Loève transform
* **LB** or **LDB**: Low-delay B – the variant of the LD conditions that uses B pictures
* **LD**: Low delay – one of two sets of coding conditions designed to enable interactive real-time communication, with less emphasis on ease of random access (contrast with RA). Typically refers to LB, although also applies to LP
* **LFNST**: Low-frequency non-separable transform
* **LIC**: Local illumination compensation
* **LM**: Linear model
* **LMCS**: Luma mapping with chroma scaling (formerly sometimes called “in-loop reshaping”)
* **LOP**: Low-complexity operating point for neural network-based filter.
* **LP** or **LDP**: Low-delay P – the variant of the LD conditions that uses P frames
* **LUT**: Look-up table
* **LTRP**: Long-term reference picture
* **MANE**: Media-aware network element
* **MC**: Motion compensation
* **MCP**: Motion compensated prediction
* **MCTF**: Motion compensated temporal pre-filtering
* **MDNSST**: Mode dependent non-separable secondary transform
* **MIP**: Matrix-based intra prediction
* **MMLM**: Multi-model (cross component) linear mode
* **MMVD**: Merge with MVD
* **MPEG**: Moving picture experts group (an alliance of working groups and advisory groups in ISO/IEC JTC 1/‌SC 29, one of the two parent bodies of the JVET)
* **MPM**: Most probable mode (in intra prediction)
* **MRL**: Multiple reference line intra prediction
* **MV**: Motion vector
* **MVD**: Motion vector difference
* **NAL**: Network abstraction layer
* **NNVC**: Neural network-based video coding (experimental software package)
* **NSQT**: Non-square quadtree
* **NSST**: Non-separable secondary transform
* **NUH**: NAL unit header
* **NUT**: NAL unit type (as in AVC and HEVC)
* **OBMC**: Overlapped block motion compensation (e.g., as in H.263 Annex F)
* **OETF**: Opto-electronic transfer function – a function that converts to input light (e.g., light input to a camera) to a representation value
* **OLS**: Output layer set.
* **OOTF**: Optical-to-optical transfer function – a function that converts input light (e.g. l,ight input to a camera) to output light (e.g., light emitted by a display).
* **ONNX**: Open Neural Network Exchange – a format used to convert code from common neural network software packages into SADL code.
* **operation point**: A temporal subset of an OLS.
* **PDPC**: Position-dependent (intra) prediction combination.
* **PERP**: Padded equirectangular projection (a 360° projection format).
* **PH**: Picture header.
* **PHEC**: Padded hybrid equiangular cubemap (a 360° projection format).
* **PMMVD**: Pattern-matched motion vector derivation.
* **POC**: Picture order count.
* **PoR**: Plan of record.
* **PROF**: Prediction refinement with optical flow
* **PPS**: Picture parameter set (as in AVC and HEVC).
* **PTL**: Profile/tier/level combination.
* **QM**: Quantization matrix (as in AVC and HEVC).
* **QP**: Quantization parameter (as in AVC and HEVC, sometimes confused with quantization step size).
* **QT**: Quadtree.
* **RA**: Random access – a set of coding conditions designed to enable relatively-frequent random access points in the coded video data, with less emphasis on minimization of delay (contrast with LD).
* **RADL**: Random-access decodable leading (type of picture).
* **RASL**: Random-access skipped leading (type of picture).
* **R-D**: Rate-distortion.
* **RDO**: Rate-distortion optimization.
* **RDOQ**: Rate-distortion optimized quantization.
* **RDPCM**: Residual DPCM
* **ROT**: Rotation operation for low-frequency transform coefficients.
* **RPL**: Reference picture list.
* **RPLM**: Reference picture list modification.
* **RPR**: Reference picture resampling (e.g., as in H.263 Annex P), a special case of which is also known as ARC or DRC.
* **RPS**: Reference picture set.
* **RQT**: Residual quadtree.
* **RRU**: Reduced-resolution update (e.g. as in H.263 Annex Q).
* **RVM**: Rate variation measure.
* **SADL**: Small adhoc deep learning library
* **SAO**: Sample-adaptive offset.
* **SBT**: Subblock transform.
* **SbTMVP**: Subblock based temporal motion vector prediction.
* **SCIPU**: Smallest chroma intra prediction unit.
* **SD**: Slice data; alternatively, standard-definition.
* **SDH**: Sign data hiding.
* **SDT**: Signal-dependent transform.
* **SE**: Syntax element.
* **SEI**: Supplemental enhancement information (as in AVC and HEVC).
* **SH**: Slice header.
* **SHM**: Scalable HM.
* **SHVC**: Scalable high efficiency video coding.
* **SIF**: Switchable (motion) interpolation filter.
* **SIMD**: Single instruction, multiple data.
* **SMVD**: Symmetric MVD.
* **SPS**: Sequence parameter set (as in AVC and HEVC).
* **STMVP**: Spatial-temporal motion vector prediction.
* **STRP**: Short-term reference picture.
* **STSA**: Step-wise temporal sublayer access.
* **TBA/TBD/TBP**: To be announced/determined/presented.
* **TGM**: Text and graphics with motion – a category of content that primarily contains rendered text and graphics with motion, mixed with a relatively small amount of camera-captured content.
* **TIMD**: Template-based intra mode derivation
* **TM**: Template matching.
* **TMVP**: Temporal motion vector prediction.
* **TS**: Transform skip.
* **TSRC**: Transform skip residual coding.
* **TT**: Ternary tree.
* **UCBDS**: Unrestricted center-biased diamond search.
* **UGC**: User-generated content.
* **UWP**: Unequal weight prediction.
* **VCEG**: Visual coding experts group (ITU-T Q.6/16, the relevant rapporteur group in ITU-T WP3/16, which is one of the two parent bodies of the JVET).
* **VCM**: Video coding for machines.
* **VPS**: Video parameter set – a parameter set that describes the overall characteristics of a coded video sequence – conceptually sitting above the SPS in the syntax hierarchy.
* **VQA**: Visual quality assessment.
* **VT**: Verification testing.
* **VTM**: VVC Test Model.
* **VUI**: Video usability information.
* **VVC**: Versatile Video Coding, the standardization project developed by JVET.
* **WAIP**: Wide-angle intra prediction
* **WCG**: Wide colour gamut.
* **WG**: Working group, a group of technical experts (usually used to refer to WG 11, a.k.a. MPEG).
* **WPP**: Wavefront parallel processing (usually synonymous with ECS).
* Block and unit names in HEVC:
  + **CTB**: Coding tree block (luma or chroma) – unless the format is monochrome, there are three CTBs per CTU.
  + **CTU**: Coding tree unit (containing both luma and chroma, synonymous with LCU), with a size of 16x16, 32x32, or 64x64 for the luma component.
  + **CB**: Coding block (luma or chroma), a luma or chroma block in a CU.
  + **CU**: Coding unit (containing both luma and chroma), the level at which the prediction mode, such as intra versus inter, is determined in HEVC, with a size of 2Nx2N for 2N equal to 8, 16, 32, or 64 for luma.
  + **PB**: Prediction block (luma or chroma), a luma or chroma block of a PU, the level at which the prediction information is conveyed or the level at which the prediction process is performed in HEVC.
  + **PU**: Prediction unit (containing both luma and chroma), the level of the prediction control syntax within a CU, with eight shape possibilities in HEVC:
    - **2Nx2N**: Having the full width and height of the CU.
    - **2NxN (or Nx2N)**: Having two areas that each have the full width and half the height of the CU (or having two areas that each have half the width and the full height of the CU).
    - **NxN**: Having four areas that each have half the width and half the height of the CU, with N equal to 4, 8, 16, or 32 for intra-predicted luma and N equal to 8, 16, or 32 for inter-predicted luma – a case only used when 2N×2N is the minimum CU size.
    - **N/2x2N** paired with **3N/2x2N** or **2NxN/2** paired with **2Nx3N/2**: Having two areas that are different in size – cases referred to as AMP, with 2N equal to 16 or 32 for the luma component.
  + **TB**: Transform block (luma or chroma), a luma or chroma block of a TU, with a size of 4x4, 8x8, 16x16, or 32x32.
  + **TU**: Transform unit (containing both luma and chroma), the level of the residual transform (or transform skip or palette coding) segmentation within a CU (which, when using inter prediction in HEVC, may sometimes span across multiple PU regions).
* Block and unit names in VVC:
  + **CTB**: Coding tree block (luma or chroma) – there are three CTBs per CTU in a P or B slice or in an I slice that uses a single tree, and one CTB per luma CTU and two CTBs per chroma CTU in an I slice that uses separate trees.
  + **CTU**: Coding tree unit (synonymous with LCU, containing both luma and chroma in a P or B slice or in an I slice that uses a single tree, containing only luma or only chroma in an I slice that uses separate trees), with a size of 16x16, 32x32, 64x64, or 128x128 for the luma component.
  + **CB**: Coding block, a luma or chroma block in a CU.
  + **CU**: Coding unit (containing both luma and chroma in P/B slice, containing only luma or chroma in I slice), a leaf node of a QTBT. It’s the level at which the prediction process and residual transform are performed in JEM. A CU can be square or rectangle shape.
  + **PB**: Prediction block, a luma or chroma block of a PU.
  + **PU**: Prediction unit, has the same size as a CU in the VVC context.
  + **TB**: Transform block, a luma or chroma block of a TU.
  + **TU**: Transform unit, has the same size as a CU in the VVC context.

## Standards, TRs, supplements and technical papers approval and publication status

* MPEG-2 | H.262 (coding specification is common text)
  + ITU-T H.262 V3 was approved in 2012-02; Amd.1 was approved in 2013-03 and was not published separately; it was instead incorporated directly into the V3 text and published 2013-09
  + ISO/IEC 13818-2:2013 V3 FDIS ballot closed 2012-05-08; FDAM 1 ballot closed 2013-04-12 and was not published separately; it was instead incorporated directly into the V3 text and published 2013-10
  + Conformance testing (not joint with ITU-T)
    - ISO/IEC 13818-4:2004 V2 FDIS closed 2004-08-22, published 2004-12-12; it specifies conformance testing for Part 1 (Systems), Part 2 (Video), Part 3 (Audio), and Part 7 (AAC)
    - ISO/IEC 13818-4:2004/Amd 3:2009 Level for 1080@50p/60p conformance testing
    - Cor 1:2007, Cor 2:2009, Cor 3:2012, Cor 4:2011 may also have video relevance
  + Reference software (not joint with ITU-T)
    - ISO/IEC TR 13818-5:2005 V2 FDIS closed 2005-07-24, published 2005-10; it specifies reference software for Part 1 (Systems), Part 2 (Video), Part 3 (Audio), Part 7 (AAC) and Part 11 (IPMP)
* AVC (twin text)
  + ITU-T H.264 V14 was Consented at 22nd meeting on 2021-04-30 (with annotated regions, shutter interval, and miscellaneous corrections), approved 2021-08-22, published 2021-10-13
  + ISO/IEC 14496-10:2020 (Ed. 9) FDIS ballot closed 2020-11-27, published 2020-12-15
  + ISO/IEC 14496-10:2022 (Ed. 10), had been forwarded from DIS directly for publication 2022-01-21 (with annotated regions, shutter interval, and miscellaneous corrections) with an editing period, submitted to ITTF in 2022-05 after consultation with ISO staff on format of graphics files, upgraded to “DIS approved for registration” in ISO Project system 2022-07-04, published 2022-11-07
  + Preliminary draft text for YCgCo-Re and YCgCo-Ro issued at 26th meeting, second draft including SMPTE ST 2128 issued at 28th meeting 2022-10, third draft issued at 29th meeting 2023-01, fourth draft issued at 30th meeting 2023-04, formal project requested and CD of 11th edition issued at 31st meeting 2023-07, DIS issued at 32nd meeting 2023-10, DAM ballot to close 2024-04-15, no action at the current meeting.
  + Conformance testing (twin text)
    - ITU-T H.264.1 V6 Approved 2016-02-13, published 2016-06-17
    - Various amendments of ISO/IEC 14496-4:2004, including:
      * ISO/IEC 14496-4:2004/AMD 6:2005 Advanced Video Coding conformance
      * ISO/IEC 14496-4:2004/AMD 9:2006 AVC fidelity range extensions conformance
      * ISO/IEC 14496-4:2004/AMD 30:2009 Conformance testing for new profiles for professional applications
      * ISO/IEC 14496-4:2004/AMD 31:2009 Conformance testing for SVC profiles
      * ISO/IEC 14496-4:2004/AMD 38:2010 Conformance testing for Multiview Video Coding
      * ISO/IEC 14496-4:2004/AMD 41:2014 Conformance testing of MVC plus depth extension of AVC
      * ISO/IEC 14496-4:2004/AMD 42:2014 Conformance testing of Multi-Resolution Frame Compatible Stereo Coding extension of AVC
      * ISO/IEC 14496-4:2004/AMD 43:2015 3D-AVC conformance testing
      * ISO/IEC 14496-4:2004/AMD 45:2016 Conformance Testing for the Multi-resolution Frame Compatible Stereo Coding with Depth Maps Extension of AVC
  + Reference software (twin text)
    - ITU-T H.264.2 V7 Approved 2016-02-13, published 2016-05-30
    - Various amendments of ISO/IEC 14496-5:2001 have been published, including:
      * ISO/IEC 14496-5:2001/AMD 6:2005 Advanced Video Coding (AVC) and High Efficiency Advanced Audio Coding (HE AAC) reference software
      * ISO/IEC 14496-5:2001/AMD 8:2006 AVC fidelity range extensions reference software
      * ISO/IEC 14496-5:2001/AMD 15:2010 Reference software for Multiview Video Coding
      * ISO/IEC 14496-5:2001/AMD 18:2008 Reference software for new profiles for professional applications
      * ISO/IEC 14496-5:2001/AMD 19:2009 Reference software for Scalable Video Coding
      * ISO/IEC 14496-5:2001/AMD 33:2015 Reference software for MVC plus depth extension of AVC
      * ISO/IEC 14496-5:2001/AMD 34:2014 Reference software of the multi-resolution frame compatible stereo coding of AVC
      * ISO/IEC 14496-5:2001/AMD 35:2015 3D-AVC Reference software
      * ISO/IEC 14496-5:2001/AMD 39:2016 Reference software for the Multi-resolution Frame Compatible Stereo Coding with Depth Maps of AVC
      * ISO/IEC 14496-5:2001/AMD 42:2017 Reference software for the alternative depth information SEI message extension of AVC
* HEVC (twin text)
  + ITU-T H.265 V7 approved 2019-11-29, published 2020-01-10
  + ISO/IEC 23008-2:2020 (Ed. 4) FDIS ballot closed 2020-07-16, published 2020-08-27
  + ITU-T H.265 V8 Consented at the 22nd meeting (shutter interval information SEI message and miscellaneous corrections), published 2020-10-13
  + ISO/IEC 23008-2:2020/AMD 1:2021 (shutter interval information SEI message) published 2021-07-12
  + ISO/IEC 23008-2:202x (Ed. 5) began as CDAM 2 High-range levels output of 25th meeting of January 2022, CDAM ballot closed 2022-04-15, conversion to 5th edition with miscellaneous corrections planned at 26th meeting of 2022-04, text submitted for DIS ballot 2022-07-10, DIS ballot closed 2023-01-10, FDIS issued 29th meeting of 2023-01, FDIS ballot opened 2023-08-06, closed 2023-10-02, pending publication
  + ITU-T H.265 V9 Consented at 31st meeting 2023-07, approved 2023-09-13, and pre-published 2023-09, published 2023-11-24.
  + Preliminary draft HEVC text for YCgCo-Re and YCgCo-Ro issued at 26th meeting 2022-04, second draft including SMPTE ST 2128 issued at 28th meeting 2022-10, third draft at 29th meeting 2023-01, fourth draft at 30th meeting 2023-04, formal work item requested and CDAM1 issued 31st meeting 2023-07, DAM issued at 32nd meeting 2023-10, DAM ballot to close 2024-04-08, no action at the current meeting.
  + Conformance testing (twin text)
    - ITU-T H.265.1 V3 approved 2018-10-14, published 2019-01-15
    - ISO/IEC 23008-8:2018 (Ed. 2) Conformance specification for HEVC, published 2018-08-06
    - ISO/IEC 23008-8:2018/AMD 1:2019 Conformance testing for HEVC screen content coding (SCC) extensions and non-intra high throughput profiles, published 2019-10-15
  + Reference software (twin text)
    - ITU-T H.265.2 V4 approved 2016-12-22, published 2017-04-10
    - ISO/IEC 23008-5:2017 (Ed. 2) Reference software for high efficiency video coding, published 2017-03-01
    - ISO/IEC 23008-5:2017/AMD 1:2017 Reference software for screen content coding extensions, published 2017-11-09
* VVC (twin text)
  + ITU-T H.266 V1 approved 2020-08-29, published 2020-11-10
  + ISO/IEC 23090-3:2021 (Ed. 1) published 2021-02-16
  + ITU-T H.266 V2 with operation range extensions, Consented 2022-01-28, Last Call began 2022-04-01, Approved 2022-04-29, pre-published 2022-06-06, published 2022-07-12
  + ISO/IEC 23090-3:2022 (Ed. 2) with operation range extensions, approval at WG level to proceed to FDIS 2022-01-21, published 2022-09-25
  + ISO/IEC 23090-3:202x (Ed. 2) / Amd.1 New level and systems-related supplemental enhancement information, CDAM 1 issued from 26th meeting, ballot closed 2022-07-14, DAM 1 issued from 27th meeting, ballot closed 2023-01-03, FDIS issued 2023-07, pending FDIS ballot
  + ITU-T H.266 V3 Consented 2023-07, approved 2023-09-29 and pre-published 2023-09, published 2023-11-29.
  + Conformance testing (twin text)
    - ITU-T H.266.1 V1 Consented 2022-01-28, Last Call began 2022-04-01, Approved 2022-04-29, pre-published 2022-05-17, published 2022-07-12
    - ISO/IEC 23090-15:2022 V1 approval at WG level to proceed to FDIS 2022-10-15, upgraded to “DIS approved for registration” in ISO Projects system 2021-10-24, upgraded to “FDIS registered for formal approval” 2022-07-11, FDIS ballot closed 2022-11-04, published 2022-11-24
    - ISO/IEC 23090-15:2022/Amd.1 Operation range extensions – DAM 1 issued from 25th meeting 2022-01-21, upgraded to “CD approved for registration as DIS” status in ISO Projects system 2022-05-31, upgraded to “DIS registered” 2022-06-22, DAM ballot closed 2022-11-15, consolidated into FDIS 2nd edition issued as an output of the 29th meeting in January 2023 (awaiting ballot at the time of this meeting)
    - ITU-T H.266.1 V2 Consented 2023-07, approved 2023-09-13 and pre-published 2023-09, published 2023-10-19.
  + Reference software (twin text)
    - ITU-T H.266.2 V1 Consented 2022-01-28, Last Call began 2022-04-01, Approved 2022-04-29, pre-published 2022-05-17, published 2022-07-12
    - ISO/IEC 23090-16:2022 V1 approval at WG level to proceed to FDIS 2022-01-21, upgraded to “DIS approved for registration” status in ISO Projects system 2022-04-21, upgraded to “FDIS registered for formal approval” 2022-04-22, FDIS ballot initiated 2022-07-24, FDIS ballot closed 2022-09-19, published 2022-10-23
* VSEI (twin text)
  + ITU-T H.274 V1 approved 2020-08-29, published 2020-11-10
  + ISO/IEC 23002-7:2021 (Ed. 1) published 2021-01-28
  + ITU-T H.274 V2 Consented 2022-01-28, Last Call began 2022-04-01, Approved 2022-05-22 (after 1 Last Call comment and Additional Review), pre-published 2022-06-17, published 2022-07-25
  + ISO/IEC 23002-7:2022 (Ed. 2) approval at WG level to proceed to FDIS 2022-01-21, upgraded to “DIS approved for registration” status in ISO Projects system 2022-05-05 and “FDIS registered for formal approval” 2022-05-08, FDIS ballot closed 2022-09-27, published 2022-10-30
  + ISO/IEC 23002-7:202x (2nd Ed.) Amd.1 Request for new edition and CD for additional SEI messages issued at 27th meeting, ballot closed 2022-10-10, DAM registered 2022-11-13, DAM ballot closed 2022-04-06, FDIS 3rd edition issued 2023-07 (pending FDIS ballot)
  + ITU-T H.274 V3 Consent 2023-07, approved 2023-09-29, pre-published 2023-10-11, pending publication.
* CICP (twin text)
  + ISO/IEC 23091-2:2021 (Ed. 2) had been forwarded from DIS directly for publication in 2021-04 and published 2021-10-18
  + ITU-T H.273 V2 (with 4:2:0 sampling alignment and corrections for range of values for sample aspect ratio, ICTCP equations for HLG, and transfer characteristics function for sYCC of IEC 61966-2-1) Consented on 2021-04-30, Last Call closed during the 23rd meeting with approval on 2021-07-14, published 2021-09-24
  + ISO/IEC 23091-2:202x (Ed. 3) Request for new edition and CD for new edition (including YCgCo-Re and YCoCg-Ro) issued at 27th meeting, ballot closed 2022-10-10, DIS registered 2022-11-13, DIS ballot closed 2023-04-06, preliminary draft text for including SMPTE ST 2128 issued at 28th meeting, incorporated into preliminary FDIS at 30th meeting 2023-04, FDIS waiting for publication of SMPTE ST 2128.
  + ITU-T H.273 Consent 2023-07, approved 2023-09, publication waiting for publication of SMPTE ST 2128.
* Conversion and coding practices for HDR/WCG Y′CbCr 4:2:0 video with PQ transfer characteristics (twin text)
  + H.Sup15 V1, approved 2017-01-27, published 2017-04-12
  + ISO/IEC TR 23008-14:2018 published 2018-08
* Signalling, backward compatibility and display adaptation for HDR/WCG video coding (twin text)
  + H.Sup18 V1, approved 2017-10-27, published 2018-01-18
  + ISO/IEC TR 23008-15:2018 published 2018-08
* Usage of video signal type code points (twin text)
  + H.Sup19 V3 approved 2021-04-30, published 2021-06-04
  + ISO/IEC TR 23091-4 (Ed. 3) published 2021-05-23
* Working practices using objective metrics for evaluation of video coding efficiency experiments (twin text)
  + HSTP-VID-WPOM V1: approved 2020-07-03, published 2020-11
  + ISO/IEC TR 23002-8 (Ed. 1) published 2021-05-20
* Film grain synthesis technologies for video applications (twin text)
  + ISO/IEC TR 23002-9 Request for subdivision and WD 1 issued at 25th meeting 2022-01-21, WD 2 issued at 27th meeting, WD 3 issued at 28th meeting, CDTR issued at 29th meeting 2023-01, consultation period ended 2023-07-09. A DTR text was issued from the 31st meeting in July 2023 but was put on hold by ISO staff editors.
* The following freely available standards are published here in ISO/IEC:  
  <https://standards.iso.org/ittf/PubliclyAvailableStandards/index.html> as of 2023-10-09:
  + ISO/IEC 13818-4:2004 Conformance for MPEG-2
  + ISO/IEC 13818-4:2004/Amd 3:2009 Level for 1080@50p/60p conformance testing
  + ISO/IEC TR 13818-5:2005 Software simulation for MPEG-2
  + Various amendments of ISO/IEC 14496-4:2004 Conformance for AVC
  + Various amendments of ISO/IEC 14496-5:2001 Reference software for AVC
  + ISO/IEC 14496-10:2022 (Ed. 10) AVC
  + ISO/IEC 23002-7:2022 (Ed. 2) – VSEI
  + ISO/IEC 23008-2:2020 (Ed. 4) HEVC
  + ISO/IEC 23090-3:2022 (Ed. 2) VVC
  + ISO/IEC 23090-15:2022 (Ed. 1) Conformance for VVC
  + ISO/IEC 23090-16:2022 (Ed. 1) Reference software for VVC
  + ISO/IEC 23091-2:2021 (Ed. 2) Video CICP
* The following standards that have been intended by JVET to be publicly available were not available at <https://standards.iso.org/ittf/PubliclyAvailableStandards/index.html> as of 2023-05-26 and this was still the case as of 2023-10-06 and 2024-01-15. (These should be checked for previously issued requests for free availability.)
  + ISO/IEC 23008-2:2020 (Ed. 4) Amd.1:2021: Shutter interval information SEI message, published 2021-07-12 (has not been requested)
  + ISO/IEC 23008-5:2017 (Ed. 2) Reference software for high efficiency video coding, published 2017-03-01
  + ISO/IEC 23008-5:2017/AMD 1:2017 Reference software for screen content coding extensions, published 2017-11-09
  + ISO/IEC 23008-8:2018 (Ed. 2) Conformance specification for HEVC, published 2018-08, published 2018-08-06
  + ISO/IEC 23008-8:2018/AMD 1:2019 Conformance testing for HEVC screen content coding (SCC) extensions and non-intra high throughput profiles, published 2019-10-15
* It appears necessary to check if all older software and conformance packages are publicly available – it might be that it was never requested, e.g. for those that were produced by JCT-3V. This topic was left TBD until the next meeting – perhaps it would be best to compile a list of all relevant software and conformance parts of AVC, HEVC, MPEG-2 aka H.262, CICP, and request these in bulk.

## Draft standards progression status

* AVC colour type indicators for YCgCo-Re, YCgCo-Ro, and SMPTE ST 2128 (IPT-PQ-C2) and the addition of support for the neural-network post-filter characteristics, neural-network activation, and phase indication SEI messages specified in Rec. ITU-T H.274 | ISO/IEC 23002-7 are in a DIS issued at 32nd meeting 2023-10, DAM ballot to close 2024-04-15, no action at the current meeting.
* HEVC 23008-2:202x (5th ed.) DAM1 New profiles, colour descriptors, and SEI messages, with colour type indicators for YCgCo-Re, YCgCo-Ro, and SMPTE ST 2128 (IPT-PQ-C2), the specification of a Multiview Main 10 profile, and the addition of support for the neural-network post-filter characteristics, neural-network activation, and phase indication SEI messages specified in Rec. ITU-T H.274 | ISO/IEC 23002-7 was issued at 32nd meeting 2023-10, DAM ballot to close 2024-04-08, no action at the current meeting.
* HEVC new levels (from JVET-Z1005) – ISO/IEC 23008-2 DIS of new edition of HEVC was issued from the April 2022 26th meeting, incorporating Amd.1 and corrigenda items (ballot closed 2023-01-10, ballot comments in the Summary of Voting document [m61834](https://dms.mpeg.expert/doc_end_user/current_document.php?id=85619&id_meeting=193)); note that Amd.1 = shutter interval SEI is already included in latest ITU-T edition of H.265. It is noted that there are potential additional items (corrigenda+tickets, YCgCo-Re and YCgCo-Ro draft, SMPTE ST 2128, multiview profiles draft) where only corrigenda items were included in the FDIS text based on ballot comments, ballot had not been started yet. ITU-T consent for a new edition is planned for July 2023. It was noted that the referencing of VSEI is also somewhat different in the ITU-T and ISO/IEC versions of HEVC and/or AVC, which might be aligned at the next convenient time (basically editorial – e.g., the ITU version of AVC specifies the annotated regions SEI message without referencing VSEI, whereas the ISO/IEC version references VSEI for the syntax and semantics of that SEI message). However, there is currently no other need for HEVC to reference the VSEI standard. An FDIS for HEVC was issued as an output of the 29th meeting in January 2023 (and it does not reference VSEI). Its ballot began 2023-08-06 and closed 2023-10-02, and it was pending publication. A new edition of H.265 (v9) was Consented in July 2023, approved 2023-09-13, and pre-published 2023-09, and published 2023-11-24 (not referencing VSEI).
* VVC new level and systems-related supplemental enhancement information (from JVET-AA2005) – VVC DAM was issued from 27th meeting, ballot closed 2023-01-03, ballot comments in the Summary of Voting document [m61833](https://dms.mpeg.expert/doc_end_user/current_document.php?id=85618&id_meeting=193). This was converted into a preliminary FDIS of VVC 3rd edition ([WG 5 N 183](https://dms.mpeg.expert/doc_end_user/current_document.php?id=86365&id_meeting=193)) at the 29th meeting of January 2023, anticipating that some alignment would be necessary with the ongoing VSEI amendment. Another preliminary FDIS was issued (WG 5 N 202) from the April meeting. The FDIS was then issued (WG 5 N 228) from the 31st meeting in July 2023. A new edition of H.266 was Consented in July 2023, approved 2023-09-29 and pre-published 2023-09, and published 2023-11-29.
* VVC Conformance testing for operation range extensions – (from JVET-Y2026) – the DAM ballot closed 2022-11-15 (ballot comments in the Summary of Voting document [m61832](https://dms.mpeg.expert/doc_end_user/current_document.php?id=85617&id_meeting=193)), and this was consolidated into an FDIS at the 29th meeting, but the ballot had not been started yet. ITU-T H.266.1 was Consented in July 2023, approved 2023-09-13 and pre-published in 2023-09, and published 2023-10-19.
* VSEI additional SEI messages (from JVET-AB2006) – VSEI DAM (JVET draft 3) was issued from the 28th meeting and a DAM ballot was issued. The FDIS of a new edition of ISO/IEC 23002-7 was issued (WG 5 N 220) from the 31st meeting in July 2023 and also reached ITU-T Consent at that meeting. H.274 v3 was approved 2023-09-29, pre-published 2023-10-11, and currently pending publication.
* Film grain synthesis technology for video applications – JVET draft 4 and the ISO/IEC 23002-9 CDTR were issued at the 29th meeting (JVET-AC2020) (a request to start work on the TR had been made at the 25th meeting), and the CDTR consultation period ended 2023-07-09. A DTR text was issued from the 31st meeting in July 2023 but was put on hold by ISO staff editors, so the ballot had not been issued. (It was noted that a second DTR could become necessary in case of comments). The publication limit date was reportedly 2023-08-09, so action to extend that date may be needed. ITU-T approval would be anticipated in April 2024.
* Video CICP new edition draft for YCgCo-Re and YCgCo-Ro (from JVET-Z1003), an ISO/IEC 23091-2 preliminary FDIS was issued from the 30th meeting and the Summary of Voting document was available as [m62572](https://dms.mpeg.expert/doc_end_user/current_document.php?id=86621&id_meeting=194) and a draft DoC had been issued as WG 5 [N 205](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/library/2/Draft%20disposition%20of%20comments%20received%20on%20ISO-IEC%20DIS%2023091-2%3A202X). There was a delay in the submittal of the FDIS due to dependency on the status of SMPTE ST 2128, which was tentatively included in the preliminary FDIS, based on an NB comment. The video CICP colour type indicator for SMPTE ST 2128 had been drafted and incorporated into the preliminary FDIS issued at the 30th meeting of April 2023. It had been reported that the specification was expected to become finalized in the SMPTE meeting in March 2023, but this had not yet happened, so the production of the FDIS was delayed. ITU-T Consent for H.273 v3 proceeded at the 31st meeting of July 2023 (to prevent undue delay since SG16 does not meet very frequently) and the text was approved in September 2023, but the text was on hold pending the publication of SMPTE ST 2128.
* A request for free availability in ISO/IEC has to be made for each edition, amendment and corrigendum, and the request needs to be approved in the WG 5 Recommendations. A request form also needs to be filled out (but the form does not need to be issued as a WG 5 document). A freely available URL for the ITU publication should be provided for the following parts:
  + For the ongoing work items, when they become finalized
  + ISO/IEC 23008-2:2020/Amd.1:2021 – HEVC FDAM issued 20th meeting (October 2020), public availability not yet requested but may not be necessary as it becomes included in next edition

## Opening remarks

Remarks during the opening session of the meeting Wednesday 17 January at 0500 UTC were as follows.

* Timing and organization of the meeting and online access, calendar posting of session plans
  + The initial number of documents was approximately the same as the previous meeting (approximately 170 by the time of opening the meeting) – parallel sessions will be necessary. Y. Ye was asked to chair sessions on EE2 related discussions. Considering that items about JVET management structure had been discussed in the December meeting of WP 3, objection were raised on this aspect by on expert. It is noted that the fact that the JVET chair asked Y. Ye for chairing those sessions (under the aspect of her being experienced in that) does not anticipate any later decision of parent bodies on that aspect. Other potential candidates for chairing parallel discussions could be F. Bossen and J. Boyce.
  + Scheduling of NNVC discussions – should be done early (some overlap with JPEG meeting this time, JPEG meets from Jan. 21, but the 5UTC and 720UTC sessions could be used)
* Plans for subsequent F2F meeting in April (Rennes), July (Sapporo), and November 2024 (Antalya).
* The meeting logistics, agenda, working practices, policies, and document allocation considerations were reviewed.
  + Access to the meeting was provided using Zoom. Recording of the meeting notes by the session chair will also permanently be shared via zoom.
  + Having text and software available is crucial (and not just arriving at the end of the meeting).
  + There were no objections voiced in the opening plenary to the consideration of late contributions.
* The results of the previous meeting and the meeting report JVET-AF1000 were reviewed. The following small issues in the meeting report were noted and were not considered sufficient to warrant issuing a revision. These are obviously left over from a previous report, and the correct information can be found in other places of the report:
  + In the summary (section 1) and in annex C (WG 5 recommendations), it is wrongly stated that the meeting in October 2025 would be under SC 29 auspices. Actually, it can be expected to be held under SG16 auspices, but date and location are still unknown. In section 11 (future meeting plans), the correct information is given.
  + In the summary (section 1), for some documents without standards deliverable status the statement that they had been “also issued as WG 5 Nxxx” was missing.
* There was a somewhat decreased number of late non-cross-check documents, not compared to the last, but compared to other previous meetings. However, not all non-cross-check documents that had been registered before the deadline were also available in time.
* There were again a few documents registered where authors’ given names were not abbreviated, and/or company affiliation was missing in the authors’ list. Participants were reminded to stick to JVET’s conventions.
* Experts were asked not to pick a specific JVET number for regular documents – this function is reserved for AHG reports, summary reports, and output docs
* Experts were asked to always register JVET documents via the “jvet-experts.org” site, not via the MPEG dms site.
* Experts were asked to inform the chair when the title of a document is changed, or if authors are added. Otherwise, that might not be correct in the meeting notes.
* It was noted that during the interim period since the last meeting, a crash happened to the password database of the jvet.org site. Therefore, it was necessary for all registered users to reset passwords. This caused problems in some cases where the registered email address had become obsolete (e.g., after change of employer).
* The primary goals of the meeting were:
  + Plan for new version of VVC software as standards part – target April 2024, for inclusion of NNPF SEI, software from TRs, layered coding, bug fixes, etc.
  + New edition video CICP FDIS (DIS ballot response in [m62572](https://dms.mpeg.expert/doc_end_user/current_document.php?id=86621&id_meeting=194), draft DoCR in [MDS22710](https://dms.mpeg.expert/doc_end_user/current_document.php?id=87856&id_meeting=194), and preliminary FDIS text in [MDS22711](https://dms.mpeg.expert/doc_end_user/current_document.php?id=87857&id_meeting=194)) – inclusion of ST 2128 descriptor needed clarification of the status, therefore it was not yet submitted for FDIS yet; was consented and approved in ITU-T, but cannot be published as long as ST 2128 is not available. This was not yet available by the time of the current meeting – therefore it was concluded to issue the FDIS in April.
  + TR on film grain synthesis technology for video applications – the DTR ballot had not been issued as expected, due to referencing and other issues. ITU consent is targeted for April 2024, possibly second DTR by April as well. Any preliminary draft at the current meeting?
  + Optimization of encoders and receiving systems for machine analysis of coded video content – no new WD of TR was issued from last meeting, more input this time, discuss how to proceed
  + Preparation of subjective tests for film grain, new content available – expert viewing to be planned for April meeting
  + Expert viewing for multi-layer verification test to be planned for April meeting
  + Any action items on reference software JM/HM? Status of MV-HEVC software and test conditions (refer to resolution of a previous meeting for the latter)?
  + New edition of AVC
    - 11th ed. 14496-10 under DIS ballot (to close 2024-04-15; response expected before April 2024 meeting)
    - ITU consent for new H.264 edition was targeted for April 2024
  + HEVC updates
    - Amd.1 of 5th ed. is under DAM ballot (to close 2024-04-08; responses expected before April 2024 meeting).
    - 5th edition was published in 2023-10. This will allow for a next edition including the ongoing Amd.1 in 2024 (converting FDAM into FDIS in April)
    - Consent for next H.265 edition was targeted for April 2024 with new elements from DAM
  + Exploration Experiments
    - Neural network-based video coding
    - Enhanced compression beyond VVC
* Liaison communication:
  + Incoming liaison statements: JPEG M66811.
* Joint meetings were expected with AG 5 (on XXX) and possibly with other groups.
* Principles of standards development were discussed.
* Scheduling of sessions was discussed – see under 2.6 and 2.15.

## Scheduling of discussions

The plans for the times of meeting sessions were established as follows, in UTC (which for this meeting was 1 hour behind the time in Geneva and Paris; 8 hours ahead of the time in Los Angeles, etc.). No session was scheduled to last longer than 2 hrs.

* 0500–0700 1st “morning” session [break after 2 hours]
* 0720–0920 2nd “morning” session

[“overday” break – nearly 12 hours]

* 2100–2300 1st “night” session [break after 2 hours]
* 2320–0120+1 2nd “night” session

Particular scheduling notes are shown below, although not necessarily 100% accurate or complete. Times are recorded in the local timezone of the meeting venue, except as otherwise noted:

* Wed. 17 January, 1st day
  + Morning sessions:
    - 0500–0600 Opening remarks, review of practices, agenda, codes of conduct, IPR policy reminder
    - 0605–0700 Reports of AHGs 1, 3-7
    - 0720–0920 Reports of AHGs 2, 8-16
  + Late-night sessions:
    - 2100–2310 EE1 summary report
    - 2330–0120+1 EE2 summary report
* Thu. 18 January, 2nd day
  + Morning sessions:
    - 0500–0710 EE2 related
    - 0725–0920 HLS/SEI 6.1
    - 0730-0920 5.2.3 EE2 related / 5.2.4 non-EE2 intra / inter (chaired by Y. Ye)
  + Late-night sessions:
    - 2100–2300 HLS/SEI 6.2 SPO SEI
    - 2100–2300 5.2.3 EE2 related / 5.2.4 non-EE2, topics other than intra / inter (chaired by Y. Ye)
    - 2320–0120+1 EE2 summary EE2-2.12, 3.x, 4.x, 5.x
    - 2320–0120+1 BoG on HLS SEI 6.2 (J. Boyce)
* Fri. 19 January, 3rd day
  + Morning sessions:
    - 0500–0700 HLS SEI 6.2/6.4
    - 0500–0700 5.2.3 EE2 related / 5.2.4 non-EE2 remaining docs (chaired by Y. Ye)
    - 0720–0920 HLS SEI 6.4/6.8
    - 0720–0920 5.2.3 EE2 related / 5.2.4 non-EE2 remaining docs (chaired by Y. Ye)

Late-night sessions:

* + - 2100–2310 HLS SEI 6.2 revisits, 6.8, non-SEI 6.9
    - 2100–2300 5.2.3 EE2 related / 5.2.4 non-EE2 remaining docs (chaired by Y. Ye)
    - 2320–0120+1 BoG on HLS 6.2/6.8/6.9 (J. Boyce)
    - 2330–0120+1 NNVC 5.1.3, 5.1.4, 5.1.5
* Mon. 22 January, 4th day
  + 0500–0800 MPEG information sharing session
  + Morning session
    - 0820–0920 NNVC remaining 5.1.4, EE revisits
  + Late-night sessions:
    - 2100–2300 HLS 6.9/6.5, revisits 6.1
    - 2100–2300 5.2.3 EE2 related / 5.2.4 non-EE2 remaining docs (chaired by Y. Ye)
    - 2320–0120+1 Further planning, HLS 6.5/6.6
* Tue 23 January, 5th day
  + Morning sessions:
    - 0500–0700 Film grain synthesis 4.12/6.7
    - 0500-0700 BoG on EE1 (E. Alshina, F. Galpin)
    - 0720–0920 Generative face video 4.16/6.3
  + Late-night sessions:
    - 2100–2300 4.6/4.9 (joint with AG5), 4.10
    - 2320–0120+1 4.2
* Wed. 24 January, 6th day
  + 0500–0600 MPEG information sharing session
  + Morning sessions:
    - 0600–0630 Joint with WG4 /WG7/VCEG JM on SEI rules and MPI/MIV
    - 0630-0700 Joint with AG5/WG4 /VCEG on multi-layer coding tests (4.5)
    - 0600-0700 BoG on EE1 (E. Alshina, F. Galpin)
    - 0720–0920 Revisits EE1 & BoG report, 4.2 next VSEI version, 4.14
  + Late-night sessions:
    - 2100–2300 Remaining docs 4.12/4.2/10/6.9
    - 2320–0120+1 Remaining doc review & revisits EE2, 4.1, 4.8, 4.11 (Y. Ye)
    - 2320–0020+1 BoG on gaming content compression (S. Puri, J. Sauer)
    - 2320–2400 BoG on FGC subjective test (M. Wien)
* Thu. 25 January, 7th day
  + Morning sessions:
    - 0500–0705 4.15 and BoG reports, other doc review 4.10/4.16/
    - 0600–0700 BoG on preparation ECM expert viewing (M. Wien)
    - 0725–0920 review draft liaison, EE1/EE2 revisits, output planning, 4.2
  + Late-night sessions:
    - 2100–2305 Review 4.8, 4.11, Revisits 6.x
    - 2315–0120+1 Other documents TBP, remaining business, AHG and output planning
* Fri. 26 January, 8th day
  + 0500–0920 (with break) JVET wrap-up plenary:
    - Approval of output docs
    - Establishment of AHGs
    - Review of meeting recommendations
    - Future planning, a.o.b.
  + 2100–2300 MPEG information sharing session
  + XXXX–XXXX WG 5 approval of meeting recommendations, closing of meeting

## Contribution topic overview

The approximate subject categories and quantity of contributions per category for the meeting were summarized as follows (note that the noted document counts do not include crosschecks and summary reports, and may not be completely accurate; documents which are allocated to multiple sections are only counted in one of them):

* AHG reports (16) (section 3)
* Project development (section 4)
  + AHG1: Deployment and advertisement of standards (2)
  + AHG2: Text development and errata reporting (4)
  + AHG3: Test conditions (0)
  + AHG3: Software development (0)
  + AHG4: Subjective quality testing and verification testing (2)
  + AHG4: Test Material (1)
  + AHG4: Codec performance with alternative test material (0)
  + AHG5: Conformance test development (1) TBP
  + AHG7: ECM tool assessment (1)
  + AHG8: Optimization of encoders and receiving systems for machine analysis of coded video content (7)
  + AHG10: Encoding algorithm optimization (1) TBP
  + AHG13: Film grain synthesis (5)
  + Implementation studies (0)
  + Profile/tier/level specification (1)
  + Gaming content compression (1)
  + Generative face video (4)
* Low-level tool technology proposals (section 5) with subtopics (number counts excluding BoG and summary reports)
  + AHG11/AHG14 and EE1: Neural network-based video coding (19) (section 5.1)
  + AHG6/AHG12 and EE2: Enhanced compression beyond VVC capability (87) (section 5.2)
* AHG9: High-level syntax (HLS) proposals (section 6) with subtopics
  + SEI messages on NNPF and processing order (11) (sections 6.1, 6.2)
  + SEI messages on other topics (30) (sections 6.2, 6.3, 6.4, 6.5, 6.6, 6.7)
  + Non-SEI HLS aspects (5) (section 6.9) 1 TBP
* Joint meetings, plenary discussions, BoG reports (4), liaison (1), summary of actions (section 6)
* Project planning (section 8)
* Establishment of AHGs (section 9)
* Output documents (section 10)
* Future meeting plans and concluding remarks (section 11)

The document counts above do not include cross-checks and summary reports.

# AHG reports (16)

These reports were discussed during 0505–0700 and 0720-0915 on Wednesday 17 Jan. 2024 (chaired by JRO).

[JVET-AG0001](https://jvet-experts.org/doc_end_user/current_document.php?id=13774) JVET AHG report: Project Management (AHG1) [J.-R. Ohm (chair), G. J. Sullivan (vice chair)]

[JVET-AG0002](https://jvet-experts.org/doc_end_user/current_document.php?id=13776) JVET AHG report: Draft text and test model algorithm description editing (AHG2) [B. Bross, C. Rosewarne (co-chairs), F. Bossen, A. Browne, S. Kim, S. Liu, J. R. Ohm, G. J. Sullivan, A. Tourapis, Y.-K. Wang, Y. Ye (vice chairs)]

1. **Ad hoc group activity**
   1. **Output documents produced**
      1. **JVET-AF1004 Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP**

This document contains a list of reported errata items for VVC, VSEI, HEVC, AVC, and Video CICP, for tracking purposes. Some of the items have been confirmed by the JVET and have been agreed to require fixing. This document also provides publication status backgrounds of these standards.

Incorporated items at the JVET-AF meeting:

* For VVC:
  + Add a missing space in "if(IntraSubPartitionsSplitType != ISP\_NO\_SPLIT )" (JVET-AF0064).
  + On coding of syntax elements of fix-valued, non-reserved bits using f(N) instead of u(N). (JVET-AF0064).
  + Replace the two instances of "in the first access unit of the CVS" with "in the first access unit of the CVS in decoding order". (JVET-AF0189)
* For VSEI:
  + Replace each of the 10 instances of "nuh\_layer\_id" in subclauses 8.22.2 and 8.23.2 with "layer identification". (from Gary J. Sullivan)
  + On cancellation of the persistence of a CTI SEI message by another CTI SEI message. (JVET-AF0064)
  + On coding of syntax elements of fix-valued, non-reserved bits using f(N) instead of u(N). (JVET-AF0064).
  + Consider adding a NOTE in VSEI telling something like: 1) AVC can reference VSEI for enabling the inclusion of only those SEI messages for which the SEI payloads are specified in VSEI but the syntax has not been extended; 2) HEVC and VVC can reference VSEI for enabling the inclusion of SEI messages for which the SEI payloads are specified in VSEI regardless of whether the syntax has been extended or not. (from Ye-Kui Wang)
* For HEVC:
  + Updated the publication status background.
  + (Perhaps for other specs such as AVC, VVC, and VSEI as well) On that the term of "true subset" should be replaced with "proper subset" or "strict subset". (from Gary J. Sullivan)
  + (Perhaps for other specs such as AVC, VVC, and VSEI as well) On that the mixed uses of Boolean operators on numerical values and vice versa should cleaned up. (from Gary J. Sullivan)
  + Removed obsoleted HEVC errata items that were included in JVET-AE1004.
* For AVC:
  + Removed obsoleted AVC errata items that were included in JVET-AE1004.
    1. **JVET-AF1006 New profiles, colour descriptors, and SEI messages for HEVC (draft 2)**

This document contains the draft text for changes on new profiles, colour descriptors, and SEI messages for the 5th edition of the High Efficiency Video Coding (HEVC) standard (Rec. ITU-T H.265 | ISO/IEC 23008-2). The changes include: 1) the support of four new profiles, namely the Multiview Main 10, Multiview Monochrome, Multiview Monochrome 10, and Multiview Monochrome 12 profiles; 2) the support of three additional colour type identifiers; 3) the HEVC-specific supports for some supplemental enhancement information (SEI) messages that may be included in HEVC bitstreams but are not to be specified in the HEVC specification, and 4) some technical corrections and editorial improvements to the 5th edition text of HEVC. The SEI messages are the neural network post-filter characteristics (NNPFC) SEI message, the neural-network post-filter activation (NNPFA) SEI message, and the phase indication SEI message, that are to be specified in the 3rd edition of the Versatile Supplemental Enhancement Information messages for coded video bitstreams (VSEI) standard (ITU‑T H.274 | ISO/IEC 23002-7).

Draft 2 incorporated items:

* In 7.2: Added a NOTE on the b(8) syntax descriptor. [JVET-AF0064]
* In 8.5.3.2.2, bullet item 6: Replaced "variable" with "variables". [Email from C. Reader]
* In D.2.24: Changed the coding of nesting\_zero\_bit from u(1) to f(1). [JVET-AF0064]
* In D.2.44: Changed the coding of mcts\_nesting\_zero\_bit from u(1) to f(1). [JVET-AF0064]
* In D.3.13, Eqn. D-15: Added a closing curly bracket after "G[ c ][ x + 1 ][ y − 1 ]". [FR008 in WG 5 N 243]
* In D.3.45: Improved the phrasing of an SEI message being indicated by the encoder (i.e., the content producer) as being "necessary". [Email from G. J. Sullivan]
* In D.3.45, D.3.46, and I.14.3.3: Appended "in decoding order" after "the first access unit of the CVS". [JVET-AF0189]
* In F.3.29: Improved the phrasing of the NOTE. [Email dicussion between A. Tourapis and Y.-K. Wang]
* In F.7.4.2.2: Improved the phrasing of NOTE 1. [Email dicussion between A. Tourapis and Y.-K. Wang]
* In F.14.2.5: Changed the coding of bsp\_nesting\_zero\_bit from u(1) to f(1). [JVET-AF0064]
* In G.11.1.1, first sentence: Changed "the Multiview Main and Multiview Main 10 profile" to "the Multiview Main or Multiview Main 10 profile". [JVET-AF0063]
* In G.11.1.1 and I.11.1.1: Removed "to" from "greater than to". [JVET-AF0063]
* In G.11.1.1: Applied the following constraint to the Multiview Main 10 profile: "All active PPSs for layers in subBitstream shall have colour\_mapping\_enabled\_flag equal to 0 only." [JVET-AF0063]
* In G.11.1.2: Changed the constraint for the multiview monochrome profiles requiring chroma\_format\_idc to be equal to 1 to requiring chroma\_format\_idc to be equal to 0. [JVET-AF0063]
* In G.11.1.2: Replaced "exensions" with "extensions". [JVET-AF0063]
  + 1. **JVET-AF1016 AVC with extensions and corrections (draft 2)**

(not available at the time of preparing this AhG report)

* + 1. **JVET-AF2002 Algorithm description for Versatile Video Coding and Test Model 21 (VTM 21)**

The JVET established the VVC Test Model 23 (VTM23) software at its 32nd meeting (13-20 October 2023, Hannover). This document serves as a source of general tutorial information on the VVC design and also provides an algorithm description and encoding method description of VTM23 software. It is noted that, as no update of the algorithm description document was released at either the 27th or the 31st JVET meetings, the numbering of the VTM software version is two higher than that of this description document. In the main body of the text, numbering refers to the software version. The VVC has been developed by a joint collaborative team of ITU-T and ISO/IEC experts known as the Joint Video Experts Team (JVET), which is a partnership of ITU-T Study Group 16 Question 6 (known as VCEG) and ISO/IEC JTC 1/SC 29/WG 11 (known as MPEG). This new standard has been designed with two primary goals. The first of these is to specify a video coding technology with a compression capability that is substantially beyond that of the prior generations of such standards, and the second is for this technology to be highly versatile for effective use in a broadened range of applications. In addition to the applications that have commonly been addressed by prior video coding standards, some key application areas for the use of this standard include in particular ultra-high-definition video (e.g., with 3840×2160 or 7620×4320 picture resolution and bit depth of 10 or 12 bits as specified in Rec. ITU-R BT.2100), video with a high dynamic range and wide colour gamut (e.g., with the perceptual quantization or hybrid log-gamma transfer characteristics specified in Rec. ITU-R BT.2100), and video for immersive media applications such as 360° omnidirectional video projected using a common projection format such as the equirectangular or cubemap projection format.

Ed. Notes:

VVC Test Model 21 (VTM21) algorithm description and encoding method v1

• Incorporated JVET-AF0111: AHG10: MTT split modes early termination

• General editorial improvements

* + 1. **JVET-AF2027 SEI processing order and processing order nesting SEI messages in VVC (draft 6)**

This document contains the draft text for changes to the Versatile Video Coding (VVC) standard (Rec. ITU-T H.266 | ISO/IEC 23090-3), to specify the SEI processing order and processing order nesting SEI messages.

**Changes yet to be integrated:**

None.

**Changes that have been integrated:**

1. Addition of po\_id (JVET-AF0061, JVET-AF0174, JVET-AF0067, JVET-AF0310) [Bytedance/HHI/MediaTek/Nokia/Dolby]
2. On the persistence scope of the SPO SEI message (JVET-AF0189, JVET-AF0049, JVET-AF0061, JVET-AF0174, JVET-AF0310) [Dolby/Nokia/Bytedance/HHI/MediaTek]
3. Addition of the processing order nesting SEI message (JVET-AF0049, JVET-AF0174, JVET-AF0310) [Nokia/HHI/Dolby/Bytedance/MediaTek]
4. Removal of the constraint requiring that there shall be at least two values of po\_sei\_processing\_order[ i ] that are not equal, which is redundant, from JVET-AF0049 and JVET-AF0062 [Nokia/Bytedance]
5. The SEI prefix indications, when present, are signalled in units of bits instead of in units of bytes, same as in the SEI prefix indication SEI message, from JVET-AF0189, JVET-AF0062, and JVET-AF0049. [Dolby/Bytedance/Nokia]
6. Move po\_sei\_prefix\_flag[ i ] from immediately before po\_sei\_payload\_type[ i ] to be immediately after po\_sei\_payload\_type[ i ], from JVET-AF0062. [Bytedance]
7. Clarify the following aspects: In the semantics of the SPO SEI message, two different types of SEI messages may have the same SEI payloadType value but are differentiated by some syntax elements in the SEI payload. For example, two NNPFC SEI messages with different nnpfc\_id values are considered as having two different SEI message types. From JVET-AF0062. [Bytedance]
8. Using separate loops for the payload type and processing order information, from JVET-AF0189. [Dolby]
9. Removing the constraint that “The value of po\_sei\_processing\_order[ po\_num\_sei\_messages\_‌minus2 + 1 ] shall not be equal to 0”, from JVET-AF0189. [Dolby]
10. Modifying the use of SeiProcessingOrderSeiList such that it determines which SEI messages are allowed to appear in an SEI processing order SEI message, and update the values in the list, including disallowing the SEI payloadType value of the decoded picture hash SEI message from being included in the list, from JVET-AF0189 and JVET-AF0070. [Dolby/Sharp]
11. **Related input contributions**

The following input contribution was noted as relevant to the work of this ad hoc group:

* [JVET-AG0079](https://jvet-experts.org/doc_end_user/current_document.php?id=13635) AHG2/AHG9: VUI extension mechanism and picture modality information for AVC and HEVC
* [JVET-AG0213](https://jvet-experts.org/doc_end_user/current_document.php?id=13769) AHG1/AHG2/AHG8: On project management related to the encoder optimization information SEI message

1. **Remaining bug tickets**

Carried over:

* [#1594](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1594) Mismatch between VVC spec and VTM for sample generation in CCLM process
* [#1607](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1607) Wrong sign (+ instead of -) in Fig 16 Flowchart for decoding a decision in the ITU text
* [#1609](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1609) NoBackwardPredFlag derivation ambiguity
* [#1617](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1617) Not initialized NumCtusInSlice[0] to 0.

New:

* [#1618](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1618) [Multilayer Profiles] Potential Mismatch of VTM22.0 & Specification Related To Derivation Process For Merge Motion Vector Difference

1. **Recommendations**

The AHG recommends to:

* Approve JVET-AF1004, JVET-AF1006, JVET-AF1016, JVET-AF2002, and JVET-AF2027 documents as JVET outputs,
* Compare the VVC documents with the VVC software and resolve any discrepancies that may exist, in collaboration with the software AHG,
* Encourage the use of the issue tracker to report issues with the text of both the VVC specification text and the algorithm and encoder description,
* Continue to improve the editorial consistency of VVC text specification and Test Model documents,
* Ensure that, when considering changes to VVC, properly drafted text for addition to the VVC Test Model and/or the VVC specification text is made available in a timely manner,
* Review bug tickets, and other AHG2 related inputs and act on them if found to be necessary.

[JVET-AG0003](https://jvet-experts.org/doc_end_user/current_document.php?id=13777) JVET AHG report: Test model software development (AHG3) [F. Bossen, X. Li, K. Sühring (co-chairs), E. François, Y. He, K. Sharman, V. Seregin, A. Tourapis (vice chairs)]

The software model versions prior to the start of the meeting were:

* [VTM 23.0](https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/-/releases/VTM-23.0) (Dec. 2023)
* [HM-18.0](https://vcgit.hhi.fraunhofer.de/jvet/HM/-/releases/HM-18.0) (Apr. 2023)
* [HM-16.21+SCM-8.8](https://vcgit.hhi.fraunhofer.de/jvet/HM/-/tags/HM-16.21+SCM-8.8) (Mar. 2020)
* [SHM 12.4](https://vcgit.hhi.fraunhofer.de/jvet/SHM/-/tags/SHM-12.4) (Jan. 2018)
* [HTM 16.3](https://vcgit.hhi.fraunhofer.de/jvet/HTM/-/tags/HTM-16.3) (Jul. 2018)
* [JM 19.1](https://vcgit.hhi.fraunhofer.de/jvet/JM/-/releases/JM-19.1) (Apr. 2023)
* [JSVM 9.19.15](https://vcgit.hhi.fraunhofer.de/jvet/jsvm/-/tags/JSVM_9_19_15)
* [JMVC 8.5](https://vcgit.hhi.fraunhofer.de/jvet/jmvc/-/tags/JMVC_8_5)
* [3DV ATM 15.0](https://vcgit.hhi.fraunhofer.de/jvet/3dv-atm/-/tags/3DV-ATM_v15.0) (no version history)
* [HDRTools 0.24](https://gitlab.com/standards/HDRTools/-/tags/v0.24) (March 2023)

Software for MFC and MFCD is only available as published by ITU-T and ISO/IEC. It is planned to create repositories with the latest versions available in ITU-T H.264.2 (02/2016). All development history is lost.

1. **Software development**

Development was continued on the GitLab server, which allows participants to register accounts and use a distributed development workflow based on git.

The server is located at:

<https://vcgit.hhi.fraunhofer.de>

The registration and development workflow are documented at:

<https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/wikis/VVC-Software-Development-Workflow>

Although the development process is described in the context of the VTM software, it can be applied to all other software projects hosted on the GitLab server as well.

Between the 31st and 32nd meetings, the GitLab server was affected by a huge number of spam issues, which were generated using newly registered accounts. Administrator user account confirmation was enabled to protect the server from spam postings. Proponents should plan for some time for server admins to confirm new accounts.

Before the 33rd meeting an issue was found, which prevents the GitLab server from sending out emails to external addresses. This prevents confirmation of new email addresses and thus the registration of new accounts. Email notifications for changes (e.g. updates on merge requests) are also not sent out. Proponents and software coordinators should check manually for updates. The issue was only identified recently and is expected to be resolved soon.

1. **VTM related activities**

The VTM software can be found at

<https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/>

The software development continued on the GitLab server. VTM versions 22.1 and 22.2 were tagged on Oct. 22, and VTM version 23.0 was tagged on Dec. 12. VTM 23.1 is expected during the 33rd JVET meeting.

VTM 22.1 was tagged on Oct. 22, 2023. Changes include:

* FIX #1611: use pic->reconstructed and pic->layerId for computing HashME condition
* FIX #1613: Fix transform skip cost for non-CTC conditions
* Update software manual on max transform size (#1614)
* Clean up IBC-related code
* Fix #1616: NNPFC SEI nnpfc\_chroma\_sample\_loc\_type\_frame value is checked even when not signalled
* Fix #1612: PO-SEI message
* JVET-AE0156: Message wrapping and importance indication for the SEI processing order SEI message
* Free vector data memory when destroying CodingStructure buffers
* Port CS mem reduction from ecm/ECM!525
* Port using actual CTU size instead of MAX\_CU\_SIZE at encoder from ecm/ECM!516
* Relax strict equality when comparing floating-point numbers
* Port picture buffer refactoring from ecm/ECM!523
* Fix unreleased RPR temp buffer
* Fix #1615: Encoding exception with VS compiler
* Fix y4m header writing when RPR is used and UpscaledOutput=2

VTM 22.2 was tagged on Oct. 23, 2023. Changes include:

* Remove macros from previous cycle

VTM 23.0 was tagged Dec. 12, 2023. Changes include:

* Add SIMD implementations of transforms
* TileIdx as a new type
* Fix crash for class F: set SPS in CodingStructure
* Fix debugBitstream functionality
* Fix out-of-bounds access in EncCu::updateRdCheckingNum
* Remove parameter UseBLambdaForNonKeyLowDelayPictures
* Remove toLast and unused variables in ClpRngs
* Fix #1622: Initialising NNPFC SEI attributes in the default constructor
* Fix #1621: Compilation issue in non-X86 environments
* JVET-AF0122-SW: Lagrange multiplier optimization for ALF
* Fix debug bitstream for multilayer
* JVET-AF0310: Processing Order Nesting SEI message
* Cleanup: replace TAB characters with spaces in config files

VTM 23.1 is expected to be tagged during the 33rd JVET meeting. Changes are expected to include bug fixes and code optimizations.

* 1. ***CTC Performance***

The following tables shows **VTM 20.0** performance over **HM 17.0** (not updated, results identical with previous versions).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **All Intra Main10** |  |  |
|  |  |  | **Over HM-17.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | -29.50% | -32.81% | -33.83% | 1273% | 166% |
| Class A2 | -29.73% | -24.39% | -21.58% | 2142% | 174% |
| Class B | -22.32% | -27.21% | -30.98% | 2397% | 175% |
| Class C | -22.89% | -19.53% | -23.19% | 3348% | 179% |
| Class E | -26.04% | -25.90% | -24.15% | 1869% | 158% |
| **Overall** | -25.50% | -25.75% | -27.02% | 2187% | 171% |
| Class D | -18.80% | -13.85% | -13.68% | 4270% | 189% |
| Class F | -39.49% | -40.23% | -42.90% | 4220% | 171% |
|  |  |  |  |  |  |
|  |  |  | **Random access Main10** |  |  |
|  |  |  | **Over HM-17.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | -40.60% | -40.71% | -47.16% | 586% | 158% |
| Class A2 | -44.02% | -41.73% | -40.78% | 679% | 170% |
| Class B | -37.41% | -50.12% | -48.51% | 672% | 161% |
| Class C | -33.86% | -36.28% | -38.21% | 923% | 164% |
| Class E |  |  |  |  |  |
| **Overall** | -38.42% | -42.87% | -43.95% | 713% | 163% |
| Class D | -31.72% | -32.60% | -31.95% | 970% | 161% |
| Class F | -46.07% | -49.76% | -50.65% | 490% | 145% |
|  |  |  |  |  |  |
|  |  |  | **Low delay B Main10** |  |  |
|  |  |  | **Over HM-17.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -30.40% | -37.43% | -35.12% | 629% | 152% |
| Class C | -28.18% | -21.23% | -21.06% | 817% | 167% |
| Class E | -31.16% | -35.72% | -28.79% | 309% | 140% |
| **Overall** | -29.85% | -31.60% | -28.85% | 575% | 154% |
| Class D | -26.75% | -16.50% | -14.88% | 881% | 179% |
| Class F | -42.17% | -44.02% | -44.14% | 435% | 137% |
|  |  |  |  |  |  |
|  |  |  | **Low delay P Main10** |  |  |
|  |  |  | **Over HM-17.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -34.52% | -38.34% | -36.12% | 526% | 160% |
| Class C | -28.07% | -17.73% | -18.09% | 670% | 176% |
| Class E | -33.06% | -37.89% | -31.10% | 280% | 146% |
| **Overall** | -32.01% | -31.36% | -28.85% | 487% | 161% |
| Class D | -26.68% | -12.73% | -11.47% | 762% | 175% |
| Class F | -40.13% | -41.24% | -41.72% | 437% | 145% |

According to common test conditions in random access configuration HM is using a GOP size of 16 pictures compared to VTM using a GOP of 32 pictures. Random access points are inserted approximately every second aligned with a GOP boundary of GOP 32 in both VTM and HM. VTM uses two more reference pictures in random access than HM (due to more memory being availably in typical level settings).

The following table shows **VTM 23.0** performance over **VTM 22.0** using SDR CTC:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **All Intra Main10** |  |  |
|  |  |  | **Over VTM-22.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | 0.00% | 0.00% | 0.00% | 94% | 91% |
| Class A2 | 0.00% | 0.00% | 0.00% | 96% | 94% |
| Class B | 0.00% | 0.00% | 0.00% | 97% | 94% |
| Class C | 0.00% | 0.00% | 0.00% | 98% | 97% |
| Class E | 0.00% | 0.00% | 0.00% | 97% | 93% |
| **Overall** | 0.00% | 0.00% | 0.00% | 96% | 94% |
| Class D | 0.00% | 0.00% | 0.00% | 98% | 98% |
| Class F | 0.00% | 0.00% | 0.00% | 90% | 97% |
|  |  |  |  |  |  |
|  |  |  | **Random access Main10** |  |  |
|  |  |  | **Over VTM-22.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | 0.00% | -0.08% | 0.01% | 93% | 94% |
| Class A2 | 0.06% | 0.07% | 0.06% | 93% | 96% |
| Class B | 0.02% | -0.03% | -0.03% | 93% | 96% |
| Class C | 0.02% | 0.03% | -0.05% | 93% | 96% |
| Class E |  |  |  |  |  |
| **Overall** | 0.02% | 0.00% | -0.01% | 93% | 96% |
| Class D | 0.03% | -0.25% | 0.06% | 96% | 97% |
| Class F | 0.14% | 0.17% | 0.12% | 91% | 96% |
|  |  |  |  |  |  |
|  |  |  | **Low delay B Main10** |  |  |
|  |  |  | **Over VTM-22.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | 0.06% | 0.14% | 0.44% | 93% | 94% |
| Class C | 0.03% | 0.02% | -0.25% | 94% | 96% |
| Class E | 0.04% | 0.20% | -0.28% | 97% | 96% |
| **Overall** | 0.04% | 0.12% | 0.03% | 94% | 95% |
| Class D | 0.04% | 0.23% | -0.73% | 97% | 98% |
| Class F | 0.21% | 0.32% | -0.22% | 92% | 96% |
|  |  |  |  |  |  |
|  |  |  | **Low delay P Main10** |  |  |
|  |  |  | **Over VTM-22.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | 0.04% | 0.24% | 0.07% | 94% | 95% |
| Class C | -0.02% | 0.22% | -0.11% | 97% | 99% |
| Class E | 0.11% | 0.39% | -0.25% | 96% | 95% |
| **Overall** | 0.04% | 0.27% | -0.07% | 95% | 96% |
| Class D | 0.04% | -0.35% | -0.34% | 98% | 98% |
| Class F | 0.30% | 0.12% | -0.20% | 92% | 96% |

For the high bit depth CTCs, there is no change in coding performance or run time between VTM 23.0 and VTM 22.0 for the low QP range. For the standard QP range (22-37), the simulations were run using VTM 23.0 with the merged fix for DualITree=0 (otherwise the software crashes).

Relative to VTM 22.0, some small differences can be seen:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access** | | | | | | | | | |
|  | **VTM23.0 + fix over VTM22.0, with CTC described in VTM-JVET-AA2018-STDQP** | | | | | | | | | |
|  |  |  | **wPSNR** |  |  | **PSNR** |  |  |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | -0.07% | -0.01% | 0.00% | -0.09% | -0.10% | 0.00% | -0.07% | -0.13% | 97% | 95% |
| Class H2 |  |  |  |  |  | 0.05% | 0.15% | 0.08% | 97% | 95% |
| **Overall** | -0.07% | -0.01% | 0.00% | -0.09% | -0.10% | 0.03% | 0.04% | -0.03% | 97% | 95% |
|  |  |  |  |  |  |  |  |  |  |  |
|  | **All Intra** | | | | | | | | | |
|  | **VTM23.0 + fix over VTM22.0, with CTC described in VTM-JVET-AA2018-STDQP** | | | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | -0.06% | 0.00% | 0.00% | -0.09% | -0.08% | 0.00% | -0.05% | -0.07% | 100% | 97% |
| Class H2 |  |  |  |  |  | 0.00% | 0.00% | 0.00% | 98% | 96% |
| **Overall** | -0.06% | 0.00% | 0.00% | -0.09% | -0.08% | 0.00% | -0.03% | -0.03% | 99% | 97% |

Results using HDR CTC show slight differences between **VTM 23.0** and **VTM 23.0**, as depicted in table below.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access** | | | | | | | | | |
|  | **Over VTM22.0** | | | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | -0.03% | 0.00% | 0.00% | -0.13% | 0.19% | 0.01% | -0.09% | -0.10% | 100% | 100% |
| Class H2 |  |  |  |  |  | 0.00% | -0.03% | 0.10% | 100% | 102% |
| **Overall** | 0.08% | 0.02% | 0.02% | 0.25% | 0.19% | 0.00% | -0.07% | -0.03% | 100% | 101% |
|  |  |  |  |  |  |  |  |  |  |  |
|  | **All Intra** | | | | | | | | | |
|  | **Over VTM22.0** | | | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | -0.08% | 0.00% | 0.00% | -0.09% | -0.06% | 0.00% | -0.06% | -0.04% | 102% | 100% |
| Class H2 |  |  |  |  |  | 0.00% | 0.00% | 0.00% | 100% | 100% |
| **Overall** | -0.08% | 0.00% | 0.00% | -0.09% | -0.06% | 0.00% | -0.04% | -0.02% | 101% | 100% |

The following tables show **VTM 22.0** performance over **HM 18.0** using HDR CTC:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access** | | | | | | | | | |
|  | **Over HM18.0** | | | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | -39.50% | -37.55% | -36.66% | -54.93% | -48.46% | -33.67% | -49.42% | -40.87% | 288% | 101% |
| Class H2 |  |  |  |  |  | -31.96% | -57.71% | -63.32% | 257% | 92% |
| **Overall** | -39.50% | -37.55% | -36.66% | -54.93% | -48.46% | -33.07% | -52.43% | -49.03% | 276% | 98% |
|  |  |  |  |  |  |  |  |  |  |  |
|  | **All Intra** | | | | | | | | | |
|  | **Over HM18.0** | | | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | -41.45% | -27.31% | -26.79% | -57.74% | -52.79% | -24.00% | -52.93% | -45.13% | 1537% | 117% |
| Class H2 |  |  |  |  |  | -21.76% | -47.21% | -50.57% | 1274% | 109% |
| **Overall** | -41.45% | -27.31% | -26.79% | -57.74% | -52.79% | -23.18% | -50.85% | -47.11% | 1436% | 114% |

* 1. ***Issues in VTM affecting conformance***

The following issues in VTM master branch may affect conformance:

* Missing HLS features (see sections below)
  1. ***Status of implementation of proposals of previous JVET meetings***

The following list contains all adoptions of the Q and R meetings that were not marked as merged (or submitted) or specification only change in the software coordinator tracking sheet:

* JVET-Q0112
* JVET-Q0154: Disallow mixing of GDR and IRAP (Disallow mixing of GDR with any non-GDR).
* JVET-Q0164
* JVET-Q0402
* JVET-R0178: Require that when no\_aps\_constraint\_flag is equal to 1, sps\_lmcs\_enabled\_flag and sps\_scaling\_list\_enabled\_flag shall be equal to 0
* JVET-R0221
* JVET-R0046: Change the description of the bitstream extraction process per the value of max\_tid\_il\_ref\_pics\_plus1[ ][ ] (aspect 1.2 per JVET-R0046-v4).
* JVET-R0065: Specify that GDR AUs shall be complete – i.e., all of the layers in the CVS shall have a picture in the AU (as with IRAP AUs).
* JVET-R0191: Update the range value for num\_ols\_hrd\_params\_minus1.
* JVET-R0222 aspect 1: Infer vps\_max\_sublayers\_minus1 to be equal to 6 when sps\_video\_parameter\_set\_id is equal to 0 (i.e. VPS is not present). The exact editorial expression is at the discretion of the editor.
* JVET-S0196 (JVET-S0144 item 17)
* JVET-S0227 (JVET-S0144 item 22)
* JVET-S0077 (JVET-S0139 item 5)
* JVET-S0174 aspect 2 (JVET-S0139 item 18.b)
* JVET-S0156 aspect 3 (JVET-S0139 item 21)
* JVET-S0139 item 26 (no source listed, text only?)
* JVET-S0188 aspect 1 (JVET-S0139 item 28)
* JVET-S0139 item 40 (item does not exist)
* JVET-S0042 (JVET-S0142 item 1.b)
* JVET-S0174 aspect 1 (JVET S0143 item 19)
* JVET-S0096 aspect 3 (JVET-S0140 item 10)
* JVET-S0096 aspect 4 (JVET-S0140 item 13)
* JVET-S0159 aspect 3 (JVET-S0140 item 16)
* JVET-S0171 (JVET-S0256)
* JVET-S0118 (JVET-S0141 item 7)
* JVET-S0102 (JVET-S0141 item 9.a)
* JVET-S0157 item 2 (JVET-S0141 item 13)
* JVET-S0157 item 4 (JVET-S0141 item 14)
* JVET-S0175 aspect 3 (JVET-S0141 item 16)
* JVET-S0175 aspect 1, 2 (JVET-S0141 item 17)
* JVET-S0175 aspects 4 and 5 (JVET-S0141 item 18)
* JVET-S0175 aspect 6 (JVET-S0141 item 19)
* JVET-S0198/ JVET-S0223 (JVET-S0141 item 24)
* JVET-S0173 aspect 2 (JVET-S0141 item 40.b)
* JVET-S0173 item 1 (JVET-S0141 item 51)
* JVET-S0173 item 3 (JVET-S0141 item 52)
* JVET-S0173 item 5 (JVET-S0141 item 53)
* JVET-S0173 item 6 (JVET-S0141 item 54)
* JVET-S0173 item 4 (JVET-S0141 item 56)
* JVET-S0176 item 4 (JVET-S0141 item 60)
* JVET-S0154 aspect 5 (JVET-S0141 item 68)
* JVET-S0154 aspect 6 (JVET-S0141 item 69)
* JVET-S0154 aspect 8 (JVET-S0141 item 71)
* JVET-S0095 aspect 5 (JVET-S0145 item 5)
* JVET-S0095 aspect 6 (JVET-S0145 item 6)
* JVET-S0100 aspect 1, depends on JVET-R0193 (JVET-S0147 item 2)
* FINB ballot comments
* Make high tier support up to 960.
  1. ***SEI TuC software***

Per the decision during the 32nd JVET meeting, an SEI TuC repository was created based on VTM-22.2. The repository is located at:

<https://vcgit.hhi.fraunhofer.de/jvet-tuc/VVCSoftware_VTM>

VTM-22.2-TuC-1.0 is expected to be tagged during the 33rd JVET meeting. Changes so far include:

* JVET-AF0088
* JVET-AF0167 MPIISEI message
* JVET-AF0107-TuC: Implementation of Encoder Optimization Info SEI message
* JVET-AF2032 SPTI SEI message

1. **HM related activities**

There had not been any further developments to HM during this meeting cycle.

The following MRs are pending [with status indicated]:

* Implement phase indication SEI message (JVET-AE0101) [waiting review]
* Port the Y4M support [one issue remains]
* Mark the current picture as short-term ref (for SCM) [need SCC expert reviewer]

The HM SCC (SCM) branch (HM-16.21+SCM-8.8) has not been updated for the recent HM versions. Updating SCM to, for example, HM-18.0+SCM-8.8 should be considered. It may though be helpful to move SCC related functionality into separate source files. Volunteer work towards merging the branches would be appreciated.

As reported in the previous reports, further information on lambda optimisation in HM would be appreciated, including comparison of allocation of bits within the GOP structures between HM and VTM.

Otherwise the [HEVC bug tracker](https://hevc.hhi.fraunhofer.de/trac/hevc/query?status=accepted&status=assigned&status=new&status=reopened&component=HM&col=id&col=summary&col=status&col=type&col=priority&col=milestone&col=time&col=reporter&report=16&order=time) lists:

* 43 tickets for “HM”, most of which are more than 5 years,
* 1 ticket for “HM RExt”,
* 9 tickets for “HM SCC”, most of which are at least 3 years old,
* 1 ticket for “RExt Text” (8 years old)
* 1 ticket for “SCC Text” (8 years old)
* 6 tickets for text (3-5 years old)
* 2 tickets for encoder description (3-9 years old)

Help to address these tickets would be appreciated.

1. **360Lib related activities**

The latest 360Lib software can be found at <https://vcgit.hhi.fraunhofer.de/jvet/360lib/-/tags/360Lib-13.6>

The following table is for the projection formats comparison using VTM-23.0 according to 360-degree video CTC (JVET-U2012) compared to that using VTM-22.0 (VTM-22.0 as anchor).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **End-to-end WS-PSNR** | | | **End-to-end S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | 0.00% | -0.01% | 0.05% | 0.00% | 0.00% | 0.05% |
| Class S2 | 0.08% | 0.16% | 0.34% | 0.08% | 0.15% | 0.34% |
| **Overall** | 0.03% | 0.06% | 0.17% | 0.03% | 0.06% | 0.16% |

The following table compares generalized cubemap (GCMP) coding and padded equi-rectangular projection (PERP) coding using VTM-23.0 (PERP as anchor).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **End-to-end WS-PSNR** | | | **End-to-end S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | -11.52% | -5.47% | -6.09% | -11.50% | -5.41% | -6.03% |
| Class S2 | -3.65% | 0.98% | 1.48% | -3.62% | 1.08% | 1.56% |
| **Overall** | -8.37% | -2.89% | -3.06% | -8.35% | -2.82% | -3.00% |

The following tables are for PERP and GCMP coding comparison between VTM-23.0 and HM-16.22 (HM as anchor), respectively.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **End-to-end WS-PSNR** | | | **End-to-end S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | -30.91% | -38.82% | -41.24% | -30.91% | -38.88% | -41.24% |
| Class S2 | -36.84% | -37.23% | -39.43% | -36.82% | -37.27% | -39.49% |
| **Overall** | -33.28% | -38.19% | -40.51% | -33.28% | -38.23% | -40.54% |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **End-to-end WS-PSNR** | | | **End-to-end S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | -34.64% | -40.04% | -42.14% | -34.59% | -40.02% | -42.09% |
| Class S2 | -38.34% | -38.18% | -40.06% | -38.32% | -38.19% | -40.11% |
| **Overall** | -36.12% | -39.30% | -41.31% | -36.08% | -39.29% | -41.30% |

1. **SCM related activities**

There had not been any further developments to SCC’s SCM during this meeting cycle.

1. **SHM related activities**

There had not been any further developments to SHVC’s SHM during this meeting cycle.

1. **HTM related activities**

There had not been any releases of HTM of MV-HEVC and 3D-HEVC.

One merge request is pending:

* JVET-AE0295: MV Main 10 profile support (waiting for proponent response)

The next release will include the following changes:

* JVET-Z0209: Early termination during calculating RDcost of depth

1. **HDRTools related activities**

There had not been any further developments to HDRTools during this meeting cycle.

1. **JM, JSVM, JMVM related activities**

There had not been any further developments to JM, JSVM or JMVM during this meeting cycle.

For the previous release of JM only basic testing was performed to confirm that the code still works, but performance testing is pending. Performance changes are unlikely but may occur due to bug fixes. A possible encoder/decoder mismatch was reported privately, but needs more investigation.

1. **Bug tracking**

The bug tracker for VTM and specification text is located at:

<https://jvet.hhi.fraunhofer.de/trac/vvc>

The bug tracker uses the same accounts as the HM software bug tracker. Users may need to log in again due to the different sub-domain. For spam fighting reasons account registration is only possible at the HM software bug tracker at

<https://hevc.hhi.fraunhofer.de/trac/hevc>

Bug tracking for HDRTools is located at:

<https://gitlab.com/standards/HDRTools/-/issues>

Please file all issues related to the VVC reference software and HDRTools into the appropriate bug tracker. Try to provide all the details, which are necessary to reproduce the issue. Patches for solving issues and improving the software are always appreciated.

1. **CTC alignment and merging**

There are currently 8 JVET CTC documents:

JVET-Y2010 VTM/HM 4:2:0 test conditions

JVET-Z2011 VTM/HM HDR test conditions

JVET-AA2018 VTM/HM high bit depth test conditions (without spreadsheet)

JVET-T2013 VTM non-4:2:0 test conditions

JVET-AA1100 HM non-4:2:0 test conditions

JVET-U2012 VTM 360 video test conditions

JVET-AC1009 SHVC test conditions

JVET-AC1015 SCM test conditions

JVET-AE1013 3DV test conditions

Further merging of HM RExt CTC into the appropriate VVC CTC was investigated (non 4:2:0 chroma formats), but proper comparable HM configuration files were not yet available by the beginning of this meeting.

The HM HDR test conditions require additional post-processing stages, due to WPSNR, deltaE100 and PSNRL100 values not being generated directly by the HM encoder. Merging of respective functions from VTM is recommended to streamline the process.

1. **Guidelines for reference software development**

No further work was conducted on Guidelines for VVC and HEVC reference software development and documents JVET-AC2003 and JVET-AC1001 remain current.

1. **Recommendations**

The AHG recommends to:

* Continue to develop reference software.
* Improve documentation, especially the software manual.
* Encourage people to test VTM and other reference software more extensively outside of common test conditions.
* Encourage people to report all (potential) bugs that they are finding.
* Encourage people to submit bit-streams/test cases that trigger bugs in VTM and other reference software.
* Encourage people to submit non-normative changes that either reduce encoder run time without significantly sacrificing compression performance or improve compression performance without significantly increasing encoder run time.
* Design and add configuration files to the VTM software for testing of HLS features.
* Review VTM-related contributions and determine whether features should be added (or removed) from the software.
* Continue to investigate the merging of branches.
* Continue to investigate merging of CTC documents.
* Verify correctness of CTC documents and issue updates as appropriate
* Keep common test conditions aligned for the different standards.

[JVET-AG0004](https://jvet-experts.org/doc_end_user/current_document.php?id=13778) JVET AHG report: Test material and visual assessment (AHG4) [V. Baroncini, T. Suzuki, M. Wien (co-chairs), W. Husak, S. Iwamura, P. de Lagrange, S. Liu, X. Meng, S. Puri, A. Segall, S. Wenger (vice-chairs)]

1. **Activities**
   1. ***Verification tests for VVC multilayer coding***

The output document JVET-AF2021 of the previous meeting documenting the updated version of the verification test plan was produced and uploaded to the document system.

The output document JVET-AF2033 of the previous meeting reporting the results of the demonstration of content layering using VVC multilayer coding was produced and uploaded to the document system.

* 1. ***Plan for subjective quality testing of FGC SEI message***

In this meeting period, candidate sequences for visual testing of filmgrain characteristics proposed at the previous meeting have been made available on the ftp site in the directory ./ahg/viewingtests/filmgrain/JVET-AF0262.

* 1. ***Test sequences***

The test sequences used for CfP/CTC are available on <ftp://jvet@ftp.ient.rwth-aachen.de> in directory “/ctc” (accredited members of JVET may contact the JVET chair for login information).

Due to copyright restrictions, the JVET database of test sequences is only available to accredited members of JVET (i.e., members of ISO/IEC MPEG and ITU-T VCEG).

The current JVET ftp site may be replaced by a new system in the near future. Potential solutions and a transition plan are currently under discussion.

1. **Related contributions**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| [JVET number](https://jvet-experts.org/doc_end_user/current_meeting.php?id_meeting=195&type_order=&sql_type=document_number) | MPEG number | [Created](https://jvet-experts.org/doc_end_user/current_meeting.php?id_meeting=195&type_order=&sql_type=document_date_time) | First upload | [Last upload](https://jvet-experts.org/doc_end_user/current_meeting.php?id_meeting=195&type_order=&sql_type=upload_document_date_time) | [Title](https://jvet-experts.org/doc_end_user/current_meeting.php?id_meeting=195&type_order=&sql_type=title) | [Source](https://jvet-experts.org/doc_end_user/current_meeting.php?id_meeting=195&type_order=&sql_type=authors) |
| [JVET-AG0004](https://jvet-experts.org/doc_end_user/current_document.php?id=13778) | m66131 | 2024-01-11 13:02:08 |  |  | JVET AHG report: Test material and visual assessment (AHG4) | V. Baroncini, T. Suzuki, M. Wien (co-chairs), W. Husak, S. Iwamura, P. de Lagrange, S. Liu, X. Meng, S. Puri, A. Segall, S. Wenger (vice-chairs) |
| [JVET-AG0071](https://jvet-experts.org/doc_end_user/current_document.php?id=13627) | m65960 | 2024-01-10 06:49:08 | 2024-01-10 15:53:19 | 2024-01-10 15:53:19 | Informal Subjective Evaluation of Low Complexity Enhancement Video Codec (LCEVC) with VVC on SDR UHD (4K) Content | O. Chubach, H.-H. Chen, C.-Y. Chen, T.-D. Chuang, Y.-W. Chen, C.-W. Hsu, Y.-W. Huang, S.-M. Lei (MediaTek) |
| [JVET-AG0224](https://jvet-experts.org/doc_end_user/current_document.php?id=13797) | m66167 | 2024-01-11 19:47:04 |  |  | Response to JEVT-AG0071 on Informal Subjective Evaluation of Low Complexity Enhancement Video Codec (LCEVC) with VVC on SDR UHD (4K) Content | [Lorenzo Ciccarelli](mailto:lorenzo.ciccarelli@v-nova.com), [Simone Ferrara](mailto:simone.ferrara@v-nova.com) |
| [JVET-AG0228](https://jvet-experts.org/doc_end_user/current_document.php?id=13801) | m66244 | 2024-01-12 16:20:30 | 2024-01-15 16:48:27 | 2024-01-16 14:37:49 | AHG4/AHG13: source for scanned film test sequences | [P. de Lagrange (InterDigital)](mailto:philippe.delagrange@interdigital.com) |

1. **Recommendations**

The AHG recommends:

* To review and consider JVET-AE0228 in the development of the draft plan for subjective quality testing of FGC SEI message.
* To review document JVET-AF0071 and JVET-AG0224 in a joint meeting with WG 4 and AG 5.
* To collect volunteers to conduct further verification tests and subjective quality tests.
* To continue to discuss and to update the non-finalized categories of the verification test plan and subjective quality test plan for FGS, including those which have not been addressed yet.
* To collect volunteers to actively contribute to the verification test development.
* To review the set of available test sequences for the verification tests as well as subjective quality tests and potentially collect more test sequences with a variety of content.
* To continue to collect new test sequences available for JVET with licensing statement.

[JVET-AG0005](https://jvet-experts.org/doc_end_user/current_document.php?id=13779) JVET AHG report: Conformance testing (AHG5) [I. Moccagatta (chair), F. Bossen, K. Kawamura, P. de Lagrange, T. Ikai, S. Iwamura, H.-J. Jhu, S. Paluri, K. Sühring, Y. Yu (vice chairs)]

1. **Activities (fix numbering)**

The AHG communication is conducted through the main JVET reflector, jvet@lists.rwth-aachen.de, with [AHG5] in message headers. However, no correspondence marked as AHG5 was sent between the 32nd and 33rd meetings.

1. **Timeline**

The progress on the Conformance testing specification is proceeding per the timeline below:

* **VVCv1 conformance:**
  + ISO/IEC FDIS 23090-15 issued from 2021-10 meeting, FDIS registered for formal approval 2022-07-11, FDIS ballot closed 2022-11-04, standard published 2022-11-24
  + H.266.1 V1 Consent 2022-01-28, Last Call began 2022-04-01, Approved 2022-04-29, pre-published 2022-05-17, standard published 2022-07-12.
* **VVCv2 conformance:**
  + ISO/IEC 23090-15/Amd.1 CDAM: 2021-10
  + ISO/IEC 23090-15/Amd.1 DAM: 2022-01
  + DAM ballot closed 2022-11-15
  + ISO/IEC FDIS 23090-15:202x 2nd edition text output of 2023-01, preparation delayed to 2023-09
  + H.266.1 V2 forwarded by JVET and Q6/16 for ITU-T Consent: 2023-07
  + H.266.1 V2 approved 2023-09-13
  + H.266.2 V2 pre-published 2023-10-06

1. **Status on bitstream submission**

The status at the time of preparation of this report is as follows:

* conformance bitstreams for VVC:
  + 104 bitstream categories have been identified
  + At least one bitstream has been submitted in each identified category
  + 283 total bitstreams have been provided, checked, and made available
  + No changes in bitstream between 32nd and 33rd meeting.
* conformance bitstreams for VVC operation range extensions:
  + 57 bitstream categories have been identified
  + 1 bitstream of 1 identified category has been re-generated
  + 128 bitstreams of 57 identified categories have been cross-checked and uploaded.
  + No changes between 32nd and 33rd meeting.

1. **Activities and Discussion**

The AHG activities are on schedule with the preliminary timeline shown in section 2.

VVC activities:

No change in bitstreams and/or packages, 2nd edition packages are available at <https://www.itu.int/wftp3/av-arch/jvet-site/bitstream_exchange/VVC2ndEd/>

VVC operation range extensions activities:

No change in bitstreams and/or packages, 2nd edition packages are available at <https://www.itu.int/wftp3/av-arch/jvet-site/bitstream_exchange/VVC2ndEd/>

VVC Multilayer activities:

The 7 additional conformance bitstreams for VVC multilayer configurations in JVET-AE2028 are decoded correctly using VTM-23.0 and are available in <https://www.itu.int/wftp3/av-arch/jvet-site/bitstream_exchange/VVCMultilayer/under_test/VTM-23.0/>.

Volunteers to cross check these 7 conformance bitstreams for VVC multilayer configurations have been identified and cross-check is ongoing.

HEVC Multiview supporting extended bit depth activities:

A merge request (MR) implementing the HEVC Multiview Main 10 profiles in JVET-AA1011 has been submitted in <https://vcgit.hhi.fraunhofer.de/jvet/HTM/-/merge_requests/5>

An implementation for the other HEVC Multiview profiles in JVET-AA1011 (Multiview Monochrome, and Multiview Monochrome 10 profiles) is still in progress.

Cross-checking of the 4 HEVC Multiview supporting extended bit depth (Multiview Main 10) bitstreams provided in AD0232 and AE0295 is contingent to the merging of the above MR. Review of the MR is on-going.

The regular JVET e-mail reflector was used for discussions ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de)).

The AHG5 chairs and JVET chairs can be reached at [jvet-conformance@lists.rwth-aachen.de](mailto:jvet-conformance@lists.rwth-aachen.de). Participants should not subscribe to this list but may send emails to it.

1. **Contributions**

JVET-AG0201 – On MV-HEVC Profiles [Y.-K. Wang, H. Liu, L. Zhang, S. Jiao, C. Hu, J. Cui, G. Xu , A. M. Tourapis, D. Podborski, S. Paluri].

1. **Ftp site information**

The procedure to exchange the bitstream (ftp cite, bitstream files, etc.) is specified in Sec 2 “Procedure” of [JVET-R2008](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=8861). The ftp and http sites for downloading bitstreams are

* VVC:

<ftp://ftp3.itu.int/jvet-site/bitstream_exchange/VVC>

<https://www.itu.int/wftp3/av-arch/jvet-site/bitstream_exchange/VVC/>

* VVC operation range extensions:

[ftp://ftp3.itu.int/jvet-site/bitstream\_exchange/VVCv2](ftp://ftp3.itu.int/jvet-site/bitstream_exchange/VVCv2/draft_conformance/draft)

<https://www.itu.int/wftp3/av-arch/jvet-site/bitstream_exchange/VVCv2>

The ftp site for uploading bitstream file is as follows.

<ftp://ftp3.itu.int/jvet-site/dropbox/>

(user id: avguest, passwd: Avguest201007)

If using FileZilla, the following configuration is suggested:

Graphical user interface, text, application, email

Description automatically generated

In the Filezilla Edit 🡪 Settings 🡪 Connection menu, it may also be necessary to set the minimum TLS level to 1.0.

1. **Recommendations**

The AHG recommends the following:

* Proceed with the cross-checking of the additional conformance bitstreams for VVC multilayer configurations (JVET-AE2028).
* Maintain and update the conformance bitstream database and contribute to report problems in JVET document 1004.
* Continue the generation, cross-checking, and documentation of the conformance streams for the HEVC multiview profiles supporting extended bit depth (JVET-AA1011).

It is noted that availability of software (HTM) and verified conformance streams for the new HEVC Multiview profiles by the next meeting would be important to progress those profiles into the standard.

[JVET-AG0006](https://jvet-experts.org/doc_end_user/current_document.php?id=13780) JVET AHG report: ECM software development (AHG6) [V. Seregin (chair), J. Chen, R. Chernyak, F. Le Léannec, K. Zhang (vice-chairs)]

1. **Software development**

ECM software repository is located at <https://vcgit.hhi.fraunhofer.de/ecm/ECM>.

ECM software is based on VTM-10.0 with enabled MCTF including the update from JVET-V0056, and GOP32, which is very close to VTM-11.0.

VTM-11.0ecm anchor <https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tree/VTM11_ANC>is used for ECM performance evaluation.

The following changes were integrated into ECM-10.1:

Fix: high QP for AA0093 (MR 519)

Fix: reconstruction buffer copied when BIF is not performed for a sub-TU (MR 522)

Fix for uninitialized values usage reported by valgrind (MR 524)

Reduce MCTF peak memory allocation by computing applyMotion in place

Fix: memory reduction: use actual CTU size instead of MAX\_CU\_SIZE in encoder (MR 520)

ALF optimization - compute only Upper triangular part of covariance matrix (MR 521)

PU memory reduction by moving intraMPM member from PU IntraPredictionData to IntraPrediction (MR 518)

Fix: CU memory reduction. move some members from CU to intra search (MR 517)

Refactor picture buffer to allocate buffers for current frame and not for the entire GOP (MR 523)

CS memory reduction by allocating temporary CodingStructure data only for the current picture (MR 525)

Shrink some deblocking buffers and reduce length of memset() initializations (MR 526)

Memory reduction for sign prediction (MR 527)

Bugfix for multi-model BVG-CCCM (JVET\_AE0100\_BVGCCCM) (MR 530)

Use transposed buffers (MR 529)

Refactor Cabac ctx (MR 528)

Fix: fill picture level recon buffer in TU reconstruction code for BIF (MR 531)

Fix for derivation of cost in IntraTMP CandList (MR 532)

Fix the problem with disabling AE0159 macro (MR 533)

Fix unreleased RPR temp buffer (MR 534)

The following adopted aspects were integrated into ECM-11.0:

JVET-AF0133: Retraining I-slice context models (Test 6.2) (MR 537)

JVET-AF0201: Incorporate decoder memory measurement (MR 547)

JVET-AF0111: Updates the threshold for MTT modes early termination (MR 538)

JVET-AF0237: SAO/ALF encoder memory reduction (lossless) (MR 539)

JVET-AF0163: TM based subblock motion refinement (Test 3.4a) (MR 546)

JVET-AF0128: LIC flag derivation for merge candidates with template costs (Test 3.2) (MR 544)

JVET-AF0190: Enabling template-based reordering tools and LIC for scaled pictures in the RPR (Test 4.1b) (MR 542)

JVET-AF0197: Luma Residual Tap in CCALF (MR 548)

JVET-AF0079: IntraTMP block vector storing (Test 2.6c) (MR 535)

JVET-AF0177: Change data type of ALF/CCALF covariance from double to float (MR 541)

JVET-AF0073: Inter cross-component prediction merge (Test 3.1d) (MR 536)

JVET-AF0059: AHG12: Fix to interpolation filter for intra prediction (MR 554)

JVET-AF0066: Enable DBV mode in single tree configuration (MR 552)

JVET-AF0159: Affine subblock BDOF refinement (MR 550)

JVET-AF0112: Dynamic Scaling of Bilateral Filter (BIF) (Test 5.1a) (MR 556)

JVET-AF0101: Lossless code optimization by moving initialization to init() function (MR 555)

JVET-AF0057: Encoder only method for robust MV derivation in DMVR (MR 557)

Optimizations:

Memory reduction by allocating picture margin according to whether RPR is used (MR 561)

Bug fixes:

Fixes to DebugStream tool (MR 562, 566, 569, 570)

Fix gcc 13.2 compiling issues (MR 560)

Fix SAO/ALF memory reduction when RPR is enabled (MR 563)

Fix tracing (MR 565)

Fix: out of bounds access in simdFilter (MR 572)

Fix: out of bounds access in BI-GPM (MR 571)

Fix: valgrind error in Affine Merge candidates management in DecCu::xDeriveCUMV (MR 574)

Fix: inconsistent block stat by JVET-AD0222 when virtual boundary is enabled (MR 575)

The following adopted aspects were integrated into VTM-11ecm anchor:

JVET-AF0165: align VTM-11.0ecm CTCs to ECM CTCs for low-delay (MR 545)

JVET-AF0111: MTT split modes early termination threshold is updated and enabled in RA and LDB CTC (MR 540)

ECM-10.1, ECM-11.0, and VTM-11ecm11.0 were tagged on November 21, 2023.

* 1. ***CTC* Performance**

In this section, ECM test results following ECM CTC configuration descried in JVET-AF2017 are summarized.

ECM-10.1 performance over ECM-10.0 anchor is summarized in the tables below.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | | |
|  | **Over ECM-10.0** | | | | | |
|  | Y | U | V | EncT | DecT | VmPeak |
| Class A1 | 0.00% | 0.01% | -0.03% | 98.0% | 98.4% | 88.2% |
| Class A2 | -0.01% | -0.15% | -0.16% | 97.7% | 97.8% | 88.0% |
| Class B | -0.01% | -0.15% | -0.16% | 93.4% | 94.3% | 57.5% |
| Class C | 0.00% | 0.03% | 0.03% | 91.5% | 93.0% | 36.5% |
| Class E | 0.01% | 0.06% | -0.15% | 93.7% | 92.6% | 44.6% |
| **Overall** | 0.00% | -0.05% | -0.09% | 94.5% | 94.9% | 57.5% |
| Class D | 0.00% | -0.09% | -0.06% | 92.7% | 92.4% | 29.9% |
| Class F | -0.09% | -0.10% | -0.13% | 99.0% | 99.2% | 44.8% |
| Class TGM | -0.21% | -0.24% | -0.27% | 100.8% | 98.9% | 56.9% |
|  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | |
|  | **Over ECM-10.0** | | | | | |
|  | Y | U | V | EncT | DecT | VmPeak |
| Class A1 | 0.02% | -0.21% | -0.01% | 97.7% | 98.8% | 78.5% |
| Class A2 | 0.01% | -0.17% | 0.00% | 96.2% | 97.8% | 78.1% |
| Class B | -0.05% | -0.01% | -0.20% | 94.5% | 98.4% | 60.1% |
| Class C | 0.01% | 0.06% | -0.07% | 92.6% | 96.3% | 40.4% |
| Class E |  |  |  |  |  |  |
| **Overall** | -0.01% | -0.07% | -0.09% | 94.9% | 97.8% | 60.1% |
| Class D | -0.01% | -0.23% | -0.16% | 93.0% | 96.0% | 32.2% |
| Class F | 0.03% | 0.10% | 0.17% | 101.8% | 101.0% | 51.5% |
| Class TGM | -0.11% | -0.16% | -0.06% | 104.0% | 103.1% | 67.5% |
|  |  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | | |
|  | **Over ECM-10.0** | | | | | |
|  | Y | U | V | EncT | DecT | VmPeak |
| Class A1 |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |
| Class B | -0.01% | 0.18% | 0.26% | 95.5% | 94.2% | 58.9% |
| Class C | -0.04% | -0.60% | -0.32% | 91.4% | 93.5% | 37.8% |
| Class E | -0.03% | 1.00% | -1.48% | 97.8% | 96.4% | 45.9% |
| **Overall** | -0.02% | 0.13% | -0.37% | 94.7% | 94.5% | 47.8% |
| Class D | 0.09% | -0.25% | -0.55% | 93.1% | 93.8% | 30.6% |
| Class F | 0.26% | 0.20% | 0.31% | 98.3% | 98.2% | 49.1% |
| Class TGM | -0.03% | 0.02% | -0.05% | 99.7% | 96.7% | 63.2% |
|  |  |  |  |  |  |  |
|  | **Low delay P Main 10** | | | | | |
|  | **Over ECM-10.0** | | | | | |
|  | Y | U | V | EncT | DecT | VmPeak |
| Class A1 |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |
| Class B | 0.02% | -0.23% | 0.09% | 96.4% | 97.7% | 58.7% |
| Class C | 0.02% | 0.05% | -0.03% | 94.5% | 96.6% | 37.5% |
| Class E | -0.09% | 0.75% | -1.10% | 96.8% | 94.3% | 45.5% |
| **Overall** | -0.01% | 0.11% | -0.25% | 95.9% | 96.4% | 47.4% |
| Class D | 0.05% | -0.51% | 0.17% | 94.7% | 95.8% | 30.4% |
| Class F | 0.18% | -0.20% | 0.39% | 98.4% | 99.1% | 48.2% |
| Class TGM | 0.05% | 0.15% | 0.05% | 99.4% | 96.4% | 62.8% |

Next tables show ECM-11.0 performance over ECM-10.1 anchor.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | | | |
|  | **Over ECM-10.1** | | | | | | |
|  | Y | U | V | EncT | DecT | EncVmPeak | DecVmPeak |
| Class A1 | -0.30% | -0.01% | -0.11% | 94.6% | 94.6% | 78.2% |  |
| Class A2 | -0.31% | -0.13% | -0.28% | 96.6% | 94.7% | 78.7% |  |
| Class B | -0.32% | -0.28% | -0.15% | 100.6% | 96.4% | 60.7% |  |
| Class C | -0.38% | -0.33% | -0.27% | 96.8% | 88.6% | 79.6% |  |
| Class E | -0.50% | -0.60% | -0.17% | 99.5% | 97.7% | 70.2% |  |
| **Overall** | -0.36% | -0.27% | -0.20% | 97.9% | 94.2% | 71.9% |  |
| Class D | -0.34% | -0.17% | -0.34% | 99.3% | 90.6% | 90.9% |  |
| Class F | -0.43% | -0.65% | -0.50% | 100.8% | 92.7% | 70.1% |  |
| Class TGM | -0.21% | -0.33% | -0.18% | 102.2% | 97.9% | 60.3% |  |
|  |  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | | |
|  | **Over ECM-10.1** | | | | | | |
|  | Y | U | V | EncT | DecT | EncVmPeak | DecVmPeak |
| Class A1 | -0.32% | -0.97% | -1.19% | 102.4% | 102.7% | 86.3% |  |
| Class A2 | -0.65% | -1.54% | -2.24% | 104.4% | 109.5% | 87.0% |  |
| Class B | -0.30% | -1.63% | -1.21% | 106.4% | 107.0% | 79.7% |  |
| Class C | -0.38% | -1.09% | -1.05% | 107.0% | 109.1% | 85.8% |  |
| Class E |  |  |  |  |  |  |  |
| **Overall** | -0.39% | -1.34% | -1.37% | 105.4% | 107.2% | 84.1% |  |
| Class D | -0.29% | -0.82% | -0.64% | 106.1% | 106.8% | 90.7% |  |
| Class F | -0.60% | -1.18% | -1.15% | 103.2% | 108.0% | 83.0% |  |
| Class TGM | -0.14% | -0.45% | -0.64% | 94.5% | 95.8% | 80.3% |  |
|  |  |  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | | | |
|  | **Over ECM-10.1** | | | | | | |
|  | Y | U | V | EncT | DecT | EncVmPeak | DecVmPeak |
| Class A1 |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |
| Class B | -0.08% | -5.40% | -5.43% | 95.9% | 101.6% | 77.5% |  |
| Class C | -0.23% | -2.93% | -2.92% | 99.5% | 101.1% | 86.4% |  |
| Class E | -0.12% | -4.41% | -2.61% | 92.6% | 97.2% | 80.9% |  |
| **Overall** | -0.14% | -4.33% | -3.89% | 96.2% | 100.3% | 81.2% |  |
| Class D | -0.16% | -4.60% | -3.34% | 101.7% | 103.3% | 94.1% |  |
| Class F | -0.53% | -3.94% | -4.09% | 98.6% | 101.3% | 83.4% |  |
| Class TGM | -0.14% | -2.26% | -2.46% | 96.7% | 106.9% | 78.9% |  |
|  |  |  |  |  |  |  |  |
|  | **Low delay P Main 10** | | | | | | |
|  | **Over ECM-10.1** | | | | | | |
|  | Y | U | V | EncT | DecT | EncVmPeak | DecVmPeak |
| Class A1 |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |
| Class B | -0.03% | -5.83% | -4.68% | 96.4% | 100.1% | 76.8% |  |
| Class C | -0.17% | -3.40% | -3.45% | 99.0% | 108.6% | 87.4% |  |
| Class E | -0.29% | -4.28% | -2.81% | 94.0% | 107.8% | 81.1% |  |
| **Overall** | -0.14% | -4.64% | -3.80% | 96.6% | 104.8% | 81.3% |  |
| Class D | -0.10% | -5.00% | -4.71% | 101.6% | 108.1% | 94.2% |  |
| Class F | -0.95% | -3.78% | -4.60% | 101.3% | 110.8% | 83.6% |  |
| Class TGM | 0.06% | -2.38% | -2.78% | 97.7% | 98.7% | 79.3% |  |

The below tables show ECM-11.0 performance comparing to VTM-11.0ecm11.0 anchor.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | | | |
|  | **Over VTM-11.0ecm11.0** | | | | | | |
|  | Y | U | V | EncT | DecT | EncVmPeak | DecVmPeak |
| Class A1 | -10.81% | -21.80% | -29.33% | 859.1% | 400.6% |  |  |
| Class A2 | -16.95% | -28.77% | -31.78% | 860.2% | 405.4% |  |  |
| Class B | -11.31% | -26.84% | -25.03% | 784.4% | 411.0% |  |  |
| Class C | -11.41% | -16.72% | -17.24% | 810.7% | 381.9% |  |  |
| Class E | -15.01% | -24.77% | -23.07% | 750.1% | 431.9% |  |  |
| **Overall** | -12.81% | -23.73% | -24.81% | 808.6% | 405.0% |  |  |
| Class D | -9.48% | -14.56% | -14.65% | 795.1% | 383.1% |  |  |
| Class F | -26.60% | -35.72% | -35.58% | 513.6% | 392.7% |  |  |
| Class TGM | -39.53% | -47.43% | -46.69% | 457.3% | 461.2% |  |  |
|  |  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | | |
|  | **Over VTM-11.0ecm11.0** | | | | | | |
|  | Y | U | V | EncT | DecT | EncVmPeak | DecVmPeak |
| Class A1 | -22.47% | -27.84% | -36.88% | 875.4% | 769.1% |  |  |
| Class A2 | -26.01% | -36.30% | -40.11% | 824.7% | 958.3% |  |  |
| Class B | -20.80% | -35.53% | -33.16% | 714.1% | 774.7% |  |  |
| Class C | -22.22% | -27.13% | -27.05% | 748.2% | 829.1% |  |  |
| Class E |  |  |  |  |  |  |  |
| **Overall** | -22.56% | -31.91% | -33.67% | 775.1% | 821.9% |  |  |
| Class D | -23.10% | -28.71% | -28.75% | 768.1% | 898.2% |  |  |
| Class F | -28.70% | -37.47% | -37.77% | 614.7% | 505.2% |  |  |
| Class TGM | -36.82% | -43.33% | -43.43% | 566.6% | 413.2% |  |  |
|  |  |  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | | | |
|  | **Over VTM-11.0ecm11.0** | | | | | | |
|  | Y | U | V | EncT | DecT | EncVmPeak | DecVmPeak |
| Class A1 |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |
| Class B | -17.76% | -41.31% | -38.53% | 661.6% | 649.7% |  |  |
| Class C | -19.49% | -31.72% | -31.66% | 618.2% | 678.1% |  |  |
| Class E | -16.28% | -30.12% | -30.39% | 587.1% | 406.5% |  |  |
| **Overall** | -17.97% | -35.32% | -34.21% | 627.8% | 586.1% |  |  |
| Class D | -21.22% | -34.09% | -33.02% | 659.2% | 772.5% |  |  |
| Class F | -24.87% | -39.63% | -38.99% | 575.6% | 491.8% |  |  |
| Class TGM | -35.47% | -44.37% | -44.64% | 538.8% | 424.4% |  |  |
|  |  |  |  |  |  |  |  |
|  | **Low delay P Main 10** | | | | | | |
|  | **Over VTM-11.0ecm11.0** | | | | | | |
|  | Y | U | V | EncT | DecT | EncVmPeak | DecVmPeak |
| Class A1 |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |
| Class B | -16.37% | -40.73% | -37.81% | 592.0% | 625.0% |  |  |
| Class C | -18.40% | -30.91% | -30.87% | 545.9% | 688.5% |  |  |
| Class E | -15.76% | -29.13% | -29.84% | 571.9% | 461.5% |  |  |
| **Overall** | -16.89% | -34.56% | -33.51% | 571.2% | 598.4% |  |  |
| Class D | -21.11% | -33.25% | -33.49% | 588.5% | 750.2% |  |  |
| Class F | -24.13% | -39.44% | -39.00% | 609.4% | 513.2% |  |  |
| Class TGM | -32.65% | -42.90% | -43.19% | 600.4% | 387.9% |  |  |

1. **ECM memory consumption**

ECM encoder memory consumption (VmPeak, GiB) is provided in ECM encoder log files and is summarized in the table below as maximum class-wise consumption rounded up to GiB.

|  |  |  |  |
| --- | --- | --- | --- |
|  | AI | RA | LB |
| class A (A1& A2) | 7 | 14 |  |
| class B | 3 | 5 | 4 |
| class C | 2 | 2 | 2 |
| class D | 2 | 2 | 2 |
| class E | 2 |  | 3 |
| class F | 3 | 5 | 4 |
| class TGM | 3 | 7 | 6 |

It is encouraged to care about memory allocation when developing and integrating tools into ECM. In particular, it is strongly recommended to re-use already existed memory wherever possible, rather than systematically allocating new memory, and allocate only the required amount of memory if new memory is needed.

1. **Recommendations**

The AHG recommends to:

* Continue to develop ECM software.
* Improve the software documentation.
* Encourage people to report all (potential) bugs that they are finding using GitLab Issues functionality <https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/issues>.
* Encourage people to submit merge requests fixing identified bugs.
* Encourage people to continue working on ECM memory consumption reduction.
* Encourage people to continue working on speeding up ECM encoder to reduce the simulation time.

It was asked what the memory consumption is compared to VTM. This was not reported, but it is estimated to be larger by a factor of approximately 2x for the current ECM software implementation.

[JVET-AG0007](https://jvet-experts.org/doc_end_user/current_document.php?id=13781) JVET AHG report: ECM tool assessment (AHG7) [X. Li (chair), L.-F. Chen, Z. Deng, J. Gan, E. François, H.-J. Jhu, X. Li, H. Wang (vice chairs)]

1. **Group off tests**
   1. ***Test settings and crosschecking***

The same four groups were used in this meeting cycle.

* Group 1: Inter template matching tools
* Group 2: Coding tools that interleave the (merge/skip/AMVP/subblock/IBC/etc) list derivation with the intra prediction/reconstruction process
* Group 3: Intra and IBC template matching (with search) related tools
* Group 4: Tools that require more processing on the neighboring reconstructed samples than VVC

Five group-off tests were performed and crosschecked on top of ECM-11. The two anchors are ECM 11.0 and VTM-11ECM11.0. The cfg files used are also attached with this report.

The testers and crosscheckers are summarized in the table below. All the tests below have been crosschecked.

|  |  |  |
| --- | --- | --- |
| Tests | Crosschecker | Tester |
| Group 1 off | Jonathan Gan (v-jonathan.gan@oppo.com) | Charles Salmon-Legagneur (charles.salmon-legagneur@interdigital.com) |
| Group 2 off | Jonathan Gan (v-jonathan.gan@oppo.com) | Xinwei Li  ([sid.lxw@alibaba-inc.com](mailto:sid.lxw@alibaba-inc.com)) |
| Group 3 off | Xiang Li (xlxiangli@google.com) | Zhipin Deng (zhipin.deng@bytedance.com) |
| Group 4 off | Xiang Li (xlxiangli@google.com) | Hong-Jheng Jhu (jhuhong-jheng@kwai.com) |
| Group 1-4 off | Lien-Fei Chen (lienfei.chen@global.tencent.com),  Xiang Li (xlxiangli@google.com) | Hongtao Wang  (hongtaow@qti.qualcomm.com) |

* 1. ***Group 1 off***

Group 1 includes inter template matching tools. The attached offgroup1.cfg was used in addition to ECM CTC settings.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | | | | |
|  | **Over ECM-11** | | | | | | **Over VTM-11ecm11** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 | 0.00% | 0.00% | 0.00% | 100.1% | 100.4% | 100% | -10.82% | -21.80% | -29.33% | 815.3% | 443.2% |
| Class A2 | 0.00% | 0.00% | 0.00% | 100.1% | 100.4% | 100% | -16.95% | -28.77% | -31.78% | 795.4% | 443.5% |
| Class B | 0.00% | 0.00% | 0.00% | 100.0% | 99.8% | 100% | -11.32% | -26.84% | -25.03% | 748.2% | 447.8% |
| Class C | 0.00% | 0.00% | 0.00% | 100.1% | 99.9% | 100% | -11.41% | -16.73% | -17.24% | 759.8% | 383.1% |
| Class E | -0.01% | -0.01% | -0.01% | 99.9% | 99.4% | 100% | -15.02% | -24.78% | -23.08% | 708.7% | 461.0% |
| **Overall** | 0.00% | 0.00% | 0.00% | 100.0% | 99.9% | 100% | -12.81% | -23.73% | -24.82% | 762.4% | 433.2% |
| Class D | -0.03% | -0.02% | -0.02% | 99.6% | 99.6% | 100% | -9.51% | -14.58% | -14.67% | 754.6% | 383.4% |
| Class F | -0.01% | -0.01% | -0.01% | 100.0% | 100.0% | 100% | -26.60% | -35.73% | -35.59% | 538.0% | 456.1% |
| Class TGM | 0.00% | 0.00% | 0.00% | 100.0% | 100.0% | 100% | -39.54% | -47.43% | -46.69% | 449.2% | 545.3% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | | | | | | |
|  | **Over ECM-11** | | | | | | **Over VTM-11ecm11** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 | 4.03% | 3.64% | 3.74% | 90.3% | 82.3% | 100% | -19.34% | -25.32% | -34.42% | 781.0% | 715.2% |
| Class A2 | 4.63% | 4.80% | 4.92% | 88.1% | 75.6% | 100% | -22.60% | -33.25% | -37.16% | 714.2% | 841.9% |
| Class B | 3.81% | 3.76% | 3.57% | 84.8% | 76.0% | 100% | -17.76% | -33.07% | -30.74% | 589.7% | 651.0% |
| Class C | 4.22% | 4.23% | 4.44% | 81.4% | 68.7% | 99% | -18.90% | -24.02% | -23.78% | 590.5% | 552.5% |
| Class E |  |  |  |  |  |  |  |  |  |  |  |
| **Overall (Ref)** | 4.13% | 4.07% | 4.11% | 85.6% | 75.1% | 100% | -19.35% | -29.15% | -30.90% | 648.4% | 668.5% |
| Class D | 3.43% | 3.79% | 3.74% | 81.1% | 67.2% | 100% | -20.41% | -25.94% | -26.00% | 630.5% | 659.2% |
| Class F | 3.18% | 3.20% | 3.14% | 90.2% | 80.9% | 100% | -26.35% | -35.42% | -35.76% | 567.8% | 429.3% |
| Class TGM | 3.78% | 3.63% | 3.61% | 91.2% | 84.6% | 100% | -34.42% | -41.23% | -41.35% | 562.6% | 441.6% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | | | | | | | |
|  | **Over ECM-11** | | | | | | **Over VTM-11ecm11** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |  |  |  |  |
| Class B | 4.69% | 5.49% | 5.37% | 75.7% | 66.8% | 100% | -13.88% | -38.16% | -35.35% | 507.7% | 441.8% |
| Class C | 4.93% | 4.87% | 5.00% | 71.7% | 57.0% | 100% | -15.51% | -28.42% | -28.22% | 442.1% | 406.2% |
| Class E | 4.57% | 5.49% | 5.24% | 80.7% | 76.3% | 100% | -12.47% | -26.38% | -26.80% | 474.4% | 326.1% |
| **Overall (Ref)** | 4.74% | 5.28% | 5.21% | 75.5% | 65.5% | 100% | -14.07% | -31.97% | -30.84% | 476.7% | 398.2% |
| Class D | 4.05% | 3.66% | 3.90% | 74.7% | 55.6% | 100% | -18.01% | -31.59% | -30.40% | 486.8% | 441.7% |
| Class F | 4.59% | 4.06% | 3.88% | 84.6% | 73.5% | 99% | -21.34% | -37.10% | -36.54% | 498.2% | 324.2% |
| Class TGM | 5.39% | 5.15% | 5.13% | 87.7% | 78.6% | 100% | -31.92% | -41.32% | -41.60% | 483.1% | 325.4% |

* 1. ***Group 2 off***

Group 2 includes coding tools that interleave the (merge/skip/AMVP/subblock/IBC/etc) list derivation with the intra prediction/reconstruction process. The attached offgroup2.cfg was used in addition to ECM CTC settings.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | | | | |
|  | **Over ECM-11** | | | | | | **Over VTM-11ecm11** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 | 0.39% | 0.67% | 0.67% | 92.8% | 95.9% | 101% | -10.45% | -21.30% | -28.83% | 753.8% | 423.9% |
| Class A2 | 0.53% | 0.77% | 0.80% | 91.7% | 96.3% | 102% | -16.50% | -28.24% | -31.25% | 726.4% | 425.3% |
| Class B | 0.51% | 0.76% | 0.79% | 90.4% | 94.8% | 103% | -10.85% | -26.24% | -24.41% | 674.6% | 425.4% |
| Class C | 0.33% | 0.43% | 0.64% | 89.3% | 94.5% | 100% | -11.11% | -16.39% | -16.72% | 679.9% | 360.8% |
| Class E | 0.74% | 0.96% | 0.99% | 91.2% | 96.0% | 101% | -14.37% | -24.03% | -22.32% | 646.2% | 446.9% |
| **Overall** | 0.49% | 0.71% | 0.77% | 90.9% | 95.4% | 101% | -12.37% | -23.19% | -24.23% | 692.0% | 413.2% |
| Class D | 0.28% | 0.55% | 0.64% | 89.4% | 89.2% | 100% | -9.23% | -14.09% | -14.14% | 674.3% | 352.0% |
| Class F | 3.64% | 3.81% | 3.94% | 91.4% | 97.3% | 100% | -24.14% | -33.61% | -33.30% | 491.6% | 437.5% |
| Class TGM | 0.91% | 0.94% | 0.93% | 90.9% | 100.2% | 100% | -38.97% | -46.94% | -46.18% | 408.8% | 553.0% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | | | | | | |
|  | **Over ECM-11** | | | | | | **Over VTM-11ecm11** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 | 2.64% | 1.21% | 1.83% | 91.4% | 96.6% | 100% | -20.42% | -26.82% | -35.73% | 800.1% | 870.8% |
| Class A2 | 1.79% | 1.52% | 2.37% | 89.5% | 97.8% | 100% | -24.68% | -35.29% | -38.66% | 725.4% | 1098.7% |
| Class B | 1.72% | 1.44% | 1.34% | 90.0% | 97.6% | 100% | -19.42% | -34.54% | -32.18% | 651.1% | 870.8% |
| Class C | 1.26% | 1.17% | 1.42% | 88.2% | 95.3% | 100% | -21.24% | -26.28% | -25.99% | 658.4% | 784.1% |
| Class E |  |  |  |  |  |  |  |  |  |  |  |
| **Overall (Ref)** | 1.80% | 1.34% | 1.67% | 89.7% | 96.8% | 100% | -21.15% | -30.94% | -32.54% | 695.4% | 887.1% |
| Class D | 0.60% | 1.11% | 0.53% | 87.5% | 94.8% | 100% | -22.63% | -27.88% | -28.33% | 687.7% | 938.7% |
| Class F | 3.99% | 3.87% | 3.81% | 88.6% | 94.8% | 100% | -25.95% | -35.17% | -35.51% | 546.6% | 478.0% |
| Class TGM | 2.08% | 2.49% | 2.68% | 90.3% | 98.5% | 100% | -35.46% | -41.95% | -41.93% | 553.6% | 508.7% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | | | | | | | |
|  | **Over ECM-11** | | | | | | **Over VTM-11ecm11** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |  |  |  |  |
| Class B | 1.88% | 4.45% | 4.67% | 83.0% | 96.1% | 100% | -16.20% | -38.71% | -35.75% | 555.8% | 631.1% |
| Class C | 1.72% | 3.32% | 2.81% | 82.6% | 91.7% | 100% | -18.12% | -29.50% | -29.71% | 514.9% | 652.5% |
| Class E | 1.33% | 3.37% | 2.95% | 86.6% | 96.7% | 101% | -15.16% | -27.84% | -28.28% | 516.0% | 422.0% |
| **Overall (Ref)** | 1.69% | 3.81% | 3.62% | 83.7% | 94.8% | 100% | -16.58% | -32.92% | -31.87% | 531.8% | 577.1% |
| Class D | 0.71% | 3.64% | 2.70% | 82.0% | 92.0% | 100% | -20.63% | -31.78% | -31.15% | 545.2% | 741.7% |
| Class F | 3.18% | 4.31% | 4.36% | 86.8% | 94.0% | 99% | -22.51% | -37.12% | -36.40% | 515.6% | 418.2% |
| Class TGM | 2.15% | 4.07% | 4.71% | 89.9% | 94.3% | 100% | -34.06% | -42.39% | -42.44% | 495.2% | 391.1% |

* 1. ***Group 3 off***

Group 3 includes intra and IBC template matching (with search) related tools. The attached offgroup3.cfg was used in addition to ECM CTC settings.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | | | | |
|  | **Over ECM-11** | | | | | | **Over VTM-11ecm11** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 | 1.03% | 2.64% | 4.45% | 79.4% | 83.7% | 99% | -9.89% | -19.88% | -26.34% | 595.0% | 238.8% |
| Class A2 | 1.89% | 4.58% | 4.57% | 79.6% | 74.0% | 98% | -15.35% | -25.74% | -28.89% | 595.4% | 223.0% |
| Class B | 1.95% | 4.11% | 4.01% | 77.9% | 76.6% | 97% | -9.60% | -24.05% | -22.19% | 629.2% | 314.1% |
| Class C | 1.75% | 2.02% | 2.31% | 78.4% | 76.2% | 100% | -9.88% | -15.12% | -15.40% | 646.3% | 298.0% |
| Class E | 3.10% | 3.99% | 3.80% | 78.3% | 77.3% | 99% | -12.38% | -21.86% | -20.23% | 605.0% | 329.8% |
| **Overall** | 1.93% | 3.46% | 3.77% | 78.6% | 77.3% | 98% | -11.13% | -21.29% | -22.16% | 617.3% | 282.4% |
| Class D | 1.30% | 1.76% | 2.09% | 80.0% | 74.5% | 100% | -8.31% | -13.08% | -12.90% | 650.7% | 294.3% |
| Class F | 4.48% | 5.64% | 5.64% | 79.5% | 71.1% | 100% | -23.42% | -32.29% | -32.13% | 447.5% | 300.6% |
| Class TGM | 8.69% | 10.55% | 10.82% | 82.8% | 60.4% | 100% | -34.38% | -42.00% | -41.04% | 397.1% | 289.6% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | | | | | | |
|  | **Over ECM-11** | | | | | | **Over VTM-11ecm11** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 | 0.95% | 2.06% | 5.33% | 94.4% | 97.8% | 100% | -21.73% | -26.40% | -34.03% | 667.3% | 560.6% |
| Class A2 | 1.15% | 3.26% | 3.11% | 94.8% | 98.0% | 100% | -25.17% | -34.31% | -38.30% | 623.9% | 701.4% |
| Class B | 1.05% | 3.97% | 3.49% | 95.2% | 98.5% | 100% | -19.97% | -32.97% | -30.86% | 688.0% | 780.8% |
| Class C | 0.84% | 1.32% | 1.44% | 94.5% | 98.9% | 100% | -21.57% | -26.14% | -25.97% | 722.1% | 834.6% |
| Class E |  |  |  |  |  |  |  |  |  |  |  |
| **Overall (Ref)** | 0.99% | 2.74% | 3.24% | 94.8% | 98.4% | 100% | -21.79% | -30.10% | -31.68% | 679.3% | 728.1% |
| Class D | 0.64% | 1.54% | 1.63% | 95.3% | 98.0% | 100% | -22.61% | -27.58% | -27.51% | 742.2% | 907.9% |
| Class F | 2.83% | 3.51% | 3.84% | 92.2% | 97.3% | 100% | -26.81% | -35.46% | -35.54% | 591.8% | 469.2% |
| Class TGM | 5.63% | 7.27% | 7.75% | 94.1% | 95.9% | 99% | -33.26% | -39.24% | -39.10% | 575.8% | 427.8% |

* 1. ***Group 4 off***

Group 4 includes tools that require more processing on the neighboring reconstructed samples than VVC.

The attached offgroup4.cfg was used in addition to ECM CTC settings.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | | | | |
|  | **Over ECM-11** | | | | | | **Over VTM-11ecm11** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 | 1.58% | 5.59% | 7.61% | 87.2% | 91.6% | 99% | -9.41% | -17.76% | -24.21% | 653.6% | 261.4% |
| Class A2 | 2.50% | 6.27% | 7.35% | 83.6% | 93.4% | 99% | -14.95% | -24.55% | -27.10% | 625.5% | 281.7% |
| Class B | 1.03% | 6.65% | 5.38% | 87.0% | 95.7% | 103% | -10.41% | -22.46% | -21.31% | 702.9% | 392.5% |
| Class C | 0.89% | 2.85% | 2.85% | 84.9% | 95.4% | 99% | -10.63% | -14.51% | -15.01% | 699.5% | 372.9% |
| Class E | 1.21% | 6.71% | 4.10% | 87.5% | 96.2% | 100% | -13.99% | -20.12% | -20.17% | 675.9% | 410.2% |
| **Overall** | 1.37% | 5.58% | 5.31% | 86.1% | 94.6% | 101% | -11.65% | -19.86% | -21.17% | 675.9% | 345.7% |
| Class D | 0.87% | 2.51% | 2.58% | 85.5% | 94.1% | 100% | -8.70% | -12.47% | -12.52% | 695.6% | 371.8% |
| Class F | 1.31% | 4.37% | 4.41% | 92.2% | 94.7% | 100% | -25.65% | -33.14% | -32.92% | 518.7% | 400.6% |
| Class TGM | 2.47% | 4.15% | 4.01% | 91.0% | 94.8% | 97% | -38.10% | -45.37% | -44.68% | 436.7% | 454.3% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | | | | | | |
|  | **Over ECM-11** | | | | | | **Over VTM-11ecm11** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 | 1.24% | 3.03% | 4.87% | 91.0% | 98.7% | 100% | -21.51% | -25.76% | -34.19% | 643.7% | 566.1% |
| Class A2 | 0.88% | 3.47% | 3.94% | 93.0% | 99.3% | 100% | -25.36% | -34.19% | -37.88% | 612.3% | 711.1% |
| Class B | 0.52% | 5.84% | 5.05% | 93.2% | 100.3% | 100% | -20.39% | -31.96% | -30.02% | 673.9% | 795.7% |
| Class C | 0.40% | 2.02% | 1.91% | 94.2% | 100.3% | 100% | -21.91% | -25.66% | -25.67% | 719.7% | 846.1% |
| Class E |  |  |  |  |  |  |  |  |  |  |  |
| **Overall (Ref)** | 0.70% | 3.79% | 3.95% | 93.0% | 99.8% | 100% | -22.01% | -29.49% | -31.27% | 666.7% | 738.8% |
| Class D | 0.44% | 2.32% | 1.87% | 95.2% | 99.6% | 100% | -22.76% | -27.06% | -27.47% | 741.2% | 922.3% |
| Class F | 0.96% | 3.32% | 3.54% | 96.0% | 100.1% | 100% | -28.08% | -35.57% | -35.78% | 616.1% | 482.7% |
| Class TGM | 0.75% | 1.57% | 1.41% | 95.0% | 99.9% | 99% | -36.38% | -42.51% | -42.70% | 581.9% | 445.6% |

* 1. ***Group 1-4 off***

In this test, all the tools in the group 1-4 are switched off. The attached offgroup1-4.cfg was used in addition to ECM CTC settings.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | | | | |
|  | **Over ECM-11** | | | | | | **Over VTM-11ecm11** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 | 3.25% | 9.51% | 13.38% | 61.6% | 73.2% | 100% | -7.90% | -14.82% | -20.27% | 461.2% | 208.8% |
| Class A2 | 5.28% | 12.23% | 12.81% | 58.4% | 66.1% | 99% | -12.59% | -20.50% | -23.48% | 437.0% | 199.4% |
| Class B | 3.78% | 11.70% | 10.44% | 58.2% | 66.8% | 96% | -7.98% | -18.93% | -17.64% | 470.2% | 274.1% |
| Class C | 3.35% | 5.88% | 6.15% | 56.1% | 62.4% | 100% | -8.48% | -12.04% | -12.32% | 461.9% | 244.0% |
| Class E | 5.76% | 11.92% | 9.27% | 59.0% | 67.3% | 101% | -10.12% | -16.20% | -16.15% | 455.4% | 287.1% |
| **Overall** | 4.18% | 10.17% | 10.18% | 58.4% | 66.8% | 99% | -9.20% | -16.52% | -17.62% | 458.7% | 244.0% |
| Class D | 2.85% | 5.02% | 5.24% | 58.0% | 58.9% | 99% | -6.90% | -10.32% | -10.23% | 471.8% | 232.8% |
| Class F | 10.77% | 15.69% | 15.92% | 64.6% | 62.2% | 100% | -19.07% | -26.56% | -26.18% | 363.5% | 263.1% |
| Class TGM | 14.79% | 18.62% | 19.29% | 66.1% | 54.8% | 98% | -30.81% | -37.90% | -36.70% | 317.1% | 262.6% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | | | | | | |
|  | **Over ECM-11** | | | | | | **Over VTM-11ecm11** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 | 9.34% | 11.00% | 17.80% | 66.8% | 77.5% | 99% | -15.21% | -20.07% | -26.47% | 472.4% | 444.1% |
| Class A2 | 9.14% | 14.17% | 15.38% | 66.5% | 73.8% | 99% | -19.24% | -27.49% | -31.05% | 437.4% | 528.0% |
| Class B | 7.44% | 15.67% | 14.57% | 64.7% | 72.8% | 99% | -14.86% | -25.58% | -23.56% | 467.4% | 577.6% |
| Class C | 6.93% | 9.39% | 9.95% | 61.6% | 64.8% | 99% | -16.78% | -20.20% | -19.64% | 470.4% | 546.7% |
| Class E |  |  |  |  |  |  |  |  |  |  |  |
| **Overall (Ref)** | 8.03% | 12.76% | 14.15% | 64.6% | 71.7% | 99% | -16.32% | -23.42% | -24.59% | 463.0% | 530.5% |
| Class D | 5.31% | 8.58% | 8.97% | 62.8% | 62.8% | 99% | -18.94% | -22.47% | -22.18% | 489.3% | 582.1% |
| Class F | 11.83% | 15.16% | 15.85% | 69.4% | 75.5% | 100% | -20.53% | -28.50% | -28.42% | 445.1% | 363.7% |
| Class TGM | 13.71% | 16.90% | 17.18% | 71.3% | 76.2% | 99% | -28.14% | -33.82% | -33.81% | 436.8% | 339.7% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | | | | | | | |
|  | **Over ECM-11** | | | | | | **Over VTM-11ecm11** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |  |  |  |  |
| Class B | 7.15% | 12.96% | 13.09% | 55.3% | 61.3% | 100% | -11.84% | -34.09% | -31.04% | 365.2% | 464.3% |
| Class C | 7.01% | 9.85% | 10.53% | 53.5% | 54.2% | 100% | -13.83% | -25.03% | -24.35% | 354.4% | 387.6% |
| Class E | 6.78% | 14.35% | 13.74% | 65.1% | 75.2% | 99% | -10.60% | -20.37% | -20.92% | 406.2% | 355.4% |
| **Overall (Ref)** | 7.01% | 12.27% | 12.40% | 57.0% | 61.9% | 99% | -12.20% | -27.64% | -26.28% | 371.3% | 408.9% |
| Class D | 5.19% | 10.00% | 10.65% | 55.6% | 52.2% | 99% | -17.06% | -27.72% | -25.80% | 391.4% | 422.6% |
| Class F | 10.57% | 14.24% | 15.19% | 66.2% | 68.0% | 99% | -17.00% | -31.21% | -29.99% | 382.9% | 308.2% |
| Class TGM | 13.52% | 17.29% | 18.26% | 71.5% | 64.4% | 100% | -26.65% | -34.93% | -34.90% | 394.7% | 286.0% |

* 1. ***Summary***

The tool-off results on top of the recent ECM versions are summarized below. Note that SCC results are not included.

1. ***Issues***
   1. ***Resolved issues***

* Software issues #56 with tool off tests was resolved
  1. ***Open issues***
* #58, [ECM-11.0 decoding mismatch when DualITree is off](https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/issues/58)
* #53, decoding mismatch was observed when AMVR is off
* The variation of runtime over VTM11ECM9 is relatively large, for which the version of compiler matters

1. ***Input* contributions**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| [JVET-AG0173](https://jvet-experts.org/doc_end_user/current_document.php?id=13729) | m66069 | 2024-01-10 20:21:16 | 2024-01-10 20:48:21 | 2024-01-10 20:48:21 | AHG7: ECM-11 results of non-CTC sequences | [X. Li (Google)](mailto:xlxiangli@google.com) |

1. ***Recommendations***

* Continue and improve tool assessment
* Resolve identified software issues related to the tool assessment
* Review all the input contributions

It was commented by one expert that he believes to have resolved the issue with the SW mismatch and will submit a merge request.

It was suggested to perform another round of subjective testing of ECM in the near future. If some of the group-off tests would be included, a careful selection would be necessary due to limited resources in on-site expert viewing. Discuss in joint meeting with AG 5

[JVET-AG0008](https://jvet-experts.org/doc_end_user/current_document.php?id=13782) JVET AHG report: Optimization of encoders and receiving systems for machine analysis of coded video content (AHG8) [C. Hollmann, S. Liu, S. Wang, M. Zhou (AHG chairs)]

1. ***Activities***

The AHG used the main JVET reflector, jvet@lists.rwth-aachen.de, for email discussion. The AHG kick-off email was sent on November 10, 2023, followed by announcements of CTC updates and document uploads. There were offline discussions and email exchanges about project timelines among co-chairs and editors, as well as editings to improve the quality of the technical report. The outcome was summarized in JVET-AG0090. There are total 10 input contriubtions related to AHG 8 mandates submitted to this meeting. They are listed in Section 3.

* 1. ***Common Test Conditions***

Common test conditions (CTC) for optimization of encoders and receiving systems for machine analysis of coded video content, were updated and summarized in output document JVET-AF2031. This document includes detailed descriptions of test datasets, anchor software and configurations, anchor generation processes, machine task networks used, test and training conditions, evaluation methodologies and metrics.

Following the discussions in the last meeting, the updated CTC includes two anchors, generated by the most recent versions of VTM software, VTM21 and VTM22, respectively. Hence two reporting templates in Excel format are enclosed in the document package for reporting VTM21 and VTM22 based results accordingly. This output document package JVET-AF2031 was uploaded on 2023-11-10 to JVET document management system and is also available at <https://vcgit.hhi.fraunhofer.de/jvet-ahg-ofm>/ofm-ctc.

* 1. ***Technical Report***

Based on the latest draft (draft 3) of the technical report (TR) JVET-AE2030, further improvements with comments have been added, and submitted as an attachment of input contribution JVET-AG0090. Besides editorial refinements, the scope in Section 1 was clarified, indicating that the described algorithms were tested on a constrained set of machine vision tasks and thus might not be applicable to all machine vision tasks.

The timeline of this TR was discussed among co-chairs and editors. The following suggested timeline was submitted to this JVET meeting for review with JVET-AG0090:

* Request for subdivision: January 2024
* Committee draft (CD): April 2024
* CD ballot: May – June 2024
* Draft TR: November 2024
* DTR ballot: December 2024 – March 2025
* Finalization: April 2025

Regarding the request for subdivision, it is proposed to request a new part in ISO/IEC 23888 (Artificial Intelligence for Multimedia) with a timeline of 24 months.

Note that there is an input contribution JVET-AG0213 suggesting to align the timelines of this TR and VSEI v4. This is planned to be discussed during this meeting.

* 1. ***Git Management***

AHG 8 related software and documents can be found at <https://vcgit.hhi.fraunhofer.de/jvet-ahg-ofm>. This repository contains two projects, one (<https://vcgit.hhi.fraunhofer.de/jvet-ahg-ofm/ofm-ctc>) containing instrucitons and information for conducting experiements and evaluation, such as evaluation scripts, machine task networks, CTC and reporting template with anchor results, while the other (<https://vcgit.hhi.fraunhofer.de/jvet-ahg-ofm/vtm-ofm>) containing implementation examples. Three software implementation examples are hosted in separate branches:

* JVET-AB0275: a region of interest-based method that uses adaptive QP to reduce the quality in background areas
* JVET-AC0086: a method that uses a pre-analysis to perform content adaptive machine vision oriented preprocessing
* JVET-AE0143: a spatial resampling algorithm and an exemplar software implementation

1. **Input *contributions***

There are 10 input contriubtions related to AHG 8 mandates. They are listed below.

|  |  |  |
| --- | --- | --- |
| **Report** | | |
| JVET-AG0008 | JVET AHG report: Optimization of encoders and receiving systems for machine analysis of coded video content (AHG8) | C. Hollmann, S. Liu, S. Wang, M. Zhou (AHG chairs) |
| **Proposal** | | |
| [JVET-AG0085](file:///C:\Users\ehollch\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\U0B9JGQQ\current_document.php%3fid=13641) | [AHG8] Continuation of study on different VTM versions | [C. Hollmann (Ericsson)](mailto:christopher.hollmann@ericsson.com) |
| [JVET-AG0090](file:///C:\Users\ehollch\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\U0B9JGQQ\current_document.php%3fid=13646) | [AHG8] Comments and editorial changes to the draft TR on optimizations for encoders and receiving systems for machine analysis of coded video content | [C. Hollmann (Ericsson)](mailto:christopher.hollmann@ericsson.com), [S. Liu (Tencent)](mailto:shanl@global.tencent.com), [J. Chen (Alibaba)](mailto:jiechen.cj@alibaba-inc.com) |
| [JVET-AG0178](file:///C:\Users\ehollch\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\U0B9JGQQ\current_document.php%3fid=13734) | AHG8: Truncating bit depth in video coding for machine tasks | [D. Ding](mailto:ddding@global.tencent.com), [X. Zhao](mailto:xinzzhao@global.tencent.com), [S. Liu (Tencent)](mailto:shanl@global.tencent.com) |
| [JVET-AG0209](file:///C:\Users\ehollch\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\U0B9JGQQ\current_document.php%3fid=13765) | AHG8: A suggestion for the performance evaluation of VCM | [S. Wang](mailto:shurun.wsr@alibaba-inc.com), [J. Chen](mailto:jiechen.cj@alibaba-inc.com), [Y. Ye (Alibaba)](mailto:yan.ye@alibaba-inc.com), [S. Wang (CityU)](mailto:shiqwang@cityu.edu.hk) |
| [JVET-AG0212](file:///C:\Users\ehollch\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\U0B9JGQQ\current_document.php%3fid=13768) | AHG8: A post-processing algorithm for machine consumption | [B. Li](mailto:libinzhe.lbz@alibaba-inc.com), [S. Wang](mailto:shurun.wsr@alibaba-inc.com), [J. Chen](mailto:jiechen.cj@alibaba-inc.com), [Y. Ye(Alibaba)](mailto:yan.ye@alibaba-inc.com), [S. Wang(Cityu)](mailto:shiqwang@cityu.edu.hk) |
| [JVET-AG0213](file:///C:\Users\ehollch\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\U0B9JGQQ\current_document.php%3fid=13769) | AHG1/AHG2/AHG8: On project management related to the encoder optimization information SEI message | [M. M. Hannuksela](mailto:miska.hannuksela@nokia.com), A. Aminlou, F. Cricri, H. Zhang (Nokia) |
| [JVET-AG0216](file:///C:\Users\ehollch\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\U0B9JGQQ\current_document.php%3fid=13772) | AHG8: Multi-layer VVC for hybrid machine-human consumption | J. Laitinen, T. Partanen, A. Mercat, J. Vanne (Tampere University), [A. Aminlou](mailto:alireza.aminlou@nokia.com), M. M. Hannuksela, F. Cricri, H. Zhang (Nokia), |
| [JVET-AG0217](file:///C:\Users\ehollch\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\U0B9JGQQ\current_document.php%3fid=13773) | AHG8: Reduced residual encoding in VVC for machine consumption | [A. Alireza (Nokia)](mailto:alireza.aminlou@nokia.com), A. Hallapuro, H. Zhang, |
| **Crosscheck** | | |
| JVET-AG0242 | Cross-check of JVET-AG0178 (AHG8: Truncating bit depth in video coding for machine tasks) | C. Hollmann (Ericsson) |

1. ***Recommendations***

The AHG recommends to:

* Review all input contributions.
* Continue investigating non-normative technologies and their suitability for machine analysis applications.
* Continue improving draft technical report on optimization of encoders and receiving systems for machine analysis of coded video content.
* Continue refining test conditions, evalution and reporting procedures.
* Discuss TR development and finalization timeline, possibly together with other related groups.

It was pointed out that a TR could have a shorter development cycle than assumed in the report above. Further, a request for subdivision should not be done too early to avoid possible cancellation. To be further discussed in context of JVET-AG0213.

[JVET-AG0009](https://jvet-experts.org/doc_end_user/current_document.php?id=13783) JVET AHG report: SEI message studies (AHG9) [S. McCarthy, Y.-K. Wang (co-chairs), T. Chujoh, S. Deshpande, C. Fogg, M. M. Hannuksela, Hendry, P. de Lagrange, G. J. Sullivan, A. Tourapis, S. Wenger (vice-chairs)]

1. **Related contributions**

A total of 43 contributions are identified relating to the mandates of AHG9. Some contributions also relate to the work of AHG1, AHG2, AHG8 and AHG16.

The number of contributions relating to each AHG9 mandate is as follows (some contributions relate to more than one mandate):

* 4 contributions relate to the mandate to study the SEI messages in VSEI, VVC, HEVC, and AVC.
  + 3 contributions relate to neural network post filter characteristics SEI messages
  + 1 contribution relates to phase indication SEI message
* 24 contributions relate to the mandate to discuss the document for the TuC for future extensions of VSEI ([JVET-2032](https://jvet-experts.org/doc_end_user/current_document.php?id=13592)).
  + 9 contributions relate to aspects of SEI processing order and processing order nesting SEI messages (These relate to [JVET-AF2027](https://jvet-experts.org/doc_end_user/current_document.php?id=13590))
  + 4 contributions relate to the source picture timing information SEI message
  + 3 contributions relate to encoder optimization information SEI message
  + 2 contributions relate to picture modality information
  + 2 contributions relate to image format metadata SEI messages
  + 1 contribution relates to the object mask information SEI message
  + 1 contribution relates to the multiplane image information SEI message
  + 3 contributions relate to progression towards a working draft of VSEI version 4
* 1 contribution relates to the mandate to collect software and showcase information for SEI messages;
* 19 contributions relate to the mandate to identify potential needs for additional SEI messages, including the study of SEI messages defined in HEVC and AVC for potential use in the VVC context.
  + 4 contribution relates to the film grain adaptive SEI message
  + 3 contributions relate to SEI messages related to generative face video
  + 2 contributions relate to image format metadata SEI messages
  + 1 contribution relates to the phase indication SEI message
  + 9 contributions relate to other SEI messages
* 0 contributions relate to the mandate to study the alignments of the same SEI messages in different standards.
* 1 contribution relate to AHG9's mandate as it allows signalling of required decoder processing of SEI messages, among other NAL units. However, it is not an SEI message proposal.

The following is a list of contributions related to the mandates of AHG9.

* 1. ***Study*** ***the SEI messages in VSEI, VVC, HEVC and AVC***
     1. ***NNPF* SEI messages**

[JVET-AG0088](https://jvet-experts.org/doc_end_user/current_document.php?id=13644) AHG9: Usage of the neural-network post-filter characteristics SEI message to define the generator NN of the generative face video SEI message [M. M. Hannuksela, F. Cricri, H. Zhang (Nokia)]

JVET-AG0088 also relates to the mandate to identify potential needs for additional SEI messages

[JVET-AG0089](https://jvet-experts.org/doc_end_user/current_document.php?id=13645) AHG9: Temporal extrapolation purpose for the neural-network post-filter characteristics SEI message [M. M. Hannuksela, F. Cricri, H. Zhang (Nokia)]

[JVET-AG0192](https://jvet-experts.org/doc_end_user/current_document.php?id=13748) AHG9: Miscellaneous NNPF items related to VUI [J. Xu, Y.-K. Wang (Bytedance)]

* + 1. **Phase *indication* SEI message**

[JVET-AG0107](https://jvet-experts.org/doc_end_user/current_document.php?id=13663) AHG9: On phase indication SEI message [T. Chujoh, T. Ikai (Sharp), K. Kawamura (KDDI)]

JVET-AG0107 also relates to the mandate to identify potential needs for additional SEI messages.

* 1. ***Discuss the document for the TuC for future extensions of VSEI***
     1. **SEI *processing* order and processing order nesting SEI messages**

[JVET-AG0052](https://jvet-experts.org/doc_end_user/current_document.php?id=13608) AHG9: On using SEI processing order for NNPF grouping [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

[JVET-AG0053](https://jvet-experts.org/doc_end_user/current_document.php?id=13609) AHG9: On activation and cancelling persistence of SEI message included in a processing order nesting SEI message [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

[JVET-AG0054](https://jvet-experts.org/doc_end_user/current_document.php?id=13610) AHG9: On the case an SEI message included in multiple sei processing order SEI messages [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

[JVET-AG0105](https://jvet-experts.org/doc_end_user/current_document.php?id=13661) AHG9: On the SEI processing order SEI message [Y. Gao, Y.-X. Bai, S.-W. Xie, M.-H. Jia, C. Huang, P. Wu (ZTE)]

[JVET-AG0165](https://jvet-experts.org/doc_end_user/current_document.php?id=13721) AHG9: On the processing order nesting SEI message [L. Chen, O. Chubach, Y. Huang, S. Lei (MediaTek)]

[JVET-AG0166](https://jvet-experts.org/doc_end_user/current_document.php?id=13722) AHG9: On grouping the SEI prefix data as a new SEI message [L. Chen, O. Chubach, Y. Huang, S. Lei (MediaTek)]

[JVET-AG0168](https://jvet-experts.org/doc_end_user/current_document.php?id=13724) AHG9: Proposed update for the SEI processing order SEI message [L. Chen, O. Chubach, Y. Huang, S. Lei (MediaTek)]

[JVET-AG0169](https://jvet-experts.org/doc_end_user/current_document.php?id=13725) AHG9: Comments on use of NNPF SEI messages in the SEI processing order SEI message [L. Chen, O. Chubach, Y. Huang, S. Lei (MediaTek)]

[JVET-AG0180](https://jvet-experts.org/doc_end_user/current_document.php?id=13736) AHG9: On the SEI processing order SEI message [Y. Sanchez, R. Skupin, C. Hellge, T. Schierl (HHI)]

* + 1. **Source *picture* timing information SEI message**

[JVET-AG0070](https://jvet-experts.org/doc_end_user/current_document.php?id=13626) AHG9: Comments on Source Picture Timing Information Message [S. Deshpande, J. Samuelsson-Allendes (Sharp)]

[JVET-AG0082](https://jvet-experts.org/doc_end_user/current_document.php?id=13638) AHG9: On Source Picture Timing SEI message [j. Samuelsson-Allendes, S. Deshpande (Sharp)]

[JVET-AG0188](https://jvet-experts.org/doc_end_user/current_document.php?id=13744) AHG9: On source picture timing information SEI message specification text [J.R. Arumugam, L. Jawale (Ittiam), P. Yin, G.J. Sullivan, S. McCarthy (Dolby)]

[JVET-AG0191](https://jvet-experts.org/doc_end_user/current_document.php?id=13747) AHG9: Reference software for source picture timing information SEI message [J.R. Arumugam, L. Jawale (Ittiam), P. Yin, S. McCarthy (Dolby)]

JVET-AG0191 also relates to the mandate to collect software and showcase information.

* + 1. **Encoder *optimization* information SEI message**

[JVET-AG0081](https://jvet-experts.org/doc_end_user/current_document.php?id=13637) AHG9: On signalling privacy protection information in SEI message [C. Kim, Hendry, J. Lim, S. Kim (LGE)]

JVET-AG0081 also relates to the mandate to identify potential needs for additional SEI messages

[JVET-AG0083](https://jvet-experts.org/doc_end_user/current_document.php?id=13639) AHG9: On feature-based optimization type [C. Kim, Hendry, J. Lim, S. Kim (LGE)]

[JVET-AG0086](https://jvet-experts.org/doc_end_user/current_document.php?id=13642) AHG9: On the encoder optimization information SEI message [M. M. Hannuksela, F. Cricri, H. Zhang (Nokia)]

* + 1. **Picture modality information**

[JVET-AG0077](https://jvet-experts.org/doc_end_user/current_document.php?id=13633) AHG9: On Picture Modality Information [J. Gao, H.-B. Teo, C.-S. Lim, K. Abe, V. Drugeon (Panasonic)]

[JVET-AG0079](https://jvet-experts.org/doc_end_user/current_document.php?id=13635) AHG2/AHG9: VUI extension mechanism and picture modality information for AVC and HEVC [J. Gao, H.-B. Teo, C.-S. Lim, K. Abe, V. Drugeon (Panasonic)]

* + 1. **Image format metadata SEI messages**

[JVET-AG0182](https://jvet-experts.org/doc_end_user/current_document.php?id=13738) AHG9: JPEG segments SEI message [P. de Lagrange, D. Doyen, E. François, F. Urban, C. Salmon-Legagneur (InterDigital)]

JVET-AF0182 also relates to the mandate to identify potential needs for additional SEI messages.

[JVET-AG0183](https://jvet-experts.org/doc_end_user/current_document.php?id=13739) AHG9: TIFF data SEI message [P. de Lagrange, D. Doyen, E. François, F. Urban, C. Salmon-Legagneur (InterDigital)]

JVET-AG0183 also relates to the mandate to identify potential needs for additional SEI messages.

* + 1. **Object mask *information* SEI message**

[JVET-AG0148](https://jvet-experts.org/doc_end_user/current_document.php?id=13704) AHG9: On object mask information SEI message [J. Chen, Y. Ye, S. Wang (Alibaba)]

JVET-AG0148 also relates to 2.2.8 Progression towards a working draft of VSEI version 4.

* + 1. **Multiplane *image* information SEI message**

[JVET-AG0232](https://jvet-experts.org/doc_end_user/current_document.php?id=13805) AHG9: Support of non-parallel MPI layers in the MPII SEI message [Y. Li (SJTU), Y.-K. Wang (Bytedance), Y. Xu, K. Yang (SJTU)]

* + 1. ***Progression* towards a working draft of VSEI version 4**

[JVET-AG0148](https://jvet-experts.org/doc_end_user/current_document.php?id=13704) AHG9: On object mask information SEI message [J. Chen, Y. Ye, S. Wang (Alibaba)]

JVET-AG0148 also relates to 2.2.6 Object mask information.

[JVET-AG0204](https://jvet-experts.org/doc_end_user/current_document.php?id=13760) AHG9: On VSEI version 4 [Hendry (LGE), S. McCarthy (Dolby), J. Chen, Y. Ye (Alibaba)]

[JVET-AG0213](https://jvet-experts.org/doc_end_user/current_document.php?id=13769) AHG1/AHG2/AHG8: On project management related to the encoder optimization information SEI message [M. M. Hannuksela, A. Aminlou, F. Cricri, H. Zhang (Nokia)]

* 1. ***Collect software and showcase information for SEI messages***

[JVET-AG0191](https://jvet-experts.org/doc_end_user/current_document.php?id=13747) AHG9: Reference software for source picture timing information SEI message [J.R. Arumugam, L. Jawale (Ittiam), P. Yin, S. McCarthy (Dolby)]

JVET-AG0191 also relates to the mandate to discuss the document for the TuC for future extensions of VSEI.

* 1. ***Identify potential needs for additional SEI messages, including study of AVC and HEVC SEI messages for use in VVC***
     1. **Film *grain* related SEI message**

[JVET-AG0101](https://jvet-experts.org/doc_end_user/current_document.php?id=13657) AHG9: Film grain adaptive SEI message [Y. Gao, S.-W. Xie, Y.-X. Bai, M.-H. Jia, C. Huang, P. Wu (ZTE)]

[JVET-AG0140](https://jvet-experts.org/doc_end_user/current_document.php?id=13696) AHG9/AHG13: FGS Extension SEI message for spatial adaptation [G.Teniou, S. Wenger, A. Hinds]

[JVET-AG0160](https://jvet-experts.org/doc_end_user/current_document.php?id=13716) AHG9/AHG13: FGS Extension SEI message useful descriptors [G. Teniou, S. Wenger, A. Hinds]

[JVET-AG0215](https://jvet-experts.org/doc_end_user/current_document.php?id=13771) AHG9/AHG13: Region-dependent film grain characteristics [P. de Lagrange, E. François, M. Le Pendu, C. Salmon-Legagneur (InterDigital)]

* + 1. ***Generative* face video SEI message**

[JVET-AG0087](https://jvet-experts.org/doc_end_user/current_document.php?id=13643) AHG9: On the generative face video SEI message [M. M. Hannuksela, F. Cricri, H. Zhang (Nokia)]

[JVET-AG0088](https://jvet-experts.org/doc_end_user/current_document.php?id=13644) AHG9: Usage of the neural-network post-filter characteristics SEI message to define the generator NN of the generative face video SEI message [M. M. Hannuksela, F. Cricri, H. Zhang (Nokia)]

JVET-AG0088 also relates to the mandate to discuss the document for the TuC for future extensions of VSEI.

[JVET-AG0203](https://jvet-experts.org/doc_end_user/current_document.php?id=13759) AHG9/AHG16: Common text for proposed generative face video SEI message [J. Chen, B. Chen, Y. Ye (Alibaba), S. Yin, S. Wang (CityU), S. McCarthy, P. Yin, G.-M. Su, A. K. Choudhury, W. Husak, G. J. Sullivan (Dolby)]

* + 1. **Image *format* metadata SEI messages**

[JVET-AG0182](https://jvet-experts.org/doc_end_user/current_document.php?id=13738) AHG9: JPEG segments SEI message [P. de Lagrange, D. Doyen, E. François, F. Urban, C. Salmon-Legagneur (InterDigital)]

JVET-AG0182 also relates to the mandate to discuss the document for the TuC for future extensions of VSEI.

[JVET-AG0183](https://jvet-experts.org/doc_end_user/current_document.php?id=13739) AHG9: TIFF data SEI message [P. de Lagrange, D. Doyen, E. François, F. Urban, C. Salmon-Legagneur (InterDigital)]

JVET-AG0183 also relates to the mandate to discuss the document for the TuC for future extensions of VSEI.

* + 1. **Phase *indication* SEI message**

[JVET-AG0107](https://jvet-experts.org/doc_end_user/current_document.php?id=13663) AHG9: On phase indication SEI message [T. Chujoh, T. Ikai (Sharp), K. Kawamura (KDDI)]

JVET-AG0107 also relates to the mandate to study the SEI messages in VSEI, VVC, HEVC and AVC.

* + 1. **Other SEI messages and aspects**

[JVET-AG0044](https://jvet-experts.org/doc_end_user/current_document.php?id=13598) AHG9: Copyright SEI message [S. Wenger, A. Hinds, G. Teniou]

[JVET-AG0045](https://jvet-experts.org/doc_end_user/current_document.php?id=13599) AHG9: AI marking SEI [S. Wenger, A. Hinds, G. Teniou]

[JVET-AG0046](https://jvet-experts.org/doc_end_user/current_document.php?id=13600) Application-required NAL Units [G. Teniou. S. Wenger]

[JVET-AG0049](https://jvet-experts.org/doc_end_user/current_document.php?id=13603) AHG9: SEI message extension of VVC for computer-generated hologram use [H. Kojima, K. Nonaka, R. Koiso, K. Kawamura, H. Kato (KDDI)]

[JVET-AG0051](https://jvet-experts.org/doc_end_user/current_document.php?id=13607) AHG9: On design for new SEI RBSP and SEI message [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

[JVET-AG0081](https://jvet-experts.org/doc_end_user/current_document.php?id=13637) AHG9: On signalling privacy protection information in SEI message [C. Kim, Hendry, J. Lim, S. Kim (LGE)]

JVET-AG0081 also relates to the mandate to discuss the document for the TuC for future extensions of VSEI.

[JVET-AG0144](https://jvet-experts.org/doc_end_user/current_document.php?id=13700) AHG9: Carriage of depth and alpha maps as HEVC single-layer bitstreams [E. Thomas, E. Potetsianakis, E. Alexiou, R. Ghaznavi-Youvalari, M. Abdoli, M-L. Champel (Xiaomi)]

[JVET-AG0167](https://jvet-experts.org/doc_end_user/current_document.php?id=13723) AHG9: Text prompt for generative AI SEI [A. Hinds, G. Teniou, S. Wenger]

[JVET-AG0184](https://jvet-experts.org/doc_end_user/current_document.php?id=13740) AHG9: Text comment SEI message [P. de Lagrange, D. Doyen, E. François, F. Urban, C. Salmon-Legagneur (InterDigital)]

* 1. ***Study the alignments of the same SEI messages in different standards***

No contributions

* 1. ***Non-SEI proposals closely related to SEI messages***

[JVET-AG0046](https://jvet-experts.org/doc_end_user/current_document.php?id=13600) Application-required NAL Units [G. Teniou, S. Wenger]

1. ***Activities***

The regular JVET e-mail reflector was used for discussions ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de)) with [AHG9] in message headers. There were no emails sent to the JVET reflector during the AHG period with [AHG9] in the message header. However, there was one email to the JVET reflector relating to the proposed copyright and AI marking SEI messages.

1. ***Recommendations***

The AHG recommends to:

* Review all related contributions; and
* Continue SEI messages studies.
* Update AHG9 mandates to include SEI-related document drafts (e.g., [JVET-AF2027](https://jvet-experts.org/doc_end_user/current_document.php?id=13590)) in addition to the TuC for future extensions of VSEI (e.g., [JVET-AF2032](https://jvet-experts.org/doc_end_user/current_document.php?id=13592))

It was asked what the status of the software related to the MPI message was. It was reported that a merge request was issued, but a merge had not yet been made. However, some problem of using the MIV content for testing this SEI message still needs to be resolved. Proponents were asked to provide downsampled content. A link to an external renderer was also provided, but not included in the software itself.

Experts were asked to raise any issues they detect on experimenting with the MPI software during the meeting.

[JVET-AG0010](https://jvet-experts.org/doc_end_user/current_document.php?id=13784) JVET AHG report: Encoding algorithm optimization (AHG10) [P. de Lagrange, A. Duenas, R. Sjöberg, A. Tourapis (AHG chairs)]

1. ***Related* contributions**

Three contributions in total, not including cross-checks, are identified relating to AHG10, and summarized in the following sections.

* 1. **Adaptive *resolution***
     1. **JVET-AG0116 – AHG12: GOP-based RPR encoder control for ECM**

This contribution is a follow-up of JVET-AF0058 that was porting the adaptive resolution algorithm available in the VTM (using the “GOPBasedRPR” configuration parameter) to ECM. Half the gain compared to the VTM case was reported, however such gain still considered as significant (around 1.5%; more than 3% in class A). Benefit was mostly reported for high resolution content and high QPs. The proponents were asked to provide full CTC results: it turns out that the objective performance results are almost neutral for the CTC (gains are still observed in class A), but the proponent reports some visual benefit.

* 1. ***Local QP* optimization**
     1. **JVET-AG0055 - CTU-Level Lagrange Multiplier and QP Adaptation for VVC Low-Delay Configuration**

This contribution is a follow-up of JVET-AF0089, that proposes a method that tries to give more quality weight to local areas that are likely to be reused in other picture (through temporal prediction). It does so by estimating a “distortion propagation factor” with a fast 1st inter pass.

During the 32nd JVET meeting, several experts commented that the contribution was interesting, and further optimization could be applied to reduce loss, optimize the interaction with BIM (Block Importance Mapping, a single-pass MCTF-dependent QP adaptation in the same spirit, already available in the VTM and the HM), etc. It was suggested for the proponents to also examine the impact on subjective quality.

In this follow-up version, scene cut detection has been introduced to mitigate the effect on the computation of the “distortion propagation factor”. BD-rate difference of -4% is reported in low-delay configurations, with a moderate increase of encoding time (around 2 to 3%).

A comparison with BIM is given, notably stating that BIM is active only on 1 picture out of 8 (the rate at which MCTF is active), while the method in this contribution impacts every picture, and has finer control over the Lagrange multiplier. As a reminder, the gains of BIM are around 2.3% in the low-delay and 2% in the random-access test conditions.

* 1. ***Residual* spatial *weighting***
     1. **JVET-AG0217 – Reduced residual encoding in VVC for machine consumption**

This contribution reports about a residual spatial weighting method (attenuating the residuals by 75% in the center of a coding block -when larger than 4 in either dimension- but preserving DC), which is said to provide gains for machine related tasks (a few % BD-rate (mAP); incomplete results). BD-rate (PSNR) loss is around 0.45%.

1. **Recommendation**

The AHG recommends that the related input contributions are reviewed and to further continue the study of encoding algorithm optimizations in JVET.

[JVET-AG0011](https://jvet-experts.org/doc_end_user/current_document.php?id=13785) JVET AHG report: Neural network-based video coding (AHG11) [E. Alshina, F. Galpin, S. Liu, A. Segall (co-chairs), J. Li, R.-L. Liao, D. Rusanovskyy, M. Santamaria, T. Shao, M. Wien, P. Wu (vice chairs)]

1. **Activities**

The AHG used the main JVET reflector, jvet@lists.rwth-aachen.de, for email. Two emails were exchanged on the reflector related to the AHG mandates.

* 1. **Common *Test Conditions***
     1. **Document**

The AHG released revised common test conditions as decided at the 32rd meeting, including the following changes:

* Clarification on averaging test results over different classes.
* Other editorial improvements.

Document was uploaded on 04.11.2023.

* + 1. **Anchor Encoding**

Anchor for the NN-based video coding activity made available on the Git repository used for the AHG activity: <https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/nnvc-ctc/-/master/Anchor%20performance/NNVC-current_Reporting_Template-with-anchor-data.xlsm?ref_type=heads>

also distributed by AhG14.

* 1. ***EE* Coordination**

The AHG finalized, conducted, and discussed the EE on NN based video coding. A summary report for the EE is available at this meeting as:

|  |  |  |
| --- | --- | --- |
| JVET-AG0023 | EE1: Summary report of exploration experiment on neural network-based video coding | E. Alshina, F. Galpin, Y. Li, D. Rusanovskyy, M. Santamaria, J. Ström, R. Chang, Z. Xie (EE coordinators) |

* 1. **Teleconferences**

The AHG conducted a joint teleconference with AHG14 and EE1 during the interim period. The teleconference was held on November 16, 2023. In this teleconference following topics have been discussed

* EE1-0, HOP retraining and luma/chroma balance changes.
* Lower complexity NN intra training cross check results.
* integration of some bug fix in NNVC-7.1.
* clarification of algorithm description for Unified Filter usage in JVET-AF2019

Re-trained with better Luma-Chroma compression performance balance HOP filter (EE1-1.0) was added to NNVC-7.1 SW as HOP2 and became comparison point for all tests in EE1 targeting improvement of HOP filter.

Lower complexity variant of NN Intra cross-check results have been reported to be successful, tool has been added to NNVC-7.1 SW, disabled by default.

* 1. **Performance *Evaluation***

The performance of the NNVC-7.1 anchor compared to VTM anchor is reported in AhG14 and EE1 summary reports. By default, NN-Intra and Low Operation Point (LOP.2) filter are enabled in NNVC-7.1.

Compression performance of some tools and their combinations available in NNVC SW or from EE1 tests on two complexity metrics (kMAC/pxl and total member for parameters) is shown in the plots below.

|  |  |
| --- | --- |
|  |  |

By available NNVC tool combination for High Operation Point 13-14% gain (for all color components) over VVC (RA configuration) is possible to achieve with ×3-4 Encoder and around ×2000 Decoder run time increment (relatively to VTM). For Low Operation Point 7% (Luma) and 14% (Chroma) gain over VVC (RA configuration) is possible to achieve with ×1.3 Encoder and around ×90 decoding run time increment (relatively to VTM). It should be noticed that all run time data reported with SADL which is used for transparency of the process, and far not yet as optimized as other AI platforms.

Major trends in EE1 tests are

* Base HOP had performance improvement using retraining and configuration changes.
* complexity reduction (visible in both kMAC/pxl and run time) with even minor performance improvement is possible both for HOP and LOP filters,
* NN-Inter complexity (both kMAC/pxl and run time) significantly improved, training cross-check has been completed,
* Content adaptivity for LOP filter (overfitting) reported significant speed up,
* Two operation points for NN-based super resolution (20 kMAC/pxl and 469 kMAC/pxl) have been demonstrated,
* New training data base was identified and used for EE1-3.1 test.

1. **Input contributions**

There are 25 input contributions related to the AHG mandates. The list of input contributions is provided below.

* 1. ***EE and Related Input Contributions***

|  |  |  |
| --- | --- | --- |
| **Reporting** | | |
| [JVET-AG0023](file:///\\Users\shanl\Documents\contribution\jvet31ae\current_document.php%3fid=13091) | EE1: Summary report of exploration experiment on neural network-based video coding | E. Alshina, F. Galpin, Y. Li, D. Rusanovskyy, M. Santamaria, J. Ström, R. Chang, Z. Xie (EE coordinators) |
| **EE Technology** | | |
| [JVET-AG0056](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13612) | EE1-1.4: Joint design of rotation and flipping on NNLF | [Z. Xie](mailto:xiezhihuang@oppo.com), [Y. Yu](mailto:yue.yu@oppo.com), [H. Yu](mailto:v-yuhaoping@oppo.com), [D. Wang(OPPO)](mailto:wangdong7@oppo.com), [R. Chang](mailto:renjiechang@tencent.com), [L. Wang](mailto:liqiangwang@tencent.com), [X. Xu](mailto:xiaozhongxu@tencent.com), [S. Liu(Tencent)](mailto:shanl@tencent.com) |
| [JVET-AG0111](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13667) | EE1-2.2 : Content-adaptive LOP filter | [R. Yang](mailto:ruiying.yang@nokia.com), [M. Santamaria](mailto:maria.santamaria_gomez@nokia.com), [F. Cricri](mailto:francesco.cricri@nokia.com), [M. M. Hannuksela](mailto:miska.hannuksela@nokia.com), [H. Zhang](mailto:honglei.1.zhang@nokia.com), [J. Lainema](mailto:jani.lainema@nokia.com), [A. Hallapuro](mailto:antti.hallapuro@nokia.com), [D. Bugdayci Sansli (Nokia)](mailto:done.bugdayci_sansli) |
| [JVET-AG0122](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13678) | EE1-3.1: Deep Reference Frame Generation for Inter Prediction Enhancement | [X. Chen](mailto:cinched@whu.edu.cn), [W. Bao](mailto:baoweijie@whu.edu.cn), [J. Jia](mailto:jiajh2021@whu.edu.cn), [Z. Chen (Wuhan Univ.)](mailto:zzchen@whu.edu.cn), [Z. Liu](mailto:zizhengliu@tencent.com), [X. Xu](mailto:xiaozhongxu@tencent.com), [S. Liu (Tencent)](mailto:shanl@tencent.com) |
| [JVET-AG0130](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13686) | EE1-4.1: Unified CNN-based super resolution for resampling-based video coding | [C. Lin](mailto:linchaoyi.cy@bytedance.com), [Y. Li](mailto:yue.li@bytedance.com), [J. Li](mailto:lijunru@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com) |
| [JVET-AG0162](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13718) | EE1-1.5: Report on implementation of HOP In-loop filter with Transformer blocks | [Y. Li](mailto:yli30@qti.qualcomm.com), [D. Rusanovskyy](mailto:dmytror@qti.qualcomm.com), [M. Karczewicz (Qualcomm)](mailto:martak@qti.qualcomm.com) |
| [JVET-AG0163](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13719) | EE1-2.3: Further complexity reduction on the joint LOP.2 | [T. Shao](mailto:tong.shao@dolby.com), [P. Yin](mailto:pyin@dolby.com), S. McCarthy (Dolby), [J. N. Shingala](mailto:jay.shingala@ittiam.com), [A. Shyam](mailto:ajayshyam@ittiam.com), A. Suneja, S. P. Badya (Ittiam) |
| [JVET-AG0174](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13730) | EE1-1.1: Report on training with HOP architecture change for EE1-0 (variant 1) | [Y. Li](mailto:yue.li@bytedance.com), C. Lin, J. Li, K. Zhang, L. Zhang (Bytedance), [D. Rusanovskyy](mailto:dmytror@qti.qualcomm.com), Y. Li, M. Karczewicz (Qualcomm), [R. Chang](mailto:renjiechang@tencent.com), L. Wang, X. Xu, S. Liu (Tencent) |
| [JVET-AG0175](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13731) | EE1-1.3: Separate models for HOP filter | [Y. Li](mailto:yue.li@bytedance.com), C. Lin, J. Li, K. Zhang, L. Zhang (Bytedance), [F. Galpin (InterDigital)](mailto:Franck.galpin@interdigital.com), [D. Rusanovskyy (Qualcomm)](mailto:dmytror@qti.qualcomm.com), [R. Chang (Tencent)](mailto:renjiechang@tencent.com) |
| **EE Technology related** | | |
| [JVET-AG0155](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13711) | EE1-Related: On Low Complexity Operational Point for In-Loop Filtering | D. Rusanovskyy, Y. Li, M. Karczewicz (Qualcomm) |
| [JVET-AG0156](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13712) | EE1-Related: On LOP2 training process | D. Rusanovskyy, Y. Li, M. Karczewicz (Qualcomm) |
| [JVET-AG0179](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13735) | EE1-1.1-related: HOP filter complexity alignment with wider activation | [Y. Li](mailto:yue.li@bytedance.com), C. Lin, J. Li, K. Zhang, L. Zhang (Bytedance) |
| **Cross Checks** | | |
| [JVET-AG0161](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13717) | Cross-check of EE1-2.3 | D. Rusanovskyy (Qualcomm) |
| [JVET-AG0226](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13799) | Crosscheck of JVET-AG0130 (EE1-4.1: Unified CNN-based super resolution for resampling-based video coding) | [Jiedong Ye](mailto:ye_jd@hust.edu.cn), Xu Li, Yiqing Zhu, Qiong Liu (HUST) |
| [JVET-AG0235](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13808) | Crosscheck of JVET-AG0122 (EE1-3.1: Deep Reference Frame Generation for Inter Prediction Enhancement) | [Z. Xie (OPPO)](mailto:xiezhihuang@oppo.com) |
| [JVET-AG0241](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13814) | Crosscheck of JVET-AG0162 (EE1-1.5: Report on implementation of HOP In-loop filter with Transformer blocks) | [D. Liu (Ericsson)](mailto:du.liu@ericsson.com) |
| [JVET-AG0294](https://jvet-experts.org/doc_end_user/current_document.php?id=13867) | Crosscheck of JVET-AG0056 (EE1-1.4) | Junru Li |

* 1. ***Non-EE Input Contributions***

|  |  |  |
| --- | --- | --- |
| **Reporting** | | |
| [JVET-AG0011](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13382) | JVET AHG report: Neural network-based video coding (AHG11) | E. Alshina, F. Galpin, S. Liu, A. Segall (co-chairs), J. Li, R.-L. Liao, D. Rusanovskyy, M. Santamaria, T. Shao, M. Wien, P. Wu (vice chairs) |
| [JVET-AG0041](file:///C:\Users\e00443164\Downloads\current_document.php%3fid=13594) | AhG11/AhG14 teleconference | [E. Alshina](mailto:elena.alshina@huawei.com), [F. Galpin](mailto:franck.galpin@interdigital.com), [D. Rusanovskyy](mailto:dmytror@qti.qualcomm.com) |
| Proposal | | |
| [JVET-AG0050](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13604) | AHG11: A Transformer-based Intra Luma Enhancement for H.266/VVC | [Hui Yuan](mailto:huiyuan@sdu.edu.cn), Wenrui Lv, [Ming Li](mailto:myron.li@oppo.com), [Dan](mailto:Zou) |
| [JVET-AG0057](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13613) | [AHG11] Study on lower-complexity NNLF | [Z. Xie](mailto:xiezhihuang@oppo.com), [Y. Yu](mailto:yue.yu@oppo.com), [H. Yu](mailto:v-yuhaoping@oppo.com), [D. Wang(OPPO)](mailto:wangdong7@oppo.com) |
| [JVET-AG0069](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13625) | AhG11: LOP with inputs transformed | [D. Liu](mailto:du.liu@ericsson.com), [J. StrÃ¶m](mailto:jacob.strom@ericsson.com), [M. Damghanian](mailto:mitra.damghanian@ericsson.com), [P. Wennersten (Ericsson)](mailto:per.wennersten@ericsson.com) |
| [JVET-AG0114](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13670) | AHG11: On new input and backbone enhancement of model for super-resolution | [Jiedong Ye](mailto:ye_jd@hust.edu.cn), Xu Li, Yiqing Zhu, Qiong Liu (HUST), [Chuan Zhou](mailto:chuan.zhou@vivo.com), M. Rafie, Zhuoyi Lv (vivo) |
| [JVET-AG0129](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13685) | AHG11: Unified CNN-based super resolution | [C. Zhou](mailto:chuan.zhou@vivo.com), [Z. Lv (vivo)](mailto:zhuoyi.lv@vivo.com) |
| Cross Checks | | |
| [JVET-AG0258](file:///D:\Users\e00443164\Documents\___JVET\JVET-AG\AhG11\current_document.php%3fid=13831) | Crosscheck of JVET-AG0069 (AhG11: LOP with inputs transformed) | [Y. Li (Qualcomm)](mailto:yli30@qti.qualcomm.com) |
| JVET-AG0241 | Crosscheck of JVET-AG0162 (EE1-1.5: Report on implementation of HOP In-loop filter with Transformer blocks) | D. Liu (Ericsson) |

1. ***Recommendations***

The AHG recommends:

* Review all input contributions.
* Continue investigating neural network-based video coding tools, including coding performance and complexity.
* Continue collecting training materials for neural network-based video coding tool development.

It was suggested to also test HOP on intra.

It was suggested to investigate combinations of NNVC with ECM (add mandate to AHG), also potentially in subjective test.

[JVET-AG0012](https://jvet-experts.org/doc_end_user/current_document.php?id=13786) JVET AHG report: Enhanced compression beyond VVC capability (AHG12) [M. Karczewicz, Y. Ye, L. Zhang (co-chairs), B. Bross, R. Chernyak, X. Li, K. Naser, Y. Yu (vice-chairs)]

1. ***Activities***

The primary activity of the AHG was the “Exploration experiment on enhanced compression beyond VVC capability” (JVET-AF2024). The combined improvements of the ECM-11.0 over VTM-11.0ecm11.0 anchorfor AI, RA and LB configurations are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | All Intra Main10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -10.81% | -21.80% | -29.33% | 859.1% | 400.6% |
| Class A2 | -16.95% | -28.77% | -31.78% | 860.2% | 405.4% |
| Class B | -11.31% | -26.84% | -25.03% | 784.4% | 411.0% |
| Class C | -11.41% | -16.72% | -17.24% | 810.7% | 381.9% |
| Class E | -15.01% | -24.77% | -23.07% | 750.1% | 431.9% |
| Overall | **-12.81%** | **-23.73%** | **-24.81%** | **808.6%** | **405.0%** |
| Class D | -9.48% | -14.56% | -14.65% | 795.1% | 383.1% |
| Class F | -26.60% | -35.72% | -35.58% | 513.6% | 392.7% |
| Class TGM | -39.53% | -47.43% | -46.69% | 457.3% | 461.2% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Random Access Main 10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -22.47% | -27.84% | -36.88% | 875.4% | 769.1% |
| Class A2 | -26.01% | -36.30% | -40.11% | 824.7% | 958.3% |
| Class B | -20.80% | -35.53% | -33.16% | 714.1% | 774.7% |
| Class C | -22.22% | -27.13% | -27.05% | 748.2% | 829.1% |
| Class E |  |  |  |  |  |
| Overall | **-22.56%** | **-31.91%** | **-33.67%** | **775.1%** | **821.9%** |
| Class D | -23.10% | -28.71% | -28.75% | 768.1% | 898.2% |
| Class F | -28.70% | -37.47% | -37.77% | 614.7% | 505.2% |
| Class TGM | -36.82% | -43.33% | -43.43% | 566.6% | 413.2% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Low Delay B Main 10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -17.76% | -41.31% | -38.53% | 661.6% | 649.7% |
| Class C | -19.49% | -31.72% | -31.66% | 618.2% | 678.1% |
| Class E | -16.28% | -30.12% | -30.39% | 587.1% | 406.5% |
| Overall | **-17.97%** | **-35.32%** | **-34.21%** | **627.8%** | **586.1%** |
| Class D | -21.22% | -34.09% | -33.02% | 659.2% | 772.5% |
| Class F | -24.87% | -39.63% | -38.99% | 575.6% | 491.8% |
| Class TGM | -35.47% | -44.37% | -44.64% | 538.8% | 424.4% |

The rate reduction for natural sequences over VTM in RA configuration for {Y, U, V} increased from ECM-10.0’s {-22.23%, -30.91%, -32.63%} to ECM-11.0’s {-22.56%, -31.91%, -33.67%}. For SCC sequences (class TGM) the rate reduction for RA configuration increased from ECM-10.0’s {-36.63%, -43.02%, -43.08%} to ECM-11.0’s { -36.82%, -43.33%, -43.43%}.

1. ***Contributions***

In addition to 40 EE2 contributions, 43 (comparing to 48 last meeting) EE2-related and AHG12-related contributions were received. The EE2-related and AHG12-related contributions can be subdivided as follows:

* 1. ***Intra (9)***

JVET-AG0075, "AHG12: Adaptive MRL Fusion", S. Blasi, J. Lainema (Nokia)

JVET-AG0078, "AHG12: Intra-prediction using Merged Histogram of Gradients", S. Blasi, I. Zupancic, P. Astola, J. Lainema (Nokia)

JVET-AG0084, "AHG12: DIMD Merge List", M. Blestel, P. Andrivon (Ofinno)

JVET-AG0106, "AHG 12: TIMD merge mode", R. G. Youvalari, M. Abdoli (Xiaomi)

JVET-AG0120, "Non-EE2: On line buffer restriction", Z. Deng, K. Zhang, L. Zhang (Bytedance)

JVET-AG0121, "Non-EE2: Block vector guided LUT for chroma prediction", J. Huo, X. Hao, M. Chen, N. Qiu, Z. Zhang, Y. Ma, F. Yang (Xidian Univ.), M. Li, F. Wang, J. Ren (OPPO)

JVET-AG0138, "Non-EE2: Chroma intra prediction mode reordering", X. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba)

JVET-AG0141, "AHG 12: Occurrence-Based Intra Coding (OBIC)", R. G. Youvalari, M. Abdoli (Xiaomi)

JVET-AG0197, "Non-EE2: Matrix based intra prediction replacing conventional intra modes", B. Ray, H. Wang, V. Seregin, M. Karczewicz, P. Garus (Qualcomm)

* 1. ***IntraTMP and IntraBC (8)***

JVET-AG0063, "EE2-1.2 related: AR-BVP for IntraTMP merge candidates", L. Zhang, Y. Yu, F. Wang, H. Yu, D. Wang (OPPO)

JVET-AG0074, "Non-EE2: IntraTMP with HMVP Candidates", C. Zhu, G. Li, T. Tang, L. Luo, H. Guo (UESTC), Y. Huo, Y. Liu (Transsion)

JVET-AG0080, "EE2-related: Extend block vector prediction for IntraTMP merge candidates", N. Qiu, J. Huo, Y. Ma, F. Yang (Xidian Univ.)

JVET-AG0113, "Non-EE2: FIBC Extension", J. Kim, J. Kang, H. Han, H. Choi (HNU), W. Lim, S.-C. Lim (ETRI)

JVET-AG0186, "Non-EE2: FIBC Extension", H.-J. Jhu, X. Xiu, W. Chen, C.-W. Kuo, N. Yan, C. Ma, X. Wang (Kwai)

JVET-AG0193, "Non-EE2: Enhancements on IntraTMP", W. Chen, X. Xiu, C. Ma, H.-J. Jhu, C.-W. Kuo, N. Yan, X. Wang (Kwai)

JVET-AG0231, "EE2-1.7b related: unrestricted 1.7b performances" F. Le Léannec, K. Naser, T. Dumas, Y. Chen, M. Radosavljević, T. Poirier, (InterDigital)

JVET-AG0243, " EE2-1.4: IntraTMP extension to DIMD", K. Naser, F. Le Léannec, T. Poirier, H. Guermoud, T. Dumas, (InterDigital)

* 1. ***Inter (10)***

JVET-AG0073, "Non-EE2: Chained motion vector prediction", Y. Kidani, H. Kato, K. Kawamura (KDDI)

JVET-AG0096, "Non-EE2: On temporal buffer handling", Z. Deng, K. Zhang, L. Zhao, L. Zhang (Bytedance)

JVET-AG0125, "AHG12: Parallel friendly use of boundary distortion for DMVR", K. Andersson, R. Yu (Ericsson)

JVET-AG0149, "Non-EE2: Improvements to subblock merge mode", J. Chen, R.-L. Liao, Y. Zheng, X. Li, Y. Ye (Alibaba)

JVET-AG0150, "EE2-related: Adaptive GPM blending", L. Zhao, K. Zhang, L. Zhang (Bytedance)

JVET-AG0172, "EE2-related: Chroma LIC derivation with template costs", T. Bae, S. Deshpande (Sharp)

JVET-AG0194, "Non-EE2: Reference filtering for inter-prediction", A. Filippov, V. Rufitskiy, K. Suverov (Ofinno)

JVET-AG0195, "Non-EE2: LIC model parameter inheritance for merge modes", C.-C. Chen, H. Huang, V. Seregin, M. Karczewicz

JVET-AG0200, "Non-EE2: Inter CCP merge mode with zero luma CBF", Z. Deng, K. Zhang, L. Zhang (Bytedance)

JVET-AG0202, "Non-EE2: Geometry partitioning mode with inter prediction and intra block copy", Y. Wang, K. Zhang, L. Zhang (Bytedance)

* 1. ***In Loop Filters (3)***

JVET-AG0065, "Non-EE2: Adaptive coefficient precision for CCALF", N. Song, Y. Yu, H. Yu, D. Wang (OPPO)

JVET-AG0198, "Non-EE2: Coding Information based ALF Classification", W. Yin, K. Zhang, Y. Wang, Z. Deng, L. Zhao, N. Zhang, M. Salehifar, L. Zhang (Bytedance)

JVET-AG0233, " EE2-related: Adaptive precision for CCALF coefficients", [N. Hu](mailto:nanh@qti.qualcomm.com), [M. Karczewicz](mailto:martak@qti.qualcomm.com), [V. Seregin](mailto:vseregin@qti.qualcomm.com), [H. Wang (Qualcomm)](mailto:hongtaow@qti.qualcomm.com)

* 1. ***Transform (4)***

JVET-AG0062, "Non-EE2: Multiple Transform Sets Selection for LFNST/NSPT", F. Wang, Y. Yu, H. Yu, D. Wang (OPPO)

JVET-AG0208, "EE2-related: On LFNST/NSPT index signalling", M. Koo, J. Zhao, J. Lim, S. Kim (LGE)

JVET-AG0230, "EE2-3.3 related: On Inter-LFNST”, S. Puri, K. Naser, C. Bonnineau, F. Le Léannec", (InterDigital)

JVET-AG0237, "EE2-3.3 related: Fix on LFNST/NSPT index signalling", M. Koo, J. Zhao, J. Lim, S. Kim (LGE)

* 1. ***Coefficients and Entropy Coding (7)***

JVET-AG0064, "AHG12: On Context modeling in Chroma Coefficient Coding", L. Xu, H. Yu, Y. Yu, D. Wang (OPPO)

JVET-AG0066, "EE2-3.2 related: On Regular Residual Coding", Y. Yu, L. Xu, J. Gan, H. Yu, L. Zhang, H. Huang, F. Wang, Z. Xie, N. Song, D. Wang (OPPO)

JVET-AG0068, "EE2-3.2 related: On CABAC bin budget", J. Gan, Y. Yu, H. Yu (OPPO)

JVET-AG0108, "AhG12 Entropy coding extension", F. Galpin, F. Lo Bianco, C. Salmon-Legagneur, M. Balcilar (InterDigital)

JVET-AG0110, "Ahg12: 16 States TCQ with State Exchange", M. Balcilar, K. Naser, Y. Chen, Franck Galpin, Fabrice Léannec (InterDigital)

JVET-AG0102, "AHG12: New context model parameters for low delay B condition", R.-L. Liao, Y. Ye, J. Chen, X. Li (Alibaba)

JVET-AG0185, "Non-EE2: Slice based Rice parameter selection for transform skip residual coding", H.-J. Jhu, X. Xiu, W. Chen, C.-W. Kuo, N. Yan, C. Ma, X. Wang (Kwai)

* 1. ***Partitioning (1)***

JVET-AG0126, "AhG12: On ECM temporal partitioning prediction", G. Laroche, P. Onno (Canon)

* 1. ***Other (1)***

JVET-AG0116, "AHG12: GOP-based RPR encoder control for ECM", K. Andersson, J. Ström, P. Wennersten, R. Yu, W. Ahmad (Ericsson)

1. ***Recommendations***

The AHG recommends to:

* To review all the related contributions.

[JVET-AG0013](https://jvet-experts.org/doc_end_user/current_document.php?id=13787) JVET AHG report: Film grain technologies (AHG13) [W. Husak, P. de Lagrange (co-chairs), A. Duenas, D. Grois, Y. He, X. Meng, M. Radosavljević, A. Segall, G. Teniou, A. Tourapis (vice-chairs)]

1. ***Discussion***

For this AHG period, there were three major topics of conversation. ISO/IEC 23002-9 DTR was edited to reflect feedback from ITU and ISO editors. The group also discussed content containing film grain for development, demonstrations, and testing of systems and processes. Also discussed were metrics and characterization for film grain synthesis. Finally, the group discussed potential future topics for the Technical Report.

Additional editing of the DTR continues due to comments that originated from the ISO and ITU editors. Most were changes due to editorial comments. Several comments related to references to AOM documents and github sites remain to be addressed.

An extensive amount of time was spent on content. Apple created a dataset using both content captured by Apple and content that is publicly available. The dataset is comprised of approximately 20 sequences with 27 variants and stored as YUVs along with proxies. Several people have copied the data and begun tests.

Disney reported on their tests on the Tears of Steel content. The tests are an effort to reduce the number of clips. The methodology and results were shared. In short, the method subtracted the original video from the video under test and analyzed the film grain characteristics. This method seems to be popular as it has been independently explored by a number of companies.

The group discussed there are outside organizations that would find value in the report. One observation is that the TR is background on film grain but not what would be considered an Engineering Guideline (EG) or a Recommended Practice (RP) in other organizations. JVET needs to consider perhaps doing best practices (ISO EG/RP) – as part of the AHG report v2. Capturing the experiences of the group and perhaps even the creatives could be useful to outside organizations.

1. **Related *contributions***

Eight contributions related to AHG13 were identified as of 01/16/2024.

* One contribution was the AHG report:
  + JVET-AG0013 JVET AHG report: Film grain technologies (AHG13)
* Seven other contributions were uploaded at the time of the report drafting:
  + JVET-AG0101 Film grain adaptive SEI message
  + JVET-AG0140 FGS Extension SEI message for spatial adaptation
  + JVET-AG0153 Proposed FGS applications
  + JVET-AG0160 FGS Extension SEI message useful descriptors
  + JVET-AG0215 Region-dependent film grain characteristics
  + JVET-AG0228 source for scanned film test sequences
  + JVET-AG0290 Compressibility analysis of Film Grain test sequences
  + JVET-AG0295 On Film Grain Synthesis Subjective Evaluation
  1. ***Contributions***

There were seven contributions uploaded other than the AHG report.

* + 1. **JVET-*AG0101* Film grain adaptive SEI message**

It is proposed to define a new film grain adaptive SEI message to specify the regions that the film grain applies, wherein:

* these regions may include the external regions or/and the internal regions.
* external regions refer to the existing available regions defined through different type of SEI messages or other approach.
* internal regions refer to the regions defined in the film grain adaptive SEI message.

In the 32nd JVET meeting, the contribution JVET-AF0142 [1] proposes to address the film grain synthesis applied locally on the picture by using the annotated regions as a mask and the contribution JVET-AF0144 [2] proposes to address the film grain synthesis applied locally on a per image basis by using the auxiliary data as a mask. Furthermore, the conclusions recorded in the meeting notes [3] confirm that the functionality as such was assessed to be useful and suggest to study the possibility of defining a new SEI message rather than changing an existing one.

Since it has been recognized that film grain may not necessarily be applicable to the entire picture but only in some parts of the picture, such as one or more annotated regions, forefront objects, etc., it would be a simple approach if these existing available regions can be used in the newly defined SEI message.

On the other hand, these existing available regions may not meet all the requirements, so it is also necessary to provide a way to describe the film grain region in the newly defined SEI message.

* + 1. **JVET-*AG0140* FGS Extension SEI message for spatial adaptation**

The proposed contribution is a follow-up of previous meeting inputs on FGS spatial adaptation. Taking into account the comments made on the importance to not modify the existing FGC SEI message, this proposal follows the suggested approach to develop a dedicated FGS Extension SEI message dealing with relative spatial application of film grain values in the picture.

During the last JVET meeting #32 in Hannover (Oct. 23) 2 contributions were presented (JVET-AF0142 and JVET-AF0144) introducing the functionality to adapt locally in the image the application of film grain values.

The proposed way forward was to revisit the proposals by developing a dedicated SEI message. The present contribution proposes a Film Grain Extension SEI message to convey the information of where film grain values are applied.

* + 1. ***JVET*-AG0153 Proposed FGS applications**

Proposal to add a new use case of FGS: Artistic intent simulation, in which FGS is used without any reference, to be differentiated from the artistic intent recovery that relies on FGS to maintain the similarity with the original content.

Film Grain Synthesis is primarily meant to optimize the video coding efficiency due to the particular nature of the grain, considered as noise with no spatio-temporal correlation. The usage of film grain as listed in the FGS Draft TR may serve the objective to maintain the artistic intent while the video source has been denoised for encoding efficiency purposes. In this case the synthetized grain is defined to emulate the original grain. The second listed scenario is the usage of FGS to mark visual artifacts such as block effects, ringing around edges or even too strong deblocking filters. Those 2 use cases are listed in clause 6.4 of the draft TR.

However, another scenario, briefly mentioned in the introduction of the Draft TR relates to the use of FGS for other artistic intent objectives (called here simulation). Film grain synthesis is a tool for increasing the depth perception, for harmonizing the cameras optics quality or simply adding an artificial movie effect, such as for some game cinematics. It is then proposed to list this type of use case and application of film grain in the Draft TR.

* + 1. ***JVET*-AG0160 FGS Extension SEI message useful descriptors**

While defining an FGS extension SEI message primarily addressing the functionality of spatial adaptation, it was felt useful to also include descriptive parameters of the film grain synthesis such as the type of grain that is meant to be reproduced, the purpose of the film grain, such as creative artistic intent or artefact masking, and its essentiality from a service viewpoint.

* + 1. **JVET-AG0215 Region-dependent film grain characteristics**

This contribution, while stating that the current FGC SEI message fits its purpose to describe the grain present on scanned film or added for artistic intent on motion pictures, and enables to recover a similar grain by driving synthesis at decoder side at reasonable cost when the grain has been lost in the encoding process for broadcast or streaming purpose, acknowledges that for other uses cases, it may be desirable to describe different film grain in different regions of the picture, for example when the picture is a composite of different sources.

A region-dependent film grain characteristics SEI message is proposed, that includes the existing FGC SEI message syntax, notably independent intensity intervals with model parameters for each interval, and adds elements to describe a list of picture regions, and associate each region with a set of intervals. The commonalities with the FGC SEI message are believed to provide a simple migration path, and simple conversion back to FGC SEI messages for parts of pictures by a network processor, for example in the context of subpictures.

* + 1. **JVET-*AG0228* source for scanned film test sequences**

This contribution reports about old film scanned by Prelinger Archives and made available on the Internet Archive website, and recommends to select a few samples to be included as candidate test sequences for FGC SEI visual testing activities planned in JVET-AD2022.

During the 30th JVET meeting, an expert viewing test was conducted to evaluate the benefit of using film grain synthesis driven by an FGC SEI message on top of VVC encodings, for a better similarity to video sources that originally contained grain and potential bitrate savings. Results were reported in JVET-AD0382, concluding that significant improvement were observed over bitrate scanning the usual MOS range, and that similarity with the source was decently preserved (-0.5 MOS) using grain synthesis on top of high-quality encoding just removing grain, leading to 90% rate savings compared to encodings just preserving grain. JVET-AD0276 reported similar conclusions based on InterDigital internal informal tests performed before the meeting. Both tests were using the same test sequences, including two sequences from a short test film by SVT shot in 2004 on 65mm film. It was questioned whether similar benefits would be observed using older film stock, that was anticipated to exhibit significantly different characteristics, likely coarser and stronger.

Since then, relevant good quality scans of older film stock was actively sought after to enable such experiments.

This contribution reports about one source where relevant material could be found, and provides information about some samples of this material.

* + 1. **JVET-AG0290 Compressibility analysis of Film Grain test sequences**

This contribution has not been uploaded.

* + 1. **JVET-*AG0295* Frequency domain Film Grain Objective Metrics with Adaptive Region Selection**

This contribution has not been uploaded.

1. ***Recommendations***

The AHG recommends:

* the related input contributions are reviewed;
* testing of FGC be discussed;
* continued conformance discussion;
* future extensions to the Technical Report;
* SEI message extensions; and
* continue the study of film grain technologies in JVET.

Upon a question, it was reported that no progress was made on the question about potential definition of conformance.

It was commented that, in producing the DTR, one problem was referencing external entities/repositories that may not be stable (was commented as inappropriate by ISO staff). It was suggested that a possible solution might be to refer to specific versions of such references. Offline discussion (G. Sullivan, W. Husak, et al.) and later report during the meeting.

[JVET-AG0014](https://jvet-experts.org/doc_end_user/current_document.php?id=13597) JVET AHG report: NNVC software development (AHG14) [F. Galpin (chair), Y. Li, Y. Li, J. N. Shingala, L. Wang, Z. Xie (vice-chairs)]

1. **Software *development***
   1. **Location**

NNVC repository is located at <https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/VVCSoftware_VTM>

NNVC software is based on VTM-11.0 with enabled MCTF including the update from JVET-V0056, GOP32, and enabling deblocking in the RDO.

NNVC-7.1 anchor at <https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/nnvc-ctc/-/blob/master/Anchor%20performance/NNVC-current_Reporting_Template-with-anchor-data.xlsm?ref_type=heads> is used for NNVC performance evaluation.

* 1. **Software *changes***

Several commits were merged in the NNVC repository.

The following changes were integrated:

* HOP related:
  + JVET-AF0155/AF0180: luma/chroma balance
  + JVET-AF0296: learning rate
  + JVET-AF0150: batch size
  + improved documentation
* LOP related:
  + JVET-AF0043: LOP.2 training and model.
* LOP/HOP:
  + Training speed-up
  + Unification of code, training scripts and configuration files
  + JVET-AF0172: cleaning scale flag
  + JVET-AF0205: padding policy
  + JVET-AF0085: residual offset adjustment
  + JVET-AF0193: decoder complexity optimization (default off)
* JVET-AF0139: NN intra lower complexity
* JVET-AF0152: SADL v7
* fixes:
  + build issues fixes.
  + default to AVX2 for SADL
  + remove deprecated code.
  1. **Software *version***

NNVC-7.1 was tagged November 29th, 2023

NNVC-7.0 was tagged November 3rd, 2023

NNVC-6.1 was tagged September 25th, 2023 (fix)

NNVC-6.0 was tagged September 6th, 2023.

NNVC-5.1 was tagged July 19th, 2023.

NNVC-5.0 was tagged May 11th, 2023.

NNVC-3.0 (a.k.a VTM-11.0\_nnvc3.0) was tagged December 1st 2022.

NCS-1.0 (a.k.a NNVC-3.0wip2) was tagged September 4th 2022 (first release containing the FilterSets, using NNVC 2.0 as a base).

VTM-11.0\_nnvc-2.0 was tagged August 4th 2022 (add deblocking in RDO).

VTM-11.0\_nnvc-1.0 was tagged May 6th 2021 (VTM-11.0 base with MCTF enabled).

1. **CTC performance**

See configurations section for naming convention.

* 1. **Comparison *to VTM***
     1. **NNVC-6.0 VTM vs NNVC-7.1 VTM**

NNVC-7.1 in VTM mode performance are the same as the ones in NNVC-5.0/NNVC-6.0.

* + 1. **NNVC-7.1 VTM vs NNVC-7.1 anchor**

The NNVC-7.1 anchor includes LOP.2 filter and Intra Prediction tools activated.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over NNVC-7.0 == 7.1** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -7.91% | -10.09% | -10.61% | -8.48% | -11.15% | -10.66% | 129% | 9031% |
| Class A2 | -6.79% | -12.68% | -9.74% | -7.00% | -11.83% | -7.35% | 126% | 8426% |
| Class B | -6.57% | -13.94% | -12.70% | -6.78% | -12.92% | -11.88% | 127% | 9076% |
| Class C | -6.58% | -14.87% | -14.43% | -7.15% | -13.24% | -12.86% | 120% | 7892% |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -6.89% | -13.17% | -12.15% | -7.26% | -12.43% | -10.99% | 126% | 8606% |
| Class D | -7.17% | -14.84% | -15.00% | -6.49% | -13.17% | -13.07% | 119% | 7530% |
| Class F | -3.71% | -9.49% | -9.41% | -4.23% | -10.19% | -9.77% | 136% | 3684% |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Low delay B Main10** | | | | | | | |
|  | **BD-rate Over NNVC-7.0 == 7.1** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 |  |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |  |
| Class B | -5.29% | -13.38% | -13.25% | -6.09% | -14.72% | -17.51% | 120% | 8941% |
| Class C | -5.53% | -13.74% | -14.88% | -7.36% | -15.51% | -16.98% | 113% | 7756% |
| Class E | -6.13% | -8.86% | -12.10% | -7.76% | -10.47% | -13.50% | 136% | 9072% |
| **Overall** | -5.58% | -12.37% | -13.51% | -6.93% | -13.92% | -16.33% | 121% | 8558% |
| Class D | -6.54% | -13.06% | -15.42% | -7.34% | -14.20% | -17.88% | 110% | 7563% |
| Class F | -3.14% | -8.73% | -8.67% | -4.24% | -9.89% | -13.21% | 129% | 4227% |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over NNVC-7.0 == 7.1** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -8.62% | -12.69% | -12.29% | -9.00% | -14.28% | -12.61% | 187% | 6908% |
| Class A2 | -7.15% | -12.70% | -11.26% | -7.47% | -12.25% | -9.23% | 186% | 5880% |
| Class B | -7.35% | -13.02% | -13.55% | -7.37% | -13.33% | -13.41% | 181% | 5701% |
| Class C | -7.53% | -13.81% | -14.47% | -7.93% | -14.27% | -14.56% | 168% | 4342% |
| Class E | -10.71% | -14.15% | -15.12% | -10.94% | -13.61% | -14.11% | 176% | 6176% |
| **Overall** | -8.13% | -13.28% | -13.42% | -8.38% | -13.56% | -12.95% | 179% | 5644% |
| Class D | -7.65% | -12.72% | -15.05% | -7.48% | -13.58% | -15.42% | 160% | 4159% |
| Class F | -5.27% | -10.79% | -10.39% | -5.08% | -11.26% | -11.78% | 141% | 4847% |

Note: Results from Tencent, crosschecked by InterDigital for gains.

* + 1. **NNVC-7.1 VTM mode vs NNVC-7.1 HOP.2**

The NNVC-7.1 where LOP.2 filter is replaced by HOP.2 and Intra Prediction tools activated.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over NNVC-7.1 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -15.29% | -8.16% | -18.63% | -16.62% | -14.19% | -19.81% | 444% | 212316% |
| Class A2 | -14.52% | -11.44% | -9.50% | -14.18% | -11.71% | -5.61% | 420% | 203481% |
| Class B | -12.30% | -12.31% | -12.15% | -11.43% | -12.31% | -11.21% | 437% | 214243% |
| Class C | -13.23% | -16.68% | -16.88% | -12.41% | -12.85% | -12.90% | 351% | 197379% |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -13.59% | -12.47% | -14.18% | -13.28% | -12.71% | -12.26% | 410% | 207086% |
| Class D | -14.34% | -15.82% | -17.40% | -11.21% | -12.00% | -11.59% | 342% | 185469% |
| Class F | -8.28% | -11.67% | -10.63% | -8.37% | -11.91% | -10.79% | 647% | 97257% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
|  |  |  |  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | | | | |
|  | **BD-rate Over NNVC-7.1 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
| Class A2 | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
| Class B | -11.10% | -6.55% | -7.45% | -11.04% | -11.29% | -12.22% | 413% | 214332% |
| Class C | -12.28% | -13.55% | -14.46% | -12.60% | -13.84% | -13.48% | 341% | 197098% |
| Class E | -13.31% | -3.89% | -9.88% | -14.03% | -8.57% | -11.80% | 907% | 272167% |
| **Overall** | -12.05% | -8.22% | -10.40% | -12.31% | -11.46% | -12.53% | 472% | 221253% |
| Class D | -13.56% | -12.14% | -14.84% | -12.38% | -12.34% | -13.48% | 324% | 184900% |
| Class F | -8.83% | -9.53% | -10.94% | -10.01% | -12.18% | -15.42% | 630% | 116303% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
|  |  |  |  |  |  |  |  |  |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over NNVC-7.1 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -12.82% | -8.37% | -15.39% | -14.38% | -13.29% | -17.63% | 512% | 151639% |
| Class A2 | -12.35% | -10.32% | -9.22% | -13.17% | -9.89% | -6.44% | 352% | 121892% |
| Class B | -11.01% | -10.87% | -11.24% | -10.94% | -11.58% | -11.47% | 322% | 114567% |
| Class C | -12.13% | -12.55% | -15.09% | -11.86% | -12.10% | -14.14% | 245% | 81849% |
| Class E | -15.62% | -14.68% | -14.86% | -15.75% | -13.67% | -14.67% | 350% | 127857% |
| **Overall** | -12.55% | -11.37% | -13.05% | -12.89% | -12.05% | -12.78% | 337% | 114638% |
| Class D | -11.72% | -11.38% | -15.01% | -10.82% | -11.27% | -12.86% | 223% | 76334% |
| Class F | -8.85% | -11.06% | -10.75% | -8.54% | -11.48% | -11.55% | 212% | 99222% |

Note: Results from Tencent, crosschecked by Oppo.

* + 1. **NNVC-7.1 VTM mode vs NNVC-7.1 VTM mode + LOP.2**

The results reflect LOP only performance.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **LOP2 BD-rate Over NNVC-7.0 == 7.** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT |
| Class A1 | -5.92% | -8.80% | -9.06% |  |  |  | 121% | 12269% |
| Class A2 | -5.59% | -12.12% | -8.92% |  |  |  | 113% | 9036% |
| Class B | -4.99% | -13.19% | -12.07% |  |  |  | 115% | 8850% |
| Class C | -5.12% | -13.71% | -13.20% |  |  |  | 105% | 8032% |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -5.33% | -12.24% | -11.14% |  |  |  | 113% | 9245% |
| Class D | -6.01% | -14.08% | -13.94% |  |  |  | 104% | 7464% |
| Class F |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | |
|  | **LOP BD-rate Over NNVC-7.0 == 7.** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT |
| Class A1 | -4.64% | -9.01% | -8.52% |  |  |  | 118% | 8037% |
| Class A2 | -4.48% | -10.12% | -8.62% |  |  |  | 108% | 6290% |
| Class B | -4.39% | -10.32% | -10.64% |  |  |  | 106% | 5134% |
| Class C | -4.64% | -11.11% | -11.70% |  |  |  | 103% | 3205% |
| Class E | -6.42% | -9.58% | -11.35% |  |  |  | 106% | 4491% |
| **Overall** | -4.84% | -10.12% | -10.31% |  |  |  | 108% | 5157% |
| Class D | -5.00% | -10.10% | -12.34% |  |  |  | 101% | 3109% |
| Class F |  |  |  |  |  |  |  |  |

Note: Results from Qualcomm, crosschecked by XX.

* + 1. **NNVC-7.1 VTM mode vs NNVC-7.1 VTM mode + HOP.2**

The results reflect HOP only performance.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over NNVC-7.0 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -13.58% | -7.16% | -17.71% |  |  |  |  |  |
| Class A2 | -13.47% | -10.84% | -8.82% |  |  |  |  |  |
| Class B | -10.95% | -11.39% | -11.44% |  |  |  |  |  |
| Class C | -11.97% | -15.75% | -15.73% |  |  |  |  |  |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -12.25% | -11.60% | -13.32% |  |  |  |  |  |
| Class D | -13.29% | -15.18% | -16.83% |  |  |  |  |  |
| Class F | -7.44% | -10.89% | -9.97% |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over NNVC-7.0 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -9.33% | -4.74% | -12.49% |  |  |  |  |  |
| Class A2 | -9.99% | -7.77% | -6.71% |  |  |  |  |  |
| Class B | -8.38% | -8.06% | -8.01% |  |  |  |  |  |
| Class C | -9.58% | -9.30% | -11.86% |  |  |  |  |  |
| Class E | -11.86% | -10.17% | -10.83% |  |  |  |  |  |
| **Overall** | -9.66% | -8.09% | -9.86% |  |  |  |  |  |
| Class D | -9.30% | -8.84% | -12.04% |  |  |  |  |  |
| Class F | -7.00% | -9.17% | -8.88% |  |  |  |  |  |

* + 1. **NNVC-7.1 VTM mode vs NNVC-7.1 VTM + NN intra low complexity**

The results reflect NN intra low complexity only performance.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over NNVC-7.1 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -3.99% | -3.52% | -3.91% | -4.20% | -3.52% | -3.80% | 165% | 429% |
| Class A2 | -2.81% | -2.37% | -2.88% | -2.81% | -2.15% | -2.53% | 171% | 573% |
| Class B | -2.92% | -2.67% | -2.83% | -2.74% | -2.43% | -2.43% | 172% | 619% |
| Class C | -2.40% | -2.08% | -2.07% | -2.53% | -1.61% | -1.52% | 163% | 596% |
| Class E | -4.14% | -3.91% | -3.51% | -3.99% | -3.45% | -2.21% | 169% | 633% |
| **Overall** | -3.17% | -2.84% | -2.96% | -3.16% | -2.55% | -2.44% | 168% | 572% |
| Class D | -2.06% | -1.69% | -1.90% | -2.07% | -1.38% | -1.01% | 158% | 627% |
| Class F | -1.74% | -1.41% | -1.35% | -1.81% | -1.26% | -1.18% | 143% | 353% |

Note: Results from InterDigital, crosschecked by Oppo.

* 1. ***Comparison to previous NNVC - NNVC-6.0 anchor***
     1. ***NNVC*-6.0 anchor vs NNVC-7.1 anchor**

The results reflect LOP progress between NNVC 6.0 (as anchor) and NNVC 7.1 (as test).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over NNVC 6.0** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -0.73% | -5.39% | -4.78% | 0.30% | -7.16% | -5.81% |  |  |
| Class A2 | -0.25% | -7.14% | -4.80% | -0.51% | -7.88% | -4.53% |  |  |
| Class B | -0.29% | -6.01% | -5.47% | -0.85% | -7.19% | -7.04% |  |  |
| Class C | 0.05% | -4.53% | -5.12% | -0.94% | -5.56% | -7.49% |  |  |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -0.28% | -5.72% | -5.10% | -0.58% | -6.89% | -6.41% |  |  |
| Class D | 0.46% | -5.54% | -5.53% | -0.78% | -6.42% | -7.38% |  |  |
| Class F | -0.66% | -3.76% | -4.60% | -1.21% | -5.62% | -6.47% |  |  |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! |
|  | **Low delay B Main10** | | | | | | | |
|  | **BD-rate Over NNVC 6.0** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! |  |  |
| Class A2 | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! |  |  |
| Class B | -0.47% | -5.48% | -5.76% | -0.99% | -8.84% | -11.46% |  |  |
| Class C | -0.31% | -2.82% | -6.31% | -1.83% | -5.38% | -13.65% |  |  |
| Class E | -1.29% | -4.85% | -7.30% | -2.27% | -8.32% | -10.72% |  |  |
| **Overall** | -0.62% | -4.44% | -6.33% | -1.59% | -7.56% | -12.00% |  |  |
| Class D | 0.02% | -2.38% | -5.89% | -1.65% | -5.19% | -11.89% |  |  |
| Class F | -1.11% | -3.84% | -4.37% | -2.13% | -6.69% | -9.37% |  |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over NNVC 6.0** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -0.18% | -5.86% | -3.93% | 1.23% | -7.29% | -5.52% |  |  |
| Class A2 | -0.57% | -5.04% | -3.57% | -0.62% | -6.25% | -4.05% |  |  |
| Class B | -0.35% | -4.77% | -5.23% | -0.37% | -6.21% | -6.83% |  |  |
| Class C | -0.21% | -4.39% | -4.86% | -0.35% | -6.55% | -7.42% |  |  |
| Class E | -0.47% | -5.44% | -6.80% | -0.86% | -6.31% | -8.28% |  |  |
| **Overall** | -0.35% | -5.02% | -4.91% | -0.22% | -6.49% | -6.52% |  |  |
| Class D | -0.08% | -4.81% | -6.03% | -0.32% | -7.18% | -8.29% |  |  |
| Class F | -0.75% | -5.00% | -4.88% | -0.80% | -6.78% | -7.49% |  |  |

* + 1. **NNVC-6.0 HOP vs NNVC-7.1 HOP**

The results reflect HOP progress between NNVC 6.0 (as anchor) and NNVC 7.1 (as test).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 HOP** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -3.44% | 18.42% | 11.54% | -3.47% | 14.01% | 11.76% |  |  |
| Class A2 | -2.95% | 17.74% | 18.25% | -2.78% | 16.39% | 18.04% |  |  |
| Class B | -1.44% | 20.17% | 17.60% | -1.09% | 19.65% | 18.91% |  |  |
| Class C | -1.65% | 16.96% | 16.42% | -1.47% | 17.49% | 18.43% |  |  |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -2.20% | 18.48% | 16.20% | -2.01% | 17.29% | 17.18% |  |  |
| Class D | -1.55% | 16.41% | 15.29% | -1.01% | 17.69% | 17.88% |  |  |
| Class F | -1.79% | 9.86% | 9.32% | -1.13% | 11.55% | 11.23% |  |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Low delay B Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 HOP** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! |  |  |
| Class A2 | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! |  |  |
| Class B | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! |  |  |
| Class C | -1.63% | 16.82% | 11.43% | -1.66% | 16.35% | 11.10% |  |  |
| Class E | -1.60% | 15.29% | -2.81% | -2.12% | 10.42% | -3.97% |  |  |
| **Overall** | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! |  |  |
| Class D | -1.82% | 15.20% | 3.94% | -1.79% | 14.20% | 3.02% |  |  |
| Class F | -1.88% | 13.07% | 4.68% | -1.54% | 12.45% | 7.95% |  |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 HOP** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -2.25% | 17.82% | 14.03% | -2.24% | 16.11% | 13.06% |  |  |
| Class A2 | -3.14% | 16.68% | 14.74% | -3.18% | 17.17% | 14.67% |  |  |
| Class B | -1.49% | 15.42% | 14.91% | -1.40% | 17.06% | 17.16% |  |  |
| Class C | -1.66% | 12.43% | 12.91% | -1.54% | 15.24% | 16.42% |  |  |
| Class E | -1.53% | 13.98% | 15.33% | -1.20% | 14.27% | 17.38% |  |  |
| **Overall** | -1.93% | 15.12% | 14.36% | -1.83% | 16.05% | 15.93% |  |  |
| Class D | -1.57% | 12.83% | 13.43% | -1.43% | 17.15% | 18.42% |  |  |
| Class F | -1.85% | 8.35% | 8.89% | -1.51% | 10.59% | 12.97% |  |  |

* 1. ***Comparison to NNVC-7.1 anchor***
     1. **NNVC-7.1 anchor vs NNVC-7.1 HOP.2**

The results below reflect LOP/HOP performance differences (LOP set as anchor, HOP as test).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over NNVC-7.1** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -8.00% | 2.04% | -8.95% | -8.89% | -3.77% | -10.33% |  |  |
| Class A2 | -8.27% | 1.41% | 0.15% | -7.66% | -0.03% | 1.77% |  |  |
| Class B | -6.15% | 1.99% | 0.66% | -4.97% | 0.79% | 0.77% |  |  |
| Class C | -7.13% | -2.16% | -2.89% | -5.65% | 0.36% | -0.14% |  |  |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -7.21% | 0.78% | -2.31% | -6.47% | -0.40% | -1.49% |  |  |
| Class D | -7.74% | -1.17% | -2.98% | -5.02% | 1.38% | 1.39% |  |  |
| Class F | -4.83% | -2.53% | -1.47% | -4.39% | -2.25% | -1.24% |  |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Low delay B Main10** | | | | | | | |
|  | **BD-rate Over NNVC-7.1** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! |  |  |
| Class A2 | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! |  |  |
| Class B | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! |  |  |
| Class C | -7.16% | 0.44% | 0.68% | -5.64% | 2.23% | 4.49% |  |  |
| Class E | -7.63% | 5.46% | 2.64% | -6.71% | 2.24% | 2.00% |  |  |
| **Overall** | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! |  |  |
| Class D | -7.52% | 1.01% | 0.76% | -5.43% | 2.83% | 5.77% |  |  |
| Class F | -5.93% | -0.76% | -2.42% | -6.08% | -2.85% | -2.61% |  |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over NNVC-7.1** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -4.59% | 4.88% | -3.46% | -5.89% | 1.08% | -5.79% |  |  |
| Class A2 | -5.62% | 2.85% | 2.26% | -6.12% | 2.73% | 3.07% |  |  |
| Class B | -3.98% | 2.77% | 2.73% | -3.83% | 2.20% | 2.29% |  |  |
| Class C | -5.02% | 1.56% | -0.72% | -4.26% | 2.48% | 0.37% |  |  |
| Class E | -5.52% | -0.68% | 0.30% | -5.29% | -0.11% | -0.70% |  |  |
| **Overall** | -4.84% | 2.29% | 0.45% | -4.89% | 1.78% | 0.15% |  |  |
| Class D | -4.42% | 1.62% | -0.04% | -3.63% | 2.86% | 2.88% |  |  |
| Class F | -3.87% | -0.22% | -0.36% | -3.63% | -0.29% | 0.42% |  |  |

* + 1. **NNVC-7.1 anchor vs NNVC-7.1 NN intra lower complexity**

The results below show the difference in performance of NN intra (as anchor) to NN intra lower complexity (as test).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over NNVC 7.1** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | 0.31% | 0.31% | 0.72% | 0.31% | 0.27% | 0.37% | 72% | 90% |
| Class A2 | -0.03% | -0.05% | -0.21% | -0.06% | -0.25% | -0.30% | 71% | 86% |
| Class B | 0.25% | 0.04% | 0.12% | 0.26% | -0.04% | -0.07% | 70% | 82% |
| Class C | 0.70% | 0.61% | 0.85% | 0.63% | 0.65% | 0.73% | 70% | 75% |
| Class E | 0.60% | 0.63% | 0.41% | 0.77% | 0.82% | 0.42% | 70% | 81% |
| **Overall** | 0.37% | 0.30% | 0.38% | 0.38% | 0.27% | 0.23% | 71% | 82% |
| Class D | 0.83% | 0.64% | 0.87% | 0.76% | 0.29% | 1.23% | 71% | 71% |
| Class F | 0.41% | 0.33% | 0.49% | 0.34% | 0.29% | 0.37% | 81% | 83% |

* 1. ***Other* tools**

Other results remain the same as tools were not changed.

1. **Discussions**
   1. **Discussions *on optimization level in NNVC***
      1. **Overall optimization level**

Current optimization level of NNVC is as follow:

* The legacy VTM code is considered less optimized than the ECM part as ECM integrated more SIMD optimization functions.
* The NN based code is mainly executed inside the SADL library, with some parts, non-SIMD optimized, executed on the NNVC side.

The complexity at encoder, as measured by the encoding time, is split in half between the VTM part and the NN based part. It is considered that the NN based part is more optimized than the VTM part since almost all the code is SIMD optimized.

The complexity at decoder, as measured by the encoding time, is mainly driven by the NN based part.

* + 1. **NN based optimization level**

The SIMD optimization parts in SADL are not all optimized at the same level.

It was noted that recent switch to AVX2 by default is slowing down NNVC significantly (about 3 or 4 times slower) compared to AVX512. A better handling of some layers for AVX2 is already prepared and solved this issue.

Other SIMD improvement have been proposed and are currently under review.

Compared to a state-of-the-art CPU inference of the same models (using optimized numpy implementation), a slowdown of a factor 2 to 3 can be observed.

1. **Configurations**

The following configurations is used to generate the different NNVC results.

The column “tested” is read as follow:

* Y: the configuration has been tested using the new NNVC-7.1 software
* P: the results are the ones from previous NNVC software basis
* N: not tested.

The column “xcheck” is read as follow:

* Y: the test has been cross-checked
* P: no cross-checked performed but results are consistent with previous version on NNVC
* N: no cross-check available

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Name | Tools | Configuration | Tested | Xcheck |
| NNVC-7.1 VTM mode | none | encoder\_xxx\_vtm.cfg | P |  |
| NNVC 7.1 Anchor/EE1 | Intra Pred + LOP.2 filterset | encoder\_xxx\_nnvc.cfg | Y | Y |
| NNVC-7.1. HOP | Intra Pred + HOP.2 | encoder\_xxx\_nnvc.cfg + nn-based/HOP.cfg | Y | Y |
| HOP only | HOP | encoder\_xxx\_vtm.cfg + nn-based/HOP.cfg | Y | Y |
| Intra lower complexity | NN Intra lower complexity | nn-based/intra2.cfg | Y | Y |
| set0 | Loop filter set #0 | encoder\_xxx\_vtm.cfg + nn-based/NnlfOption\_1.cfg | P |  |
| set1 | Loop filter set #1 | encoder\_xxx\_vtm.cfg + nn-based/NnlfOption\_2.cfg | P |  |
| set0+rdo | Loop filter set #0 + Rdo | nn-based/NnlfOption\_1.cfg + --EncNnlfOpt=1 | P |  |
| set1+rdo+intra+temporal filter | Loop filter set #1 + Temporal filter | nn-based/NnlfOption\_2.cfg  + --NnlfSet1Multiframe=1 | P |  |
| Sr | Super-resolution | nn-based/nnsr.cfg + nn-based/nnsr\_classAx.cfg | P |  |
| Pf | Adaptive post-filters | nn-based/nnpf/nnpf\_xxx.cfg | P |  |

1. **Recommendations**

The AHG recommends to:

* Continue to develop NNVC software.
* Improve the software documentation.
* Encourage people to report all (potential) bugs that they are finding using GitLab Issues functionality <https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/VVCSoftware_VTM/-/issues>

Encourage people to submit merge requests fixing identified bugs.

[JVET-AG0015](https://jvet-experts.org/doc_end_user/current_document.php?id=13788) JVET AHG report: Gaming content compression (AHG15) [S. Puri, J. Sauer (co-chairs), R. Chernyak, A. Duenas, L. Wang (vice chairs)]

1. **Testing conditions for gaming content**

This section describes simulations’ setup that is planned to be used within AhG15.

The simulations will be performed prior to the January meeting.

Anchor: VTM-11ecm11 (<https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tree/VTM-11.0ecm11.0>)

Test: ECM-11.0 (<https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tree/ECM-11.0>)

Configurations: AI/LDB/LDP

Testing conditions: JVET CTC with QP set [27, 32, 37, 42]

Content: 5s; TBD

Extra parameters: with and without ClassF cfg file

1. **Gaming content sequence survey**

A plenitude of candidate sequences for gaming content compression were identified by the group. The sequences along with their source are listed in the following table: In addition to the rendered view (YUV) for some sequence auxiliary information is available (depth, optical flow, etc.).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sequence name** | **Provided by** | **Resolution, FPS, color space, bit depth** | **Availability** | **Auxiliary information available** |
| Level1 | InterDigital | 1920x1080, 60p, 420, 10bits | JVET ftp | yes |
| Darktree | InterDigital | 1920x1080, 60p, 420, 10bits | JVET ftp | yes |
| ArenaOfValor | Tencent | 1920x1080, 60p, 420, 8bits | JVET ftp | no |
| ARPG | Huawei | 1920x1080, 60p, 10s (600 Frames), YUV420(8 bits) | to be provided | yes |
| DesertTown1 | Huawei | 1920x1080, 60p, 10s (600 Frames), YUV420(8 bits) | to be provided | yes |
| DesertTown2 | Huawei | 1920x1080, 60p, 10s (600 Frames), YUV420(8 bits) | to be provided | yes |
| Sun\_Temple1 | Huawei | 1920x1080, 60p, 10s (600 Frames), YUV420(8 bits) | to be provided | yes |
| Sun\_Temple2 | Huawei | 1920x1080, 60p, 10s (600 Frames), YUV420(8 bits) | to be provided | yes |
| CSGO | Twitch/SA4 | 1920x1080, 60p, 10s (600 Frames), YUV420(8 bits) | [h264](https://media.xiph.org/video/derf/twitch/H264/) [uncompressed](https://media.xiph.org/video/derf/twitch/Uncompressed/) / [ReferenceSequences](https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/) | to be checked |
| DOTA2 | Twitch | 1920x1080, 60p, 10s (600 Frames), YUV420(8 bits) | [h264](https://media.xiph.org/video/derf/twitch/H264/) [uncompressed](https://media.xiph.org/video/derf/twitch/Uncompressed/) | to be checked |
| EuroTruckSimulator2 | Twitch | 1920x1080, 60p, 10s (600 Frames), YUV420(8 bits) | [h264](https://media.xiph.org/video/derf/twitch/H264/) [uncompressed](https://media.xiph.org/video/derf/twitch/Uncompressed/) | to be checked |
| Fallout4 | Twitch | 1920x1080, 60p, 10s (600 Frames), YUV420(8 bits) | [h264](https://media.xiph.org/video/derf/twitch/H264/) [uncompressed](https://media.xiph.org/video/derf/twitch/Uncompressed/) | to be checked |
| GTAV | Twitch | 1920x1080, 60p, 10s (600 Frames), YUV420(8 bits) | [h264](https://media.xiph.org/video/derf/twitch/H264/) [uncompressed](https://media.xiph.org/video/derf/twitch/Uncompressed/) | to be checked |
| Hearthstone | Twitch | 1920x1080, 60p, 10s (600 Frames), YUV420(8 bits) | [h264](https://media.xiph.org/video/derf/twitch/H264/) [uncompressed](https://media.xiph.org/video/derf/twitch/Uncompressed/) | to be checked |
| Minecraft | Twitch/SA4 | 1920x1080, 60p, 10s (600 Frames), YUV420(8 bits) | [h264](https://media.xiph.org/video/derf/twitch/H264/) [uncompressed](https://media.xiph.org/video/derf/twitch/Uncompressed/) / [ReferenceSequences](https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/) | to be checked |
| Rust | Twitch | 1920x1080, 60p, 10s (600 Frames), YUV420(8 bits) | [h264](https://media.xiph.org/video/derf/twitch/H264/) [uncompressed](https://media.xiph.org/video/derf/twitch/Uncompressed/) | to be checked |
| Starcraft | Twitch/SA4 | 1920x1080, 60p, 10s (600 Frames), YUV420(8 bits) | [h264](https://media.xiph.org/video/derf/twitch/H264/) [uncompressed](https://media.xiph.org/video/derf/twitch/Uncompressed/) / [ReferenceSequences](https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/) | to be checked |
| Witcher3 | Twitch | 1920x1080, 60p, 10s (600 Frames), YUV420(8 bits) | [h264](https://media.xiph.org/video/derf/twitch/H264/) [uncompressed](https://media.xiph.org/video/derf/twitch/Uncompressed/) | to be checked |
| Baolei-Man | Tencent/SA4 | 1920x1080, 60p, 10s (600 Frames), YUV420(8 bits) | [ReferenceSequences](https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/) | to be checked |
| Baolei-Balloon 4K | Tencent/SA4 | 4096x2160, 60p, 10s (600 Frames), YUV420(8 bits) | [ReferenceSequences](https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/) | to be checked |
| Baolei-Yard 4K | Tencent/SA4 | 4096x2160, 60p, 10s (600 Frames), YUV420(8 bits) | [ReferenceSequences](https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/) | to be checked |
| Baolei-Woman | Tencent/SA4 | 1920x1080, 60p, 10s (600 Frames), YUV420(8 bits) | [ReferenceSequences](https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/) | to be checked |
| Jianling-Temple | Tencent/SA4 | 1920x1080, 60p, 10s (600 Frames), YUV420(8 bits) | [ReferenceSequences](https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/) | to be checked |
| Jianling-Beach | Tencent/SA4 | 1920x1080, 60p, 10s (600 Frames), YUV420(8 bits) | [ReferenceSequences](https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/) | to be checked |
| Heroes of the Storm part 1 | Kingston/SA4 | 1920x1080, 30p, 30s (900 Frames), YUV420(8 bits) | [ReferenceSequences](https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/) | to be checked |
| Project CARS | Kingston/SA4 | 1920x1080, 30p, 30s (900 Frames), YUV420(8 bits) | [ReferenceSequences](https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/) | to be checked |
| WoW part 2 | Kingston/SA4 | 1920x1080, 30p, 30s (900 Frames), YUV420(8 bits) | [ReferenceSequences](https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/) | to be checked |

1. **Auxiliary information for gaming sequences**

Various additional information may be available from the gaming engine. This can include:

|  |  |  |
| --- | --- | --- |
| **Type** | **Format** | **Used in** |
| depth map | per-pixel, 14 bit | JVET-Y0041 |
| optical flow | quarter-pel | JVET-Y0041 |
| camera information | 4x4 array, 32bit | JVET-AF0187 |
| depth map | per-pixel, 32bit | JVET-AF0187 |
| motion vector | per-pixel, 32bit | JVET-AF0187 |
| camera information | 4x4 array, 32bit | DesertTown/ Sun\_Temple when released |
| depth map | per-pixel, 32bit (R32) | DesertTown/ Sun\_Temple when released |
| motion vector | per-pixel, 2x32bit (R32G32) | DesertTown/ Sun\_Temple when released |
| albedo texture | per-pixel, 4x8bit (R8G8B8A8) | DesertTown/ Sun\_Temple when released |
| normal map | per-pixel, 3x32bit (R32G32B32) | DesertTown/ Sun\_Temple when released |

1. **Sequence segments to be simulated**

Of each sequence a segment of 5 seconds was selected for encoding for January meeting. This data should help identifying the most interesting sequences for a new class of gaming content sequences.

The table below lists the selected segments for each sequence. Note that for some sequence more than one potentially interesting segment was identified.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sequence name** | **Frame count** | **Frame rate** | **Bit depth** | **Start frame** | **End frame** |
| Level1 | 600 | 60 | 10 | 0 | 299 |
| Darktree | 600 | 60 | 10 | 300 | 599 |
| ArenaOfValor | 600 | 60 | 8 | 0 | 299 |
| ARPG | 600 | 60 | 8 | 100 | 399 |
| DesertTown1 | 600 | 60 | 8 | 200 | 499 |
| DesertTown2 | 600 | 60 | 8 | 150 | 449 |
| Sun\_Temple1 | 600 | 60 | 8 | 40 | 339 |
| Sun\_Temple2 | 600 | 60 | 8 | 50 | 349 |
| CSGO | 3600 | 60 | 8 | 1100 | 1399 |
| DOTA2 | 3600 | 60 | 8 | 1734 | 2033 |
| EuroTruckSimulator2 | 3600 | 60 | 8 | 1000 | 1299 |
| Fallout4 | 3600 | 60 | 8 | 600 | 899 |
| GTAV | 3600 | 60 | 8 | 300 | 599 |
| Hearthstone | 3600 | 60 | 8 | 3099 | 3398 |
| Minecraft | 3600 | 60 | 8 | 600 | 899 |
| Rust | 3600 | 60 | 8 | 3000 | 3299 |
| Starcraft | 3600 | 60 | 8 | 2900 | 3199 |
| Witcher3 | 3600 | 60 | 8 | 1300 | 1599 |
| Baolei-Man | 600 | 60 | 8 | 300 | 599 |
| Baolei-Balloon 4K | 600 | 60 | 8 | 300 | 599 |
| Baolei-Yard 4K | 600 | 60 | 8 | 60 | 359 |
| Baolei-Woman | 600 | 60 | 8 | 120 | 419 |
| Jianling-Temple | 600 | 60 | 8 | 0 | 299 |
| Jianling-Beach | 600 | 60 | 8 | 0 | 299 |
| Heroes of the Storm part 1 | 300 | 30 | 8 | 150 | 299 |
| Project CARS | 300 | 30 | 8 | 0 | 149 |
| WoW part 2 | 300 | 30 | 8 | 150 | 299 |

1. **Coding performance results**

Encoding ongoing. Will be added when available.

1. **Recommendations**

The AHG recommends to:

* Continue to evaluate candidate sequences for gaming content compression
* Continue to refine testing conditions (use classF yes/no, potentially adjust QPs)
* Identify if/how auxiliary information can be used for coding of gaming content
* Review input documents on gaming content compression

Further discussion on test material in joint meeting with AG 5. After availability of results, a BoG meeting (J. Sauer, S. Puri) should summarize them before the joint meeting is held.

It should be clarified if it is possible to transfer the newly proposed content to the JVET ftp (licensing conditions are included in the telco report).

[JVET-AG0016](https://jvet-experts.org/doc_end_user/current_document.php?id=13789) JVET AHG report: Generative face video compression (AHG16) [Y. Ye (chair), H.-B. Teo, Z. Lyu, S. McCarthy, S. Wang (vice chairs)]

1. **Conference calls**

AHG16 conducted two conference calls during this AHG period, with 42 and 32 participants in the first and second conference call, respectively. The following three contributions were reviewed at these conference calls:

* JVET-AG0042 “AHG 16: Proposed Common Software Tools and Testing Conditions for Generative Face Video Compression” is related to the mandates on testing conditions and software tools for GFVC experimentation; it proposes talking face video test sequences, coding configurations for VVC and GFVC, quality and distortion metrics, and proposes a software tool to evaluate and compare the GFVC coding performance with that of VVC Main 10 profile;
* JVET-AG0048 “AHG16: Interoperability Study on Parameter Translator of Generative Face Video Coding” is related to the mandate on interoperability study; it trains parameter translator networks that allow “mismatched” encoder and decoder (i.e. encoder and decoder can rely on different types of GFVC algorithm and parameters) to interoperate. The software implementation of JVET-AG0048 is based on the GFVC software tool provided by JVET-AG0042;
* JVET-AG0187 “AHG16: Study text for common test conditions and evaluation procedures for generative face video coding (draft 1)” is related to the mandate on testing conditions. It provides a template for drafting the GFVC testing conditions and reference configurations.

JVET-AG0043 “AHG16: Report on AhG meeting on generative face video compression” provides detailed notes on the discussions around the above contributions and the participant lists of the two conference calls. A brief summary of the main conclusions of the AHG16 conference calls is as follows:

* AHG16 recommends the following initial testing conditions to be used for GFVC:
  + 8-bit, SDR, low delay configurations;
  + 15 VOXCELEB test sequences and 18 CFVQA test sequences as identified in JVET-AG0042, where the 18 CFVQA sequences also include the 3 extra “bad case” sequences provided in JVET-AG0042-v3, all sequences are in 256x256 resolution;
  + VVC QP: 37, 42, 47, 52;
  + GFVC QP (for base pictures): 22, 32, 42, 52;
* Software tool: recommend to adopt the GFVC software tool proposed in JVET-AG0042 as the AHG16 software tool;
* Interoperability: recommend to adopt the translator networks proposed in JVET-AG0048 as part of the AHG16 software tool to provide an interoperability example;
* Encourage identification and study of other datasets, including test sequences with higher resolutions;
* Continue to study the behavior of GFVC within a wider bit rate range, not just the QP points suggested above.

AHG16 requests for JVET to discuss the establishment of the AHG16 software tool and testing conditions considering the above recommendations. AHG16 also requests GitHub repository to host AHG16 software tool and disk space to host the test sequences on JVET’s ftp site.

1. **Related contributions**

Besides the three contributions listed above (JVET-AG0042, JVET-AG0048, and JVET-AG0187), the following contributions are identified as being related to the activities of AHG16:

[JVET-AG0087](https://jvet-experts.org/doc_end_user/current_document.php?id=13643) AHG9: On the generative face video SEI message [M. M. Hannuksela, F. Cricri, H. Zhang (Nokia)]

[JVET-AG008](https://jvet-experts.org/doc_end_user/current_document.php?id=13644)8 AHG9: Usage of the neural-network post-filter characteristics SEI message to define the generator NN of the generative face video SEI message [M. M. Hannuksela, F. Cricri, H. Zhang (Nokia)]

[[JVET-A](https://jvet-experts.org/doc_end_user/current_document.php?id=13695)G0139](https://jvet-experts.org/doc_end_user/current_document.php?id=13595) AHG16: Depthwise separable convolution for generative face video compression [R. Zou, R.-L. Liao, B. Chen, J. Chen, Y. Ye (Alibaba)]

[JVET-AG0203](https://jvet-experts.org/doc_end_user/current_document.php?id=13759) AHG9/AHG16: Common text for proposed generative face video SEI message [J. Chen, B. Chen, Y. Ye (Alibaba), S. Yin, S. Wang (CityU), S. McCarthy, P. Yin, G.-M. Su, A. K. Choudhury, W. Husak, G. J. Sullivan (Dolby)]

1. **Recommendations**

The AHG recommends to:

* Discuss the establishment of the AHG16 software tool and testing conditions in JVET.
* Review related contributions.

It was commented that it would be beneficial to have quality metric numbers reported (could however only be done once the test conditions are agreed).

# Project development (28)

## AHG1: Deployment and advertisement of standards (2)

Contributions in this area were discussed at 0055–0120 on Thursday 25 Jan. 2024 (chaired by Y. Ye).

[JVET-AG0020](https://jvet-experts.org/doc_end_user/current_document.php?id=13605) Deployment status of the HEVC standard [G. J. Sullivan]

This information contribution contains a survey of deployed products and services using the HEVC standard and the formal specifications in which it is supported, along with a brief introduction to the standard written for broad readership. Revision marking is included to show changes relative to JVET-AF0020-v1 of October 2023.

A number of updates were reviewed, two significant ones being the following:

* In December 2023, Futuresource Consulting issued a report about HEVC commissioned by InterDigital, reporting:
  1. That “Over 96% of consumer video products sold today have HEVC playback capability, totalling 2.3 billion devices annually; and this is rising year over year”.
  2. Installed-base penetration of HEVC decoding was estimated at 90.4%, totalling 8.1 billion devices and up from 58% four years earlier, including 4.1 billion smartphones (93% explicitly identified as hardware support), 1.2 billion smart TVs, 733 million set-top boxes, 254 million media streaming devices, 132 million games consoles (40%), 513 million tablets (a majority in a declining segment), and 1.2 billion PCs and laptops.
  3. Sales within 2023 estimated at 417 million HEVC-enabled TVs, set-top boxes and media streaming devices.
  4. A current usage share of 17% for HEVC (about double the percentage share from four years earlier).
  5. Common usage of HEVC for 77% (10 of 13) identified streaming service providers for movies and episodic video content (Amazon Freeverse, Amazon Prime, Apple TV+, Discovery Go, Disney+, HBO Max, Hulu, Netflix, Paramount+, and Peacock), with higher usage of HEVC for the more rapidly growing UHD services than for HD.
* Stereoscopic multi-view HEVC (MV-HEVC) emerged prominently in 2023.
  1. In June 2023, Apple released a beta version of an Apple HEVC Stereo Video specification for stereoscopic applications including Apple Vision Pro headsets for release in early 2024, using 8-bit or 10-bit multi-view HEVC (a.k.a. MV-HEVC). Support for MV-HEVC capture and encoding for the iPhone 15 Pro phone and iOS 17.2 was announced in November 2023 and released in December 2023.
  2. Support for MV-HEVC was announced by Alteon.io in November 2023.
  3. Disney announced support of MV-HEVC for Apple Vision Pro headsets with its Disney+ streaming service in January 2024, including HDR with UHD resolution for each eye.

Comments and discussions:

* There are multiple codec choices in the marketplace nowadays.
* Futuresource Consulting’s survey on the service providers seems to be focused on the North American market.
* Ateme recently added support for MV-HEVC encoding

[JVET-AG0021](https://jvet-experts.org/doc_end_user/current_document.php?id=13606) Deployment status of the VVC standard [G. J. Sullivan]

This contribution contains a survey of deployed products and services, publicly available software source code, related tools, and formal specifications supporting the VVC standard (Rec. ITU-T H.266 | ISO/IEC 23090-3).

Revision marking is included to show changes relative to JVET-AF0021-v3 of October 2023.

A number of updates (on SEI Robotics and MainConcept) were reviewed.

## AHG2: Text development and errata reporting (4)

Contributions in this area were discussed (also on Wed for AG0327) at 0900–0925 on Thursday 25 Jan. 2024 (chaired by JRO).

[JVET-AG0204](https://jvet-experts.org/doc_end_user/current_document.php?id=13760) AHG9: On VSEI version 4 [Hendry (LGE), S. McCarthy (Dolby), J. Chen, Y. Ye (Alibaba)]

The current version of the TuC for future extensions of VSEI contains aspects related to VVC subclauses, VUI syntax and semantics, and many SEI messages. Some of these aspects, particularly some SEI messages, might be mature enough for inclusion in a next version of the VSEI specification. It is proposed to start work at this meeting on a preliminary working draft or working draft for the next version of VSEI to support technologies in the TuC that are now determined to be mature. It is also proposed to continue the TuC to facilitate study of additional technologies.

[JVET-AG0327](https://jvet-experts.org/doc_end_user/current_document.php?id=13900) AHG9: List of technologies under consideration for future extensions of VSEI [S. McCarthy (Dolby)] [late]

This contribution provides the following table:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Spec** | **Type** | **Group** | **Technology** | **Version 1  (JVET-AE2032)** | **Version 2 (JVET-AF2032)** | **Potential WD candidate** |
| VSEI | VUI |  | modality type |  | y | (x) |
| VSEI | SEI | NNPF SEI messages | NNPFC SEI message application information signalling | y | y | x |
| general post-processing filter process and neural-network post-filter using PPFs | y | y |  |
| neural-network post-filter group characteristics SEI message | y | y |  |
| neural-network post-filter group activation SEI message | y | y |  |
| post-processing filter gain SEI message |  | y |  |
| VSEI | SEI |  | encoder optimization SEI message | y | y | x |
| VSEI | SEI |  | source picture timing information SEI message | y | y | x |
| VSEI | SEI |  | object mask information SEI message | y | y | (x) |
| VSEI | SEI | Image format metadata SEI messages | EXIF metadata SEI message |  | y |  |
| JFIF metadata SEI message |  | y |  |
| XMP metadata SEI message |  | y |  |
| VSEI | SEI |  | multiplane image information SEI message |  | y |  |
| ?? |  |  | large supplemental enhancement information message |  | y |  |

The initial selection of potentially mature candidates was discussed and agreed Wed. 24 Jan. 0845-0855. (x) needs more review or clarification. TuC adoptions of the current meeting should be added if mature, e.g. temporal extrapolation purpose for NNPFC. Revisit.

Prelim. WD Jan. 2024

CD April 2024 + request

DIS July 2024

FDIS + ITU consent Jan 2025

It was noted that errata item are also to be included.

[JVET-AG0213](https://jvet-experts.org/doc_end_user/current_document.php?id=13769) AHG1/AHG2/AHG8: On project management related to the encoder optimization information SEI message [M. M. Hannuksela, A. Aminlou, F. Cricri, H. Zhang (Nokia)]

See section 6.5

[JVET-AG0308](https://jvet-experts.org/doc_end_user/current_document.php?id=13881) CICP extensions for depth and alpha signaling [D. Podborski, A. Tourapis (Apple)] [late]

This contribution extends the colour primaries and matrix coefficients code points in CICP to also allow the support of depth and alpha information. The intent of this extension is to enable carriage of such information in single layer bitstreams using existing video coding specifications, such as AVC and HEVC, without the need of specifying any additional metadata.

This is suggested as a simple approach to signal that only an alpha map or depth map is carried in a monochrome sequence. For the depth map, usually the SEI expressing the translation of the values into depth values would be needed. It needs to be clarified if the existing SEI message for depth maps can be used in single-layer streams in AVC and HEVC, and how it could be associated with the depth map, or if an extension would be necessary. It is agreed that this is a desirable functionality (to be supported with AVC, HEVC and VVC), but further study appears necessary to understand what is the best way (e.g., JVET-AG0144 has the same target) of achieving it, and which changes would be necessary for the different specs.

## AHG3: Test conditions (0)

This section is kept as a template for future use.

## AHG3: Software development (0)

This section is kept as a template for future use.

## AHG4: Subjective quality testing and verification testing (2)

Contributions in this area were discussed in joint meeting (see section 7.4).

[JVET-AG0071](https://jvet-experts.org/doc_end_user/current_document.php?id=13627) Informal Subjective Evaluation of Low Complexity Enhancement Video Codec (LCEVC) with VVC on SDR UHD (4K) Content [O. Chubach, H.-H. Chen, C.-Y. Chen, T.-D. Chuang, Y.-W. Chen, C.-W. Hsu, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-AG0224](https://jvet-experts.org/doc_end_user/current_document.php?id=13797) Response to JVET-AG0071 on Informal Subjective Evaluation of Low Complexity Enhancement Video Codec (LCEVC) with VVC on SDR UHD (4K) Content [L. Ciccarelli, S. Ferrara (V-Nova)] [late]

## AHG4: Test material (1)

Contributions in this area were discussed at 2115–2135 on Tuesday 23 Jan. 2024 (chaired by JRO and MW).

[JVET-AG0228](https://jvet-experts.org/doc_end_user/current_document.php?id=13801) AHG4/AHG13: source for scanned film test sequences [P. de Lagrange (InterDigital)] [late]

This contribution reports about old film scanned by Prelinger Archives and made available on the Internet Archive website, and recommends to select a few samples to be included as candidate test sequences for FGC SEI visual testing activities planned in JVET-AD2022.

It was reported by S. Wenger that a good selection of materials from Internet archive is tagged as public domain, so it could be used. This would however not be the case for all material, would need to be checked on a case by case basis.High quality versions of sequences are also available, though not uncompressed.

It was suggested to take action on selecting some sequences and investigate their compression characteristics. If not possible during the meeting, it should be added to AHG mandates.

## Codec performance with alternative test materials (0)

This section is kept as a template for future use.

## AHG5: Conformance test development (1)

Contributions in this area were discussed at 2100–2115 on Thursday 25 Jan. 2024 (chaired by JRO).

[JVET-AG0329](https://jvet-experts.org/doc_end_user/current_document.php?id=13902) On update of Multilayer conformance test for multiple profiles signalling [S. Iwamura, S. Nemoto, A. Ichigaya (NHK), F. Urban, P. de Lagrange (InterDigital)] [late]

Additional conformance bitstream for multilayer coding configurations is defined in JVET-AE2028. During the crosscheck of the conformance bitstreams, it was found that the bitstreams are not well formed, especially the VPS does not contain intended data. This contribution proposes to update those conformance test with revised configuration and also proposes VTM encoder only changes to solve the issues identified during the regeneration of the conformance bitstream.

No need to change the bitstream description in JVET-AE2028. Bitstreams to be included in repository.

Software merge request was submitted already.

## AHG7: ECM tool assessment (1)

Contributions in this area were discussed at 2150–2240 on Tuesday 23 Jan. 2024 (chaired by JRO and MW).

[JVET-AG0173](https://jvet-experts.org/doc_end_user/current_document.php?id=13729) AHG7: ECM-11 results of non-CTC sequences [X. Li (Google)]

In JVET-AF meeting, one of AHG7 mandate was modified to “Collect simulation results on non-CTC sequences (e.g., those used in previous verification tests), and identify a set of non-CTC sequences that would be appropriate for additional testing”. This proposal reports the ECM-11 results of the sequences used in the subjective tests in JVET-AB2029 and the sequences with mixed content.

Firstly, the four UHD sequences (DrivingPOV3, Marathon2, MountainBay2, TallBuildings2) and four HD sequences (Beatriz, EuroTruckSimulator2, DOTA2, Meridian2) used in JVET-AB2029 visual quality comparison of ECM/VTM encoding are tested. Note that seven UHD sequences and six HD sequences were tested in JVET-AB2029. But three UHD and two HD sequences were also used in the ECM CTC. The duplicated sequences were excluded in this test.

Secondly, the three mixed content sequences defined in JVET-AC1015 Common test conditions for SCM screen content coding were also tested as there are no mixed content sequences in ECM CTC.

The following observations were made:

1. The average UHD and HD luma coding gains of non-CTC sequences are about 3-4% lower than those of the CTC sequences (20.5% vs 24.2% for UHD and 16.5% vs 20.8% for HD) while the average coding time ratios of non-CTC sequences are larger than those of CTC sequences
2. The average luma gain over group 1-4 off with non-CTC sequences is smaller than that with CTC sequences.
3. The average luma coding gain of the mixed content is between the average coding gain of class F and class TGM in CTC
4. The average coding gain of VMAF (and VMAF NEG) is generally in line with the average coding gain of luma PSNR
5. The encoding time ECM-11 is much slower than that of VTM-11ecm11, which will make bitrate matching of subjective test more difficult

it is proposed to

* Use the four UHD and four HD non-CTC sequences for the potential subjective tests
* Also test group 1-4 tool off in potential subjective tests
* Replace some class F and TGM sequences with mixed content sequences if it is not preferred to test more sequences in ECM CTC

It was commented that the assumption that overtuning has happened in ECM encoding may be questionable. Further investigation and more detailed analysis on this would be necessary (e.g., by tool-on/off tests). It is however noted that the benefit of such an investigation would be questionable. Naturally, when using same test sequences for a longer period of time (allowing tracking of progress), some degree of overfitting may happen. This is why we are conduction verification test with different material to show that the development has nevertheless fulfilled its purpose. Likewise, when a CfP is made, typically some new test material should be added beyond/replacing the sequences that had been used in exploration.

A comment was made that variations in terms of compression when using different sequences is often observed and typical, depending on characteristics.

Concern was expressed about the group-based analysis, turning off tools from 4 groups (1-4).

Having a second configuration with disabling template matching based tools might give some interesting information, to learn what they contribute in terms of visual quality (if manageable in terms of amount of testing).

BoG (M. Wien) to develop a plan for the subjective testing to be conducted in April.

## AHG8: Optimization of encoders and receiving systems for machine analysis of coded video content (7+1)

Contributions in this area were discussed at 2240–2300 and 2320-0135+1 on Tuesday 23 Jan. 2024, at 2230-2310 on Wednesday 24 Jan. 2024, and at 0540-0600 on Thursday 25 Jan. 2024 (chaired by JRO).

[JVET-AG0085](https://jvet-experts.org/doc_end_user/current_document.php?id=13641) [AHG8] Continuation of study on different VTM versions [C. Hollmann (Ericsson)]

In this contribution the study originally presented in JVET-AE0096 is reported to be continued. Results are made available for VTM-21, VTM-22, and VTM-23. The performance of VTM-21 and VTM-22 is reported to be identical to VTM-20, whereas VTM-23 provides an additional benefit in machine vision task performance compared to VTM-20 for random access and low delay.

Results for ECM-11 are also reported. On SFU-HW, using ECM-11 is reported to give a BD-rate result of around −14% for random access and −10% for both low delay and all intra configurations. On TVD, ECM‑11 is reported to give a BD-rate result of around −15% for random access, −7% for low delay, and −16% for all intra configurations.

In version 2 of this contribution results for ECM-11 for TVD in AI configuration were added.

Results are indicating unexpectedly large variations in performance over the different VTM version, both on average, and even more on a sequence-by-sequence basis.

Compared to JVET-AE0096, results with newer VTM versions, and with ECM are added. For the machine task judged by BD rate with MAP and MOTA criteria, gain is observed relative to VTM, but less than for a PSNR criterion.

[JVET-AG0090](https://jvet-experts.org/doc_end_user/current_document.php?id=13646) [AHG8] Comments and editorial changes to the draft TR on optimizations for encoders and receiving systems for machine analysis of coded video content [C. Hollmann (Ericsson), S. Liu (Tencent), J. Chen (Alibaba)]

In this contribution editorial changes to the draft technical report on optimizations for encoders and receiving systems for machine analysis of coded video content are proposed. Furthermore, the progression of the technical report is discussed and it is proposed to issue a request for subdivision at this meeting for a new part in ISO/IEC 23888 (Artificial Intelligence for Multimedia) with a 24-month timeline.

The following timeline is proposed for the development of the TR:

* Request for subdivision: January 2024
* Committee draft (CD): April 2024
* CD ballot: May – June 2024
* Draft TR: November 2024
* DTR ballot: December 2024 – March 2025
* Finalization: April 2025

Regarding the request for subdivision, it is proposed to request a new part in ISO/IEC 23888 (Artificial Intelligence for Multimedia) with a timeline of 24 months.

Check if references to VVC and VSEI need to be added (e.g., for some definitions on terminology).

Reference to software may need revision.

Output from this meeting: Request for subdivision, WD (not preliminary any more).

It was commented that the current version is mainly for object detection and object tracking. Later extensions may also refer to other tasks.

In case that more relevant contributions would be received by next meeting, the timeline could still be modified.

[JVET-AG0178](https://jvet-experts.org/doc_end_user/current_document.php?id=13734) AHG8: Truncating bit depth in video coding for machine tasks [D. Ding, X. Zhao, S. Liu]

Currently, video content is usually represented with a bit depth of eight or higher, such as 8-bit or 10-bit, widely used for video communication tailored to human vision. However, for machine vision tasks like object tracking and detection, the quality score, for instance, measured by detection rate, exhibits less sensitivity to bit depth. Hence, a lower sample value range, represented by fewer bits, can better balance bitrate and machine task score.

This work introduces a method based on truncating bit depth to enhance RD performance in line with common test conditions of optimization of encoders and receiving systems for machine analysis of coded video content. Specifically, on the encoder side, the luma sample values are right shifted by 1 bit, while the chroma sample values remain unchanged. On the decoder side, the reconstructed luma samples are conditionally shifted back (left shift) accordingly.

BD rate coding gain (Pareto MOTA) on the AI, LD, RA configurations are respectively -33.43%, -25.54%, -31.70% in TVD dataset. The encoding time on AI, LD, RA configurations are 48.45%, 67.64%, 64.20% of the anchor.

It was commented that for the example of the basketball drill sequence, at the highest rate point the performance is less than for the 10-bit anchor. It was asked if there is a saturation effect? The proponents believe that this would not be the case, but it was not tested, as the QP values of the test points are defined in CTC.

The cross-checker reports that similar behaviour was observed for other variants of basketball sequences. He believes that this may partially also be due to the ground truth defined for those, also the performance is relatively low for those (25% MAP at highest rate point).

Also traffic shows similar behaviour. At the lowest rate point, MAP goes down to 5%, which appears not very useful.

It was suggested that it would be better to make the truncation sequence dependent. However, finding a criterion to decide whether the bit depth should be truncated or not might not be straightforward (the encoder should not directly use knowledge about the exact machine task that is operated, according to testing conditions.

It was suggested that inserting zero bits in the left shift might not be optimum. For example, when converting 8 bit to 10 bit, inserting “10” would be better (average of values that might have been truncated).

It was commented that for proposals like this which intend to save bit rate, it would be relevant that the range of quality of the classification retained. Therefore, it would be more useful to compare RD curves at approximately same MAP, rather than same QP.

In this context, it was suggested that it would be logical to use lower QP for the lower bit depth, as it brings the rates to similar amount.

Further study to make the results more consistent.

[JVET-AG0242](https://jvet-experts.org/doc_end_user/current_document.php?id=13815) Cross-check of JVET-AG0178 (AHG8: Truncating bit depth in video coding for machine tasks) [C. Hollmann (Ericsson)] [late]

[JVET-AG0209](https://jvet-experts.org/doc_end_user/current_document.php?id=13765) AHG8: A suggestion for the performance evaluation of VCM [S. Wang, J. Chen, Y. Ye (Alibaba), S. Wang (CityU)]

In current common test conditions (CTC) of VCM, for each machine task, only a single machine vision model is used to evaluate the VCM performance. However, even for a specific machine task, there are a variety of machine vision models in the real-world. It is asserted that evaluating VCM performance with only one model may not be stable according to the experimental results in this contribution. And thus, it is proposed to consider including more machine vision models for each machine task in VCM CTC.

Using more than one model for the machine task is definitely desirable, and it is obvious that the model performance has high impact on the evaluation criteria. However, regarding RD behaviour, it may be natural that graphs are crossing when one of the models is more affected by compression artifacts. Therefore, there is a high probability that results are “inconsistent” according to the analysis of the proposal. May be better to analyse the results from different models in a more general way, e.g. if they show a consistent tendency.

A good approach would be to look for models working well on uncompressed data, and then investigate how they react on compression artifacts.

Add mandate to AHG

Needs to be discussed with WG 4 to keep testing conditions aligned.

It was commented that test conditions were modified by WG 4 in the meantime.

[JVET-AG0212](https://jvet-experts.org/doc_end_user/current_document.php?id=13768) AHG8: A post-processing algorithm for machine consumption [B. Li, S. Wang, J. Chen, Y. Ye (Alibaba), S. Wang (CityU)]

A learning-based video post-processing algorithm in which an enhancement neural network is utilized to process the reconstructed pictures to improve machine task performance is proposed in this contribution. According to the experimental results, it is reported that the proposed algorithm achieves 18.46% (RA), 26.35% (LD), 33.70% (AI) BD-rate saving for object detection task on SFU dataset, and 27.86% (RA), 42.96% (LD), 33.70% (AI) BD-rate savings for object tracking task on TVD dataset. Additional simulation results show that the proposed algorithm can also achieve obvious performance improvement when evaluated with a different machine task model rather than the one specified in VCM CTC.

It was asked if tests were performed with uncompressed data? No, but the effect of the filter is adapted depending on QP, and at high rates performance is almost identical. Also, improvement in classification quality is mainly at low bit rates.

Decoding time is significantly increased (approx.10x), run on GPU

Results confirmed by crosscheck – one remark is the need for large GPU memory

Pure postprocessing, no information on that in the bitstream.

Model was trained with CTC classification model, but better result on alternative classification model. Explanation for that? Might it be the case that the alternative model has better performance on uncompressed data?

Would it also give benefit for human viewing? PSNR results?

Proponents are requested to provide additional results on uncompressed performance and PSNR.

Postfilter is an interesting element of completely non-normative processing that should be included in the TR. Not clear if this is the best representative approach, due to high complexity. It was agreed to include software in AHG8 software branch for further study by other experts.

[JVET-AG0314](https://jvet-experts.org/doc_end_user/current_document.php?id=13887) Crosscheck of AG0212 (AHG8: A post-processing algorithm for machine consumption) [H. Zhang (Nokia)] [late]

[JVET-AG0213](https://jvet-experts.org/doc_end_user/current_document.php?id=13769) AHG1/AHG2/AHG8: On project management related to the encoder optimization information SEI message [M. M. Hannuksela, A. Aminlou, F. Cricri, H. Zhang (Nokia)]

Other contributions on related SEI messages and review of this one – see section 6.6

In a follow-up discussion, it was agreed that a reference to this SEI message (text included in JVET-AG0213) shall be included in the TR draft, provided that EOI SEI would also be included in the WD for VSEIv4.

[JVET-AG0216](https://jvet-experts.org/doc_end_user/current_document.php?id=13772) AHG8: Multi-layer VVC for hybrid machine-human consumption [J. Laitinen, T. Partanen, A. Mercat, J. Vanne (Tampere University), A. Aminlou, M. M. Hannuksela, F. Cricri, H. Zhang (Nokia)] [late]

This contribution investigates VVC multi-layer coding to facilitate hybrid machine-human consumption, where machine-optimized input is coded as the base-layer and the original content is encoded as an enhancement-layer for human consumption. Machine-optimized input is generated using pre-processing by extracting *Region-of-Interest* (*ROI*) areas from the background. Two multi-layered schemes are tested by blurring or greying-out the background. For each multi-layered schemes four different spatial scaling ratios of 1x, 0.75x, 0.5x, and 0.25x are tested and compared to coding of original content using VVC (VVC only), and simulcasting of machine-optimized and original contents using VVC (simulcasting). The comparison for machine and human consumption are performed using mAP and PSNR, respectively.

In addition to BD-rate comparisons between the tested schemes, the contribution presents break-even points, defined as the fraction of time that the full multi-layer bitstream is required for human consumption, where the multi-layered schemes are beneficial over the VVC-only case. Based on the presented results, it is concluded that both the VVC multi-layer coding and simulcast are valid options for hybrid machine-human consumption. The contribution proposes to add a section on hybrid machine-human consumption in the technical report on optimization of encoders and receiving systems for machine analysis of coded video content.

In v2, the following items have been added:

* A detailed analysis of 0.25x downsampling of machine consumption using extended QP
* The proposed text to be included in the technical report
* Excel files of results
* Presentation slides

A potential application is security monitoring, where the machine (getting base layer) is analysing the video all the time, and the human only sporadically.Would not be beneficial for storage, where typically the single layer VVC at quality for human consumption would be stored. For streaming, a benefit is claimed when the human does not watch all the time. Potentially, instead using scalable coding, a switch to a higher-quality single layer stream could be made.

Another expert commented that in terms of coding efficiency there could also be cases where using the human stream as base layer is better.

More evidence about the benefit would be necessary for inclusion in the technical report. There are many side assumptions to demonstrate that bit rate is saved.

Further study.

[JVET-AG0217](https://jvet-experts.org/doc_end_user/current_document.php?id=13773) AHG8: Reduced residual encoding in VVC for machine consumption [A. Alireza, A. Hallapuro, H. Zhang [Nokia] [late]

This contribution proposes to scale down residual values of inter blocks at the VTM encoder for machine consumption. The value of the scaling function is set to unity at the block boundary, while its value decreases toward center of the block to 0.7. This scaling guarantees continuity of the reconstructed frame at boundaries of coding units, while encoding the residual of the coding unit potentially with a smaller number of bits. DC component of the residual block is deducted first, scaling is applied to the deducted residual, and then DC component is added back to the scaled residual. The proposed method is not applied to small coding units (i.e., width or height less than or equal to 4/2 for RA/LD). Distortion calculation and rate-distortion optimization is performed according to the existing process in VTM encoder. The contribution proposes to add this method to the technical report on optimization of encoders and receiving systems for machine analysis of coded video content. Average pareto BD-Rate results are as follows:

Object detection (SFU-HW): RA= -0.56%, LD = -2.84%

Object tracking (TVD): RA= -4.02%, LD = -2.41%

In v2, the following items have been added:

* Improvement of the text and figures
* Complete results (SFU and TVD)
* The proposed text to be included in the technical report (section 4)
* Excel file of results

Tests are finalized and cross-checked.It was commented that gains are rather small (e.g., compared to QP adaptation which is another way of encoder adaptation).

Proponents believe that the method can be combined with the QP adaptation approach.

One cross-checker mentions that for some sequences a relatively high loss occurs.

Unlike the QP adaptation that is already included in the TR, the proposed method does not need object detection, so it is content agnostic and task agnostic. It was commented that the proposed method could also get more gain when the decision was based on identifying objects.

It was requested to test the method on top of the QP adaptation approach (in an object agnostic way, or if deemed useful also taking into account different residual scaling in case of objects), to see if the method provides additional benefit.

[JVET-AG0317](https://jvet-experts.org/doc_end_user/current_document.php?id=13890) Crosscheck of JVET-AG0217 (AHG8: Reduced residual encoding in VVC for machine consumption) [C. Hollmann (Ericsson)] late]

[JVET-AG0326](https://jvet-experts.org/doc_end_user/current_document.php?id=13899) Crosscheck of JVET-AG0217 (AHG8: Reduced residual encoding in VVC for machine consumption) [D. Ding (Tencent)] [late]

## AHG10: Encoding algorithm optimization (1)

Contributions in this area were discussed at 2115–2135 on Thursday 25 Jan. 2024 (chaired by JRO).

[JVET-AG0055](https://jvet-experts.org/doc_end_user/current_document.php?id=13611) AHG10: CTU-Level Lagrange Multiplier and QP Adaptation for VVC Low-Delay Configuration [H. Guo, C. Zhu, L. Luo, J. Chen (UESTC), Y. Huo, Y. Liu (Transsion)]

This contribution proposes an updated version of the method in contribution JVET-AF0089.

This contribution proposes a CTU-level Lagrange multiplier and QP adaptation for VVC low-delay configuration. The proposed method was implemented on VTM-23.0. Compared to the original VTM-23.0, the reported overall PSNR-Y, U, V BD-rate and {EncT, DecT} variations in low-delay B and low-delay P are shown below, respectively.

Over VTM-23.0:

* LDB: -3.59%, -4.72%, -3.78% {103%, 100%}
* LDP: -3.44%, -4.47%, -3.05% {102%, 101%}

A bug was found recently, modified results in newest version of the contribution (from which the abstract above was taken). Upon question, it was rather a parameter tuning.

Useful clarification of difference compared to BIM (block importance mapping)

It was asked why for some sequences losses are observed. Proponents believe this was mainly for screen content, but it was commented that it also appears for some cases with camera material.

Interesting tool, useful progress, but it was suggested to further investigate about the non-homogeneous behaviour over different sequences.

It was suggested to make the code available to AHG10, but this would need to be checked company internally.

It was commented that it should be avoided to make parameter tuning specifically for the test set.

It would be highly appreciated if also RA would have benefit.

Further investigation was recommended.

[JVET-AG0280](https://jvet-experts.org/doc_end_user/current_document.php?id=13853) Crosscheck of JVET-AG0055 (AhG10: CTU-Level Lagrange Multiplier and QP Adaptation for VVC Low-Delay Configuration) [J. Huo, Z. Zhang (Xidian Univ.)] [late]

## AHG13: Film grain synthesis (5+5)

Contributions in this area were discussed at 0500–0705 on Tuesday 23 Jan. 2024 and at 2115–2230 on Tuesday 23 Jan. 2024 (chaired by JRO).

[JVET-AG0101](https://jvet-experts.org/doc_end_user/current_document.php?id=13657) AHG9: Film grain adaptive SEI message [Y. Gao, S.-W. Xie, Y.-X. Bai, M.-H. Jia, C. Huang, P. Wu (ZTE)]

It is proposed to define a new film grain adaptive SEI message to specify the regions that the film grain applies, wherein:

* these regions may include the external regions or/and the internal regions.
* external regions refer to the existing available regions defined through different type of SEI messages or other approach.
* internal regions refer to the regions defined in the film grain adaptive SEI message.

What is the benefit of the “external” signalling? As it is rectangular, the signalling overhead of “internal” is not large, and the annotated regions SEI includes many elements that are not at all needed.

It was commented that the establishment of such a combination is inrtroducing an undesirable relationship.

[JVET-AG0140](https://jvet-experts.org/doc_end_user/current_document.php?id=13696) AHG9/AHG13: FGS Extension SEI message for spatial adaptation [G. Teniou, S. Wenger, A. Hinds (Tencent)]

The proposed contribution is a follow-up of previous meeting inputs on FGS spatial adaptation. Taking into account the comments made on the importance to not modify the existing FGC SEI message, this proposal follows the suggested approach to develop a dedicated FGS Extension SEI message dealing with relative spatial application of film grain values in the picture.

Has replicated syntax and semantics of existing SEI, extending it by region information (one or several bounding boxes, max. number of regions limited to 16 by syntax element). It also includes an approach to use alpha map for the overlay of the film grain.

[JVET-AG0153](https://jvet-experts.org/doc_end_user/current_document.php?id=13709) [AHG13] Proposed FGS applications [G. Teniou (Tencent)]

Proposal to add a new use case of FGS: Artistic intent simulation, in which FGS is used without any reference, to be differentiated from the artistic intent recovery that relies on FGS to maintain the similarity with the original content.

It was commented that separating artistic intent into two purposes might be confusing.

[JVET-AG0160](https://jvet-experts.org/doc_end_user/current_document.php?id=13716) AHG9/AHG13: FGS Extension SEI message useful descriptors [G. Teniou, S. Wenger, A. Hinds (Tencent)]

While defining an FGS extension SEI message primarily addressing the functionality of spatial adaptation, it was felt useful to also include descriptive parameters of the film grain synthesis such as the type of grain that is meant to be reproduced, the purpose of the film grain, such as creative artistic intent or artefact masking, and its essentiality from a service viewpoint.

It was asked what the benefit is that a decoder could conclude from receiving such information? An exampole was given that a decoder might not use the film grain that comes with the bit stream, if it has by itself a good implementation of 35mm grain, for example.

Essentiality is also signalled (4 levels).

It was requested to provide more evidence about benefit and use cases.[JVET-AG0215](https://jvet-experts.org/doc_end_user/current_document.php?id=13771) AHG9/AHG13: Region-dependent film grain characteristics [P. de Lagrange, E. François, M. Le Pendu, C. Salmon-Legagneur (InterDigital)] [late]

This contribution, while stating that the current FGC SEI message fits its purpose to describe the grain present on scanned film or added for artistic intent on motion pictures, and enables to recover a similar grain by driving synthesis at decoder side at reasonable cost when the grain has been lost in the encoding process for broadcast or streaming purpose, acknowledges that for other uses cases, it may be desirable to describe different film grain in different regions of the picture, for example when the picture is a composite of different sources.

A region-dependent film grain characteristics SEI message is proposed, that includes the existing FGC SEI message syntax, notably independent intensity intervals with model parameters for each interval, and adds elements to describe a list of picture regions, and associate each region with a set of intervals. The commonalities with the FGC SEI message are believed to provide a simple migration path, and simple conversion back to FGC SEI messages for parts of pictures by a network processor, for example in the context of subpictures.

The proposal would also allow different characteristics of film grain in different rectangular regions (not supported by JVET-AG0101 and JVET-AG0140). This includes a common part of the model, and definition of “intensity intervals” to individually control the parameters per region. This would not be the same as controlling the strength by an alpha map, but the effect might be similar.

On the other hand, an alpha map would support more than rectangular regions.

It was commented that all region-based approaches add more complexity and may not be easy to implement. Even frame-based film grain synthesis has not reached a wide deployment so far. The most critical point is sample dependency, where adding another step for alpha blending, or changing the model locally might be a problem. Also on/off requires additional check.

In software, an implementation might be less critical than for the case of dedicated hardware. It was argued that it is up to the decoding device to run it or not (a counter-argument is made that for certain applications the expectation that film grain would be overlaid is important). Use cases may be different, though (more evidence about that would be needed).

It was commented that, as the synthesis related to FGC is computation-intense, the tradeoff with the additional benefit should carefully be considered, and more evidence of the benefit should be provided.

It was requested by proponents to include some approach in TuC. This would require software implementation to also demonstrate the benefit and investigate the complexity impact. Proponents are asked to consider combining their approaches and working on an implementation in software, which would allow more detailed study. See further notes under JVET-AG0328.

[JVET-AG0290](https://jvet-experts.org/doc_end_user/current_document.php?id=13863) [AHG13] Compressibility analysis of Film Grain test sequences [S. Paluri, J. Kim, D. Podborski, A. M. Tourapis (Apple)] [late]

This document presents results on the compressibility of the film grain content introduced in JVET-AF0262. The sequences are compressed using the FFMPEG software and the encoding statistics are analyzed. Such information could be utilized when selecting content for different experiments or evaluations within JVET or other groups.

This analysis is deemed to be extremely helpful in the context of selecting test cases for the planned subjective test, to have a variety of content, e.g. with different levels of complexity, and identify which sequences are hard to encode when film grain is present. BoG to prepare version of JVET-AD2022?

In a follow-up discussion, it was suggested to also add some sequences with film grain in CTC, as it is known that they could be difficult to encode (not removing the film grain, not using synthesis).

[JVET-AG0295](https://jvet-experts.org/doc_end_user/current_document.php?id=13868) AHG13: Frequency domain Film Grain Objective Metrics with Adaptive Region Selection [X. Meng, W. Zhang, S. Labrozzi (Disney Streaming)] [late]

In the entertainment industry, film grainis widely present in TV and motion pictures. However, there are no widely used objective metrics to evaluate the quality of the synthesized grainy video, which highly limits the development of film grain related technologies. In the 32rd JVET meeting, we proposed a subband based film grain assessment method (SFGA). Although SFGA not only enhances our understanding of film grain technology, but also provides a valuable tool for fine-tuning film grain parameters in the modeling and synthesis process, it has several shortcomings that need to be addressed.

* First, the SFGA is conducted on the grain-only image, which means that the clean/denoised image is needed. However, the clean image is usually unavailable, and the denoiser may be not strong enough to extract the film grain from the grainy image properly.
* Second, the SFGA is performed on the manually selected flat regions. However, in the practical use cases, manually selecting all the flat regions for each image is unrealistic. So adaptive region selection is essential.

In this proposal, we extend the frequency domain film grain objective metric to directly take effect on the grained video. Specifically, an automatic grain region selection scheme is proposed. Experimental results show that the film grain metric with region selection on content+grain has a better performance compared to the SFGA proposed in JVET-AF0209. This also greatly improves the completeness of the method and expands the application scenarios.

The correlation of the metric with MOS was found to be higher for plain film grain (not overlaid to movie). Therefore, in case of film grain overlaid, the analysis is concentrating on flat regions.Pearson rank correlation is roughly 80% with the new method.

Software might be provided (needs to ask for permission), which could be highly valuable in the preparation of quality viewing.

Flatness is classified via edge maps. It was commented that this might possibly classify textured regions as “flat”.

Currently only used for luma. It was commented that considering chroma (e.g., in RGB domain) might also be beneficial.

It was commented that potentially using the new material with ground truth might be interesting.

[JVET-AG0316](https://jvet-experts.org/doc_end_user/current_document.php?id=13889) Update on Technical Report “Film grain synthesis technology for video applications” [A. Norkin (Netflix)] late]

This document informs JVET that AOMedia has finalized the AFGS1 specification, which is referenced by the Technical Report “Film grain synthesis technology for video applications” [1]. The document also includes information on what TR editors may consider updating in the TR, if possible.

It is reported that some modifications have been made in the latest version of AFGS1, and as a consequence it is suggested to implement an update in the TR.As AFGS1 has been finalized, it is believed that one of the main concerns of ISO editors has been resolved. There was also a comment that it is only accessible via a github link, which should however be less critical if the document is stable. This could be achieved if a specific version is referenced.

New version (draft 6) of the technical report to be issued.

The newly added subclause (7.4.2.7) describes a feature of the latest version of AFGS1 that is not supported by the AV1 version. This may be in conflict with another statement in 9.3 of the TR, that it would be equivalent with the implementation in AV1.

There are also changes in naming syntax elements, which should not be critical.

It was agreed to include the new subclause, and potentially remove it when there are complaints by ISO editors.

[JVET-AG0324](https://jvet-experts.org/doc_end_user/current_document.php?id=13897) AHG4/AHG13/AHG3: film grain tuning tool tutorial [P. de Lagrange (InterDigital)] [late]

This contribution describes how to use the “FGC SEI designer” graphical tool to display and interactively edit an FGC SEI configuration file. That tool is available in the JVET gitlab (in “JVET AHG Film Grain Technologies” group). The intent is to encourage experimentation and showcase the capabilities of the FGC SEI in conjunction with the implementation provided in the same repository, which is asserted to be hardware-friendly and interoperable with different film grain models.

It is also asked whether interactive testing based on this tool could be used in the context of AHG4/AG5, to evaluate the grain fidelity aspect considered in the second category of the FGC SEI visual testing plan (JVET-AD2022).

Was presented at 2135 on Tuesday Jan. 23 (chaired by JRO and MW)Software tool is available on proponents’ github, and also in JVET gitlab.

The tool could be highly useful in the preparation of visual test for film grain.

It was commented that professionals creating artistic film grain would likely not use this tools, as they have their own tools, and further developing those is part of their job.

It was reported that not all film grain on digital materials is created by artists, there are also cameras that insert the film grain right after capture.

[JVET-AG0328](https://jvet-experts.org/doc_end_user/current_document.php?id=13901) AHG9/AHG13: unified proposal on film grain regions characteristics SEI message [G. Teniou, S. Wenger, A. Hinds (Tencent), P. de Lagrange, E. François, D. Doyen, M. Le Pendu, C. Salmon-Legagneur (InterDigital), P. Wu, Y. Gao, S. Xie, Y. Bai, M. Jia, C. Huang (ZTE)] [late]

This contribution proposes a new SEI message for conveying Region-dependent film grain characteristics. It is a unified proposal based on contributions JVET-AG0140, JVET-AG0215 and JVET-AG0101. The proposal describes a region-dependent film grain characteristics SEI message, that includes the existing FGC SEI message syntax, notably independent intensity intervals with model parameters for each interval and adds elements to describe a list of picture regions, and associate each region with a set of intervals.

If agreeable by the group, it is proposed to document this proposed solution in the TuC so that proponents can use it for running experiments as requested during the 1st presentation of the different contributions.

The proponents will provide a software implementation of the proposed SEI.

Regions can overlap. Max number of regions increased relative to JVET-AG0140.

Agreed to include in TuC, not in WD for VSEIv4.

## Implementation studies (0)

Contributions in this area were discussed at XXXX–XXXX on XXday XX Jan. 2024 (chaired by XXX).

## Profile/tier/level specification (1)

Contributions in this area were discussed at 0905–0920 on Wednesday 24 Jan. 2024 (chaired by JRO).

[JVET-AG0201](https://jvet-experts.org/doc_end_user/current_document.php?id=13757) On MV-HEVC profiles [Y.-K. Wang, H. Liu, L. Zhang, S. Jiao, C. Hu, J. Cui, G. Xu (Bytedance), A. M. Tourapis, D. Podborski, S. Paluri (Apple)]

It is asserted that, although the syntax of MV-HEVC allows for an output layer set (OLS) in an MV-HEVC bitstream to have both texture layers and (depth or auxilary) layers, where an auxilary layer could carry alpha or depth, the definitions of the existing Multiview Main profile and the being-developed Multiview Main 10 profile do not allow *having both auxiliary (alpha or depth) layers with chroma\_format\_idc equal to 0 and texture layers in an OLS* (referred to as “*the wanted combination*” below for simplicity).

It is therefore proposed to add a second Multiview Main 10 profile and also a second Multiview Main profile that both allow *the wanted combination*. The proposal also includes the addition of the Multiview Monochrome 16 profile and updates to the being-developed Multiview Monochrome, Multiview Monochrome 10, and Multiview Monochrome 12 profiles, and all these monochrome profiles allow *the wanted combination*.

The exact proposed text changes are in the attachment of this contribution.

There were no specific comments made on the technical content. Proponents were suggested to try initiating a ballot comment on the DIS, such that it could be discussed at the next meeting.

It was commented that the term “wanted combination” is unclear, as it does not tell what the combination is about.

## Gaming content compression (AHG15) (1+1)

Contributions in this area were discussed at XXXX–XXXX on XXday XX Jan. 2024 (chaired by XXX).

[JVET-AG0047](https://jvet-experts.org/doc_end_user/current_document.php?id=13601) AHG15: Report on AhG meeting on gaming content compression [S. Puri, J. Sauer, R. Chernyak, A. Duenas, L. Wang]

Was basis of discussions in BoG JVET-AG0331.

[JVET-AG0302](https://jvet-experts.org/doc_end_user/current_document.php?id=13875) AHG15: Description of gaming content sequences proposed by Huawei [Z. Lin, W. Chen, Y. Zhao, K. Cai, J. Sauer, E. Alshina (Huawei)] [late]

Was discussed in BoG JVET-AG0331.

## Generative face video (AHG16) (4+4)

Contributions in this area were discussed at 0720–0940 on Tuesday 23 Jan. 2024, and 0605–XXXX on Thursday 25 Jan. 2024 (chaired by JRO).

[JVET-AG0042](https://jvet-experts.org/doc_end_user/current_document.php?id=13595) AHG 16: Proposed Common Software Tools and Testing Conditions for Generative Face Video Compression [B. Chen, J. Chen, R.-L. Liao, Y. Ye (Alibaba), S. Wang (CityU)]

At the 32nd JVET meeting, a new Ad hoc Group on Generative Face Video Compression (GFVC) has been established with mandates to investigate aspects related to software implementation, test conditions, coordinated experimentation, interoperability study and so on. In this contribution, we propose a unified software package that incorporates various GFVC methods that use different face video representations and allows coding to be performed using VVC Main 10 profile. We also propose test data with suitable copyright license and corresponding test conditions for experimentation. Based on the proposed test conditions, we provide rate-distortion performance showing that GFVC can achieve significantly better reconstruction quality than the state-of-the-art VVC standard at ultra-low bitrate ranges. The proposed software package is available at <https://github.com/Berlin0610/GFVC_Software>.

In version 2 of this contribution, the distortion measurement has been amended according to the discussion at the AHG conference call on 12/7 by adding PSNR and SSIM calculation in the YUV420 domain. The github repo has been updated with the latest metric calculation functionality. Rate distortion performance in terms of PSNR and SSIM (in RGB 444 and in YUV 420) has been added to this contribution. The xls attachment has also been revised to show per-sequence rate-distortion performance and rate-distortion plots.

In version 3 of this contribution, some extra sequences that are not included in the currently proposed test data but have complex head motion and camera motion are tested to provide some examples GFVC cannot handle.

Update of contribution that had already been presented in AHG telco.

The arithmetic encoder is conventionally encoding the parameters (not integrated with NN).

The currently available sequences are 256x256 – would be desirable to have content with higher resolution.

It could be considered to use cropped parts (head and shoulder) from some videoconferencing sequences that had been provided in the context of the verification test.

The software performs extraction and synthesis related to the proposed SEI, but the software currently is writing a raw bitstream, not implementing the transport via SEI in VVC bitstream yet.

What is the decoder complexity? Some synthesis models exist that have been demonstrated in real-time (not the three approaches supported by this software).

Crosscheck JVET-AG0323 confirms results.

It was agreed to make the software available in an AHG repository related to VTM.

[JVET-AG0323](https://jvet-experts.org/doc_end_user/current_document.php?id=13896) Cross-check of JVET-AG0042 (AHG 16: Proposed Common Software Tools and Testing Conditions for Generative Face Video Compression) [J.-Y Thong, H.-B. Teo (Panasonic)] [late]

[JVET-AG0043](https://jvet-experts.org/doc_end_user/current_document.php?id=13596) AHG16: Report on AhG meeting on generative face video compression [J.-R. Ohm, Y. Ye, H.-B. Teo, Z. Lyu, S. McCarthy, S. Wang]

Was subsumed in JVET-AG0016, no need for presentation.

[JVET-AG0048](https://jvet-experts.org/doc_end_user/current_document.php?id=13602) AHG16: Interoperability Study on Parameter Translator of Generative Face Video Coding [S. Yin, B. Chen, J. Chen, R.-L. Liao, Y. Ye (Alibaba), S. Wang (CityU)]

At the 32nd JVET meeting, a new Ad hoc Group on Generative Face Video Compression (GFVC) has been established with mandates to investigate aspects related to software implementation, test conditions, coordinated experimentation, interoperability study and so on. In this contribution, we investigate the issue of interoperability among different GFVC systems. We use face parameter translators integrated into the decoder side to convert different types of received face parameters to decoder-specific ones. We also evaluate the rate-distortion performance of the proposed parameter translator mechanism following the proposed test conditions in JVET-AG0042. We show that different face parameters from different GFVC algorithms can be effectively adapted to one another while retaining promising coding performance. We make the translator implemtnation available at <https://github.com/xyzysz/GFVC_Software-Decoder_Interoperability>

Parameter translator is a fully connected network with 4 layers, 6 different translators with approx. 30K parameters and 60kMAC per picture.

Metrics DISTS and LPIPS are used, which are reported to be often used in context of assessing quality in this area. The parameter translator is not using those metrics, only tries to match the features generated by the other encoder as close as possible.

It was commented that results might be better with a training where the output is considered. This might be possible, as both parameter translator and synthesis are non-normative and can be optimized for a given decoding device.

Generally, translation appears to work, but there is some quality drop. However, in some cases (in particular coding with FOMM, decoding with one of the others), even a quality improvement can be observed.

It was agreed to make the translator software available in the AHG repository related to VTM.

[JVET-AG0087](https://jvet-experts.org/doc_end_user/current_document.php?id=13643) AHG9: On the generative face video SEI message [M. M. Hannuksela, F. Cricri, H. Zhang (Nokia)]

This contribution proposes the following asserted improvements and fixes to the generative face video (GFV) SEI message proposed in JVET-AF0234:

1. It is proposed to avoid the mandatory definition of TranslatorNN in each base picture and instead gate the definition of the TranslatorNN with a flag. The definition is required to be present in the first GFV SEI message that is included in or follows an IRAP picture unit in output order.
2. It is proposed to avoid dummy pictures and instead allow multiple GFV SEI messages in a picture unit, e.g., carried within suffix SEI NAL units. Additional ingredients of this item include the following:
   1. Each new instance of a GFV SEI message will invoke the generative neural network (NN) inference.
   2. To differentiate between repetitions and new instances of GFV SEI messages, it is proposed to add a counter syntax element (gfv\_cnt) in the syntax.
3. It is proposed to enable referencing the same drive picture multiple times. This item includes the following ingredients:
   1. It is proposed to indicate whether the current picture is a drive picture with a new flag, gfv\_drive\_pic\_flag,
   2. The semantics of gfv\_drive\_pic\_fusion\_flag continue to indicate if the drive picture is used as input to GenerativeNN.
4. It is proposed to remove:

NOTE 5 – When current decoded picture corresponds to a driving picture, it should be marked as not for output purpose.

About item 1: It was commented that this parameter translator would be a generic one (sent somehow out of band), different from the translator implementation presented in JVET-AG0048.

Agreed. that the suggested modification is useful (enable the capability of sending parameter translator description only at an IRAP).

About item 2:

Agreed. This modification avoids sending dummy pictures and saves bits, also makes the SEI mor “self contained” in determining the timing of the output.

It was commented that additionally the timing of the output needs to be specified (which was so far taken care about by the means of dummy picture). Further study on that aspect.

About items 3/4:

It was commented that this may somehow modify the concept of drive pictures, and whether they are used for fusion or not. The definition of drive picture might also require clarification. Is “drive picture” a good name? After offline communication, item 3 was a wrong interpretation of drive pictures For item 4, the following modification is suggested:

NOTE 5 – When current decoded picture corresponds to a driving picture, the current decoded picture is not intended to be displayed.

The agreed items to be included in the generative face SEI (in TuC)

[JVET-AG0088](https://jvet-experts.org/doc_end_user/current_document.php?id=13644) AHG9: Usage of the neural-network post-filter characteristics SEI message to define the generator NN of the generative face video SEI message [M. M. Hannuksela, F. Cricri, H. Zhang (Nokia)]

This contribution proposes to enable the use of a neural-network post-filter (NNPF) defined in the NNPF characteristics (NNPFC) SEI message as the generator neural network (GenerativeNN) of the generative face video (GFV) SEI message design proposed in JVET-AF0234. Ingredients of this proposal include the following:

1. The use of the NNPF auxiliary input mechanism to input face parameters and matrices to the NNPF inference, including the following aspects:
   1. nnpfc\_auxiliary\_inp\_idc & 2 greater than 1 specifies that the auxiliary input is derived from SEI message(s) of a given payload type value (nnpfc\_sei\_aux\_payload\_type).
   2. When nnpfc\_sei\_aux\_payload\_type is equal to the payload type value of the GFV SEI message, nnpfc\_sei\_aux\_id is present and indicates the gfv\_id value of the GFV SEI message for which this NNPF serves as the GenerativeNN.
   3. When nnpfc\_sei\_aux\_payload\_type is equal to the payload type value of the GFV SEI message, the auxiliary inputs are inputDriveKeyPoint, inputDriveMatrix, inputBaseKeyPoint, and inputBaseMarix as derived in the semantics of the GFV SEI message.
2. Addition of generative modelling among the purposes indicated by nnpfc\_purpose.
   1. Generative modelling is proposed to be added as the next available bit in nnpfc\_purpose.
   2. It is proposed to constrain generative modelling and general visual quality enhancement to be mutually exclusive.

The NNPF is suitable for this purpose, as the input would be base pictures and drive pictures, and parameter as aux information, output are the synthesized pictures.

It was commented that it would be interesting to analyse the size of the NNPF representation expressed by NNR. On the other hand, if an URL of an existing synthesis method would be signalled, the advantage of having precise matching between encoder and decoder would be retained.

It is not the goal of this contribution to replace anything in the existing SEI.

Generally this option of signalling the synthesis method is agreed to be beneficial, and would get rid of any interoperability problem.

It was commented that another approach could be to make the GFV SEI self contained in this regard, to include the whole signalling in it, similar as already for the parameter translator. This might be simpler than being forced to implement the entire NNPF SEI additionally.

A new purpose would need to be defined in NNPF, and if that purpose is generically “generative modelling”, it might be much less constrained than for the mere purpose of face generation.

Further study on these aspects.

[JVET-AG0139](https://jvet-experts.org/doc_end_user/current_document.php?id=13695) AHG16: Depthwise separable convolution for generative face video compression [R. Zou, R.-L. Liao, B. Chen, J. Chen, Y. Ye (Alibaba)]

This contribution proposes to introduce the depthwise separable convolution into the CFTE method of generative face video compression. More specifically, partial 3x3 convolution layers in CFTE are replaced by depthwise separable convolution to reduce the number of parameters and computation complexity. It is reported that compared with the original CFTE, the proposed method reduces number of parameters and kMACs per pixel by 41% and 4%, respectively. The overall coding performance impact of the proposed method for {DISTS, LPIPS, RGB444\_PSNR, RGB444\_SSIM, YUV420\_PSNR, YUV420\_SSIM} is { -62.16%, -60.17%, -5.49%, -33.75%, -2.64%, -28.97%} when comparing with VTM-22.2 and {-4.46%, -5.16%, -1.82%, -10.69%, -8.61%, -1.55% } when comparing with the original CFTE.

Interesting improvement of CFTE (both in terms of complexity and performance).

It was commented that the GFC models are not by themselves part of the software package, but need to be downloaded from some external URL, such that they can be run within the software.

It is up to proponents to provide the new model as an additional one, or replacing the previous CFTE model (it was suggested to better keep the previous model. This could also be useful to investigate interoperability between different versions of a model).

It was agreed to modify the GFVC software to support this model.

In the future, a more general mechanism to link with more external models (or allow a user of the software to configure it for using additional models) would be beneficial.

[JVET-AG0187](https://jvet-experts.org/doc_end_user/current_document.php?id=13743) AHG16: Study text for common test conditions and evaluation procedures for generative face video coding (draft 1) [S. McCarthy, P. Yin (Dolby), B. Chen, Y. Ye (Alibaba), S. Wang (CityU)]

This document proposes draft text to define common test conditions and software reference configurations for testing of generative video face coding. The conditions and configurations are tentatively recommended for use in applicable technical contributions to JVET meetings.

It was commented that also the ratio of bits that are used for base picture and other pictures should be reported. As the QP for base pictures is lower than for the I picture in VVC in some cases, this might cause different latency.

It was suggested to investigate whether there is sufficient overlap in terms of quality, as otherwise computing BD rate would be meaningless. An information about the overlap should be added in the .xls.

It was agreed to convert this into an output document. It is understood that this is a first version, to be extended in particular also into the direction of higher quality ranges (as discussed in AHG).

[JVET-AG0203](https://jvet-experts.org/doc_end_user/current_document.php?id=13759) AHG9/AHG16: Common text for proposed generative face video SEI message [J. Chen, B. Chen, Y. Ye (Alibaba), S. Yin, S. Wang (CityU), S. McCarthy, P. Yin, G.-M. Su, A. K. Choudhury, W. Husak, G. J. Sullivan (Dolby)]

This contribution updates the common text of generative face video (GFV) SEI message proposed in JVET-AF0234 to 1) make the signaling of translator neural network optional by adding a gating flag; 2) support predictive signaling of matrix elements to align with the key-point coordinate signaling; 3) refine the interface between the translator neural network and generative neural network. The proposed GFV SEI message is intended for ultra-low bitrate face video compression applications such as video conferencing, live entertainment, and face animation.

The aspect 1 (gating flag for signalling the translator) was already agreed to be modified in context of JVET-AG0087.

Aspect 2 (predictive signalling of matrix elements) is agreed.

Aspect 3 (more flexible interface between translator and model in terms of numbers of keypoints and matrix elements) is agreed.

It was agreed to include the generative face SEI into TuC. This shall include the aspects that were agreed from JVET-AG0087.

Sufficient evidence has been demonstrated that the concept could work via an SEI message.

Software to be provided that implements that SEI message in the context of the generative face software, including transport of the data in a VVC stream.

# Low-level tool technology proposals (106)

## AHG11/AHG14: Neural network-based video coding (19+2)

### Summary, BoG reports, and information documents (3)

Contributions in this area were discussed at 2100–2310 on Wednesday 17 Jan. 2024 (chaired by JRO).

[JVET-AG0023](https://jvet-experts.org/doc_end_user/current_document.php?id=13790) EE1: Summary report of exploration experiment on neural network-based video coding [E. Alshina, F. Galpin, Y. Li, D. Rusanovskyy, M. Santamaria, J. Ström, R. Chang, Z. Xie (EE coordinators)]

This report summarizes the activities of the Exploration Experiment 1 (EE1) performed between the 32nd and 33rd JVET meetings to evaluate Neural Network-based Video Coding (NNVC) technologies, analyze their performance, evaluate their complexity aspects, and clarify training procedure. One teleconference has been conducted during this meeting circle.

Introduction

This round of EE1 tests includes:

* EE1-1: HOP
  + EE1-1.0: HOP re-training and luma/chroma balance changes ([JVET-AF0155](about:blank), [JVET-AF0180](about:blank)), possible tuning from JVET-AF0150, JVET-AF0296
  + EE1-1.1: EE1-0 with architecture change variant 1 (JVET-AF0102, JVET-AF0103, JVET-AF0182), as defined in [JVET-AF0307](https://jvet-experts.org/doc_end_user/current_document.php?id=13571).
  + EE1-1.2: EE1-0 with architecture change variant 2 (JVET-AF0102, JVET-AF0103, JVET-AF0153), as defined in [JVET-AF0307](https://jvet-experts.org/doc_end_user/current_document.php?id=13571).
  + EE1-1.3: Comparison test for HOP single model and two models (JVET-AF0183)
  + EE1-1.4: Study joint inference design of [JVET-AF0154](about:blank) (rotation) and [JVET-AF0086](about:blank) (flipping)
  + EE1-1.5: HOP In-loop filter with transformer blocks (JVET-AF0158)
* EE1-2: LOP
  + EE1-2.1: LOP/HOP fast training (JVET-AF0043, fast Stage III training)
  + EE1-2.2: LOP Content adaptive (JVET-AF0056)
  + EE1-2.3: Further complexity reduction of LOP (JVET-AF0071)
* EE1-3: inter prediction
  + EE1-3.1: Deep Reference Frame Generation for Inter Prediction Enhancement (JVET-AF0208)
* EE1-4: super-resolution
  + EE1-4.1: Unified CNN super resolution for resampling-based video coding (JVET-AF0143).

The anchor for EE1 tests is the default configuration of NNVC-7.0 as defined by AhG11/AhG14 (NN-intra and low complexity NN-filter are enabled by default) in JVET-AE2016. Anchor performance and reference point for HOP NN-filters was be provided by AhG14 (shown in the Table 1).

*Table 1 EE1 anchor performance vs VTM.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access** | | | | | | | |
|  | **BD-rate Over NNVC-7.1 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -7.9% | -10.1% | -10.6% | -8.5% | -11.2% | -10.7% | 129% | 9031% |
| Class A2 | -6.8% | -12.7% | -9.7% | -7.0% | -11.8% | -7.3% | 126% | 8426% |
| Class B | -6.6% | -13.9% | -12.7% | -6.8% | -12.9% | -11.9% | 127% | 9076% |
| Class C | -6.6% | -14.9% | -14.4% | -7.1% | -13.2% | -12.9% | 120% | 7892% |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | **-6.9%** | **-13.2%** | **-12.1%** | **-7.3%** | **-12.4%** | **-11.0%** | 126% | 8606% |
| Class D | -7.2% | -14.8% | -15.0% | -6.5% | -13.2% | -13.1% | 119% | 7530% |
| Class F | -3.7% | -9.5% | -9.4% | -4.2% | -10.2% | -9.8% | 136% | 3684% |
|  |  |  |  |  |  |  |  |  |
|  | **Low delay B** | | | | | | | |
|  | **BD-rate Over NNVC-7.1 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class B | -5.3% | -13.4% | -13.3% | -6.1% | -14.7% | -17.5% | 120% | 8941% |
| Class C | -5.5% | -13.7% | -14.9% | -7.4% | -15.5% | -17.0% | 113% | 7756% |
| Class E | -6.1% | -8.9% | -12.1% | -7.8% | -10.5% | -13.5% | 136% | 9072% |
| **Overall** | **-5.6%** | **-12.4%** | **-13.5%** | **-6.9%** | **-13.9%** | **-16.3%** | 121% | 8558% |
| Class D | -6.5% | -13.1% | -15.4% | -7.3% | -14.2% | -17.9% | 110% | 7563% |
| Class F | -3.1% | -8.7% | -8.7% | -4.2% | -9.9% | -13.2% | 129% | 4227% |
|  |  |  |  |  |  |  |  |  |
|  | **All Intra** | | | | | | | |
|  | **BD-rate Over NNVC-7.1 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -8.6% | -12.7% | -12.3% | -9.0% | -14.3% | -12.6% | 187% | 6908% |
| Class A2 | -7.1% | -12.7% | -11.3% | -7.5% | -12.3% | -9.2% | 186% | 5880% |
| Class B | -7.3% | -13.0% | -13.5% | -7.4% | -13.3% | -13.4% | 181% | 5701% |
| Class C | -7.5% | -13.8% | -14.5% | -7.9% | -14.3% | -14.6% | 168% | 4342% |
| Class E | -10.7% | -14.1% | -15.1% | -10.9% | -13.6% | -14.1% | 176% | 6176% |
| **Overall** | **-8.1%** | **-13.3%** | **-13.4%** | **-8.4%** | **-13.6%** | **-13.0%** | 179% | 5644% |
| Class D | -7.7% | -12.7% | -15.0% | -7.5% | -13.6% | -15.4% | 160% | 4159% |
| Class F | -5.3% | -10.8% | -10.4% | -5.1% | -11.3% | -11.8% | 141% | 4847% |

Summary **of all test results**

Summary of all EE1 tests results is shown in Table 2.

The key aspects of each EE1 test are described below.

HOP filter category:

* EE1-1.0 is retraining of existing unified HOP filter architecture with better Luma-Chroma performance balance. As agreed during AhG11/AhG14 teleconference ([JVET-AG0041](https://jvet-experts.org/doc_end_user/current_document.php?id=13594)) it became new NNVC-7.1 HOP2 filter design and comparison point for all technologies targeting replacement HOP filter. This became the basis of all other experiments in the HOP category.
* EE1-1.1 [JVET-AG0174](https://jvet-experts.org/doc_end_user/current_document.php?id=13730) key aspects: group convolutions in residual block (JVET-AF0182), alternation of two types of residual blocks and decomposition rank reduction, i.e. C31=48 (JVET-AF0102), 1×1 convolutions instead of 3×3 over IPB and BS inputs (JVET-AF0103).
* EE1-1.2 withdrawn
* EE1-1.3 [JVET-AG0175](https://jvet-experts.org/doc_end_user/current_document.php?id=13731) key aspects: two separate models for Luma and Chroma with total complexity 473 (366 for Luma, 107 for Chroma) no changes of HOP architecture beside of output and number residual blocks.
* EE-1.4 [JVET-AG0056](https://jvet-experts.org/doc_end_user/current_document.php?id=13612) key aspects: both rotation and flipping are combined to geotransform for input / output NNLF, tested in combination with both LOP and HOP. 0.4% gain in LDB configuration.
* EE-1.5 [JVET-AG0162](https://jvet-experts.org/doc_end_user/current_document.php?id=13718) key aspects: simplified (w/o softmax and layer normalization) transformers based attention module is added to two (#8 and #15) out of 26 residual blocks (additionally “rank of decomposition” C31 is reduced to 48 and alternating BB type, similar to EE1-1.1).

LOP filter category:

* EE1-2.1 withdrawn
* EE1-2.2 [JVET-AG0111](https://jvet-experts.org/doc_end_user/current_document.php?id=13667) key aspects: content adaptivity for LOP, overfitting multiplier in activation in K layers (K depends on bit-rate overhead), only class C results provided so far (0.5% Luma gain in RA cfg). The test is based on the LOP1 architecture.
* EE1-2.3 [JVET-AG0163](https://jvet-experts.org/doc_end_user/current_document.php?id=13719) key aspects: skip connection is modified to have no start/end point between the two 1x1 convolutions and the two 1x1 conv layers could be fused into one single 1x1 convolutional layer. Training strategy (batch size and learning rate) is modified compared to anchor.

NN-Inter category:

* EE1-3.1 [JVET-AG0122](https://jvet-experts.org/doc_end_user/current_document.php?id=13678) key aspects: deep reference frame generation with down-sampled input, ‘small IFRNet’ and up-sampling resulting optical flow, reduced number of channels, removal of some convolution operation. The complexity of model is reduced from 504 kMAC/pxl to 69 kMAC/pxl during this EE. Training uses license free HAA500 data base.

Super-resolution category:

* EE1-4.1 [JVET-AG0130](https://jvet-experts.org/doc_end_user/current_document.php?id=13686) key aspects: Super-resolution filter in NNVC has 469kMAC/pxl complexity. Two variants of SR filter with Low complexity (20 kMAC/pxl) and same as existing SR filter complexity (469 kMAC/pxl), both based on unified filter architecture are tested in absence of other NNVC tools (NNVC configured as VTM).

*Table 2 Summary of results for all EE1 tests vs NNVC-7.1 default configuration* (NNLF-LOP2 & NN-Intra).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Test** | **Random Access** | | | **All Intra** | | | **kMAC/pxl** | **Param** | **Source** |
| Y | U | V | Y | U | V |  |  |  |
| NNVC-7.1 [LOP] | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 25.1 | 1.6 | [JVET-AG0014](https://jvet-experts.org/doc_end_user/current_document.php?id=13597) |
| EE1-1.0 [HOP] | -7.2% | 0.8% | -2.3% | -4.8% | 2.3% | 0.5% | 485 | 3 | [JVET-AG0014](https://jvet-experts.org/doc_end_user/current_document.php?id=13597) |
| EE1-1.1 [HOP] | -7.4% | -0.9% | -2.7% | -5.0% | 2.0% | 0.5% | 474 | 2.9 | [JVET-AG0174](https://jvet-experts.org/doc_end_user/current_document.php?id=13730) |
| EE1-1.3 [HOP] | -7.1% | -1.7% | -3.5% | -4.8% | 0.5% | -0.8% | 481 | 2.9 | [JVET-AG0175](https://jvet-experts.org/doc_end_user/current_document.php?id=13731) |
| EE1-1.4 [HOP] | -7.3% | 0.7% | -2.7% |  |  |  | 485 | 3 | [JVET-AG0056](https://jvet-experts.org/doc_end_user/current_document.php?id=13612) |
| EE1-1.5 [HOP] | -7.2% | -7.0% | -8.0% | -5.5% | -3.2% | -4.2% | 483 | 2.9 | [JVET-AG0162](https://jvet-experts.org/doc_end_user/current_document.php?id=13718) |
| EE1-1.4 [LOP] | -0.1% | -0.4% | -0.4% |  |  |  | 25.1 | 1.6 | [JVET-AG0056](https://jvet-experts.org/doc_end_user/current_document.php?id=13612) |
| EE1-2.2 | - | - | - |  |  |  |  |  | [JVET-AG0111](https://jvet-experts.org/doc_end_user/current_document.php?id=13667) |
| EE1-2.3 [LOP] | -0.3% | 2.6% | 2.1% | 0.0% | 1.8% | 1.5% | 24.6 | 1.6 | [JVET-AG0163](https://jvet-experts.org/doc_end_user/current_document.php?id=13719) |
| EE1-3.1 [LOP] | -0.6% | -0.3% | -0.3% |  |  |  | 94.1 | 4.5 | [JVET-AG0122](https://jvet-experts.org/doc_end_user/current_document.php?id=13678) |

Performance and complexity for technologies targeting HOP filter architecture or inference improvements in NNVC-7.1 are summarized in Table 3. BD-rate reported relatively to NNVC-1.7- HOP (HOP2 filter and NN-Intra enabled).

*Table 3 Summary of results for HOP filter tests in EE1 vs NNVC-7.1 HOP configuration* (NNLF-HOP2 & NN-Intra).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Test** | **Random Access** | | | **All Intra** | | | **kMAC/pxl** | **Param** | **Source** |
| Y | U | V | Y | U | V |
| EE1-1.0 | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 485 | 3 | [JVET-AG0014](https://jvet-experts.org/doc_end_user/current_document.php?id=13597) |
| EE1-1.1 | -0.2% | -1.7% | -0.3% | -0.2% | -0.3% | 0.0% | 474 | 2.9 | [JVET-AG0174](https://jvet-experts.org/doc_end_user/current_document.php?id=13730) |
| EE1-1.3 | 0.1% | -2.4% | -1.1% | 0.1% | -1.8% | -1.3% | 481 | 2.9 | [JVET-AG0175](https://jvet-experts.org/doc_end_user/current_document.php?id=13731) |
| EE1-1.4 | -0.1% | -0.1% | -0.4% |  |  |  | 485 | 3 | [JVET-AG0056](https://jvet-experts.org/doc_end_user/current_document.php?id=13612) |
| EE1-1.5 | 0.0% | -7.6% | -5.7% | -0.7% | -5.5% | -4.7% | 483 | 2.9 | [JVET-AG0162](https://jvet-experts.org/doc_end_user/current_document.php?id=13718) |

*Table 4 Summary of results for Super-resolution technologies tested vs VTM (NNVC-7.1 all tools disabled)*.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Test** | **Random Access** | | | **All Intra** | | | **kMAC/pxl** | **Param** | **Document** |
| Y | U | V | Y | U | V |
| EE1-4.1.1a | -1.6% | -0.6% | 0.1% | -1.2% | -0.2% | -1.2% | 20 | 0.08 | [JVET-AG0130](https://jvet-experts.org/doc_end_user/current_document.php?id=13686) |
| EE1-4.1.1b | -2.2% | -0.6% | 0.3% | -1.2% | -2.7% | -1.2% | 469 | 1.9 | [JVET-AG0130](https://jvet-experts.org/doc_end_user/current_document.php?id=13686) |

HOP category:

1.1 is basically a simplification, but was slightly modified such that complexity in kMAC/pix became the same as 1.0, and provides some gain in luma and chroma, run time reduction. Training by three independent parties, cross-checked. Decision: Adopt JVET-AG0174

1.3 has small loss in luma, gain in chroma. The approach of using separate models for luma and chroma could also be combined with the simplification of 1.1 Investigate this aspect in next EE.

1.4 (flipping/rotation signalled on block level) is only applied to B slices, no gain was found for I slices. Gain relatively small for RA, reported larger for LB. Encoding runtime is slightly larger, decoding run time is increased, as the filter is enabled more frequently. It was commented that this is less related to exploring NN technology, but rather giving chances of using it more frequently. The gain is relatively small, and it adds some complexity in software for encoder and decoder. It was also commented that by a more extensive training strategy this might also be implicitly done in the NN LF. No action.

1.5 has interesting gain in chroma, and also luma for AI, and it is using transformers in attention mechanisms. Further, though integerization is known to be problematic for transformers, a simplification was implemented that is quantized. Training crosscheck not done yet – further investigate in next EE.

LOP category:

2.2 (adaptive post filter) is based on LOP1 (not the newest LOP). Shows interesting gain, and latency has also been reduced by performing overfitting on segments of 1 s rather than whole sequence as in previous approach. Time to train the model was also reduced. Further investigate in EE with newest LOP, which likely will need some modification of the overfitting process. It was further commented that results are not complete yet, and it will be interesting to review those. JVET-AG0111 was reviewed after completion on Wednesday 24 Jan. 0800. In RA, it provides 1.4% luma gain over NNVC anchor (LOP1 + intrapred enabled), 7.9% with NN tools off. With overfitting, encoder runtime is roughly doubled. The filter is run in-loop, and a new APS is defined for that. Further speedup of encoder should be investigated.

It was asked if there would be problems basing it on LOP2? Implementation was not done yet. This should be investigated in next EE (no inclusion in software at this meeting).

2.3 is a complexity reduction of the current LOP. Some gain in luma for RA, but loss in chroma. For AI, loss in chroma. A different training strategy was used, but it was reported that when using it for the base model as well, still luma gain is retained. Cross-checkers reported that the unbalance of luma and chroma performance should be better understood. Contribution was reviewed (see notes under JVET-AG0XXX).

NN-Inter category:

3.1 complexity was largely reduced relative to previous version, but the gain also dropped from -2.3% to -0.6% (RA luma). An alternative dataset was used for training. Proponents report that the performance with the usual (BVI) set was not satisfactory (may not be appropriate for the optical flow inherent in the network). By using yet another set (Vimeo which would not be license free), a gain of -1.3% would be possible. The previous approach also used Vimeo for training. Further investigate in EE. It was also suggested that when later adopting such an approach to NNVC, it might be good to have a higher-operation point as well.

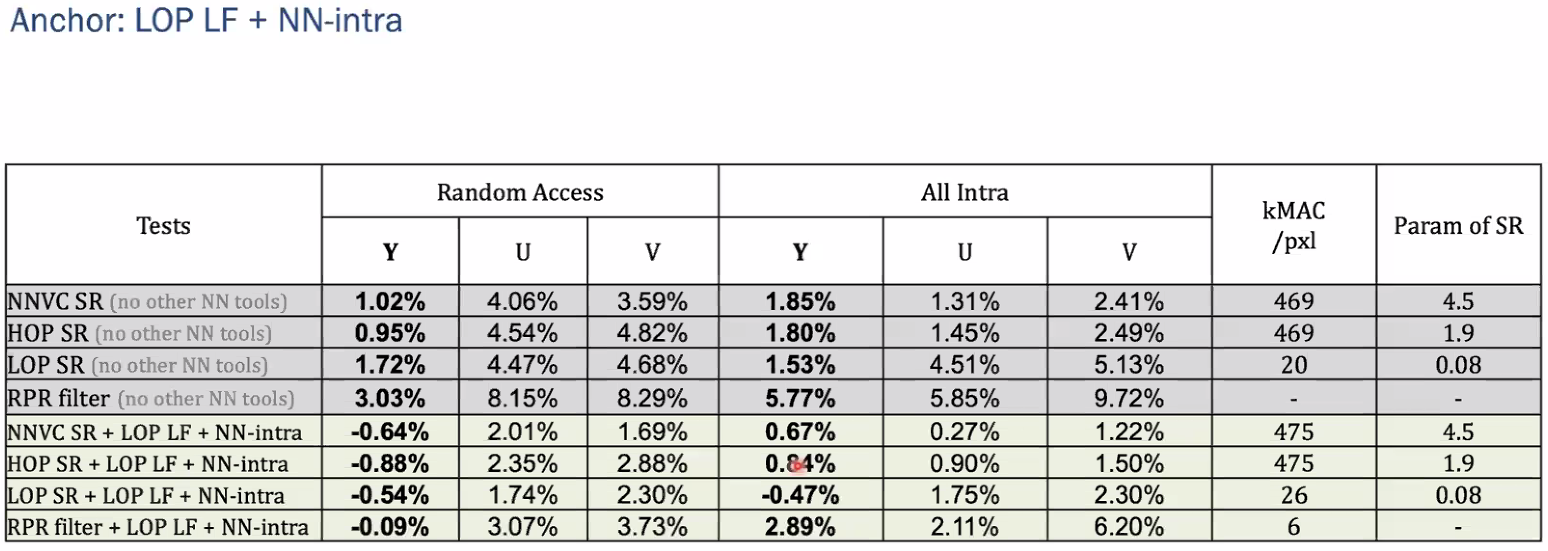
It was commented that for tools that go more towards end-to-end optimization and need optimization for different types of motion the currently available training sets may not be sufficient.

It was suggested to issue a call for training material. BoG JVET-AG0225 was tasked working on that. Clips do not need to be long, but 65 frames (RA period) should be available.

Super-resolution category:

4.1 is testing unified solutions (LOP and HOP). HOP similar in performance to SR currently in NNVC SW, Was not tested in combination with loop filters – results are against VTM anchor, not NNVC. It was also commented that a modified encoder decision (additional RD check) was used than in current SR and RPR. The proponent commented that this was only modified for LOP in AI.

Proponents were asked to provide results on top of NNVC anchor. An update of JVET-AG0130 was presented Wednesday 24 Jan. 0820.



Additional results are shown where original resolution is always preserved in QP22 and 27, and gain is mor consistent.

Concern was expressed about the chroma performance drop. Chroma is also coded in reduced resolution.

No re-training was done for the configuration on top of NNVC. It was suggested to investigate this aspect.

It is the first time that results are presented combining SR and LF, which at least gives some benefit in luma. Compared to the SR approach currently in NNVC, both HOP and LOP show benefit, either in compression (HOP) or in significant reduction of complexity (LOP).

Adopt to NNVC provided that training crosscheck is confirmed. It was reported that JVET-AG0226 confirms that both LOP and HOP (quantized models) were successfully crosschecked with <0.1% deviation.

As the compression benefit of HOP relative to LOP (in CTC, NNVC anchor) is relatively small, and the complexity is significantly larger, it was suggested that only LOP should be adopted to NNVC. It shall replace the existing SR approach, which also has an unreasonable performance/complexity tradeoff compared to LOP.

Decision: Adopt JVET-AG0130 LOP configuration to NNVC SW and JVET-AG2019, disabled by default.

It was commented that the software needs improvement to follow the NNVC software guidelines.

In the ongoing EE on SR (new proposals), an equivalent training strategy should be used, or the HOP/LOP configuration should also be retrained.

General remark on all EEs: It was agreed to impose a deadline to finalization of EE1 results by limiting write access to the GIT (plus 24 hrs grace period).

[JVET-AG0041](https://jvet-experts.org/doc_end_user/current_document.php?id=13594) AhG11/AhG14 teleconference [E. Alshina, F. Galpin, D. Rusanovskyy]

[JVET-AG0315](https://jvet-experts.org/doc_end_user/current_document.php?id=13888) Call for training materials for neural network-based video coding tool development [E. Alshina, F. Galpin, S. Liu, A. Segall (co-chairs), J. Li, R.-L. Liao, D. Rusanovskyy, M. Santamaria, T. Shao, M. Wien, P. Wu (vice chairs)] late]

TBP – to be reviewed during approval of output documents Friday 26 Jan.

### EE1 contributions: Neural network-based video coding (8)

Contributions in this area were discussed in the context of the EE summary report JVET-AG0023, or in the BoG JVET-AG0XXX unless noted otherwise.

[JVET-AG0056](https://jvet-experts.org/doc_end_user/current_document.php?id=13612) EE1-1.4: Joint design of rotation and flipping on NNLF [Z. Xie, Y. Yu, H. Yu, D. Wang (OPPO), R. Chang, L. Wang, X. Xu, S. Liu (Tencent)]

[JVET-AG0294](https://jvet-experts.org/doc_end_user/current_document.php?id=13867) Crosscheck of JVET-AG0056 (EE1-1.4) [J. Li (Bytedance)] [late] [miss]

[JVET-AG0111](https://jvet-experts.org/doc_end_user/current_document.php?id=13667) EE1-2.2: Content-adaptive LOP filter [R. Yang, M. Santamaria, F. Cricri, M. M. Hannuksela, H. Zhang, J. Lainema, A. Hallapuro, D. Bugdayci Sansli (Nokia)]

[JVET-AG0300](https://jvet-experts.org/doc_end_user/current_document.php?id=13873) Crosscheck of JVET-AG0111 (EE1-2.2: Content-adaptive LOP filter) [S. P. Badya, A. Shyam, J. N. Shingala (Ittiam)] [late]

[JVET-AG0313](https://jvet-experts.org/doc_end_user/current_document.php?id=13886) Crosscheck of JVET-AG0111 (EE1-2.2: Content-adaptive LOP filter) [Z. Dai (OPPO)] [late]

[JVET-AG0122](https://jvet-experts.org/doc_end_user/current_document.php?id=13678) EE1-3.1: Deep Reference Frame Generation for Inter Prediction Enhancement [X. Chen, W. Bao, J. Jia, Z. Chen (Wuhan Univ.), Z. Liu, X. Xu, S. Liu (Tencent)]

[JVET-AG0235](https://jvet-experts.org/doc_end_user/current_document.php?id=13808) Crosscheck of JVET-AG0122 (EE1-3.1: Deep Reference Frame Generation for Inter Prediction Enhancement) [Z. Xie (OPPO)] [late]

[JVET-AG0130](https://jvet-experts.org/doc_end_user/current_document.php?id=13686) EE1-4.1: Unified CNN-based super resolution for resampling-based video coding [C. Lin, Y. Li, J. Li, K. Zhang, L. Zhang (Bytedance)]

[JVET-AG0226](https://jvet-experts.org/doc_end_user/current_document.php?id=13799) Crosscheck of JVET-AG0130 (EE1-4.1: Unified CNN-based super resolution for resampling-based video coding) [J. Ye, X. Li, Y. Zhu, Q. Liu (HUST)] [late]

[JVET-AG0162](https://jvet-experts.org/doc_end_user/current_document.php?id=13718) EE1-1.5: Report on implementation of HOP In-loop filter with Transformer blocks [Y. Li, D. Rusanovskyy, M. Karczewicz (Qualcomm)]

[JVET-AG0241](https://jvet-experts.org/doc_end_user/current_document.php?id=13814) Crosscheck of JVET-AG0162 (EE1-1.5: Report on implementation of HOP In-loop filter with Transformer blocks) [D. Liu (Ericsson)] [late]

[JVET-AG0163](https://jvet-experts.org/doc_end_user/current_document.php?id=13719) EE1-2.3: Further complexity reduction on the joint LOP.2 [T. Shao, P. Yin, S. McCarthy (Dolby), J. N. Shingala, A. Shyam, A. Suneja, S. P. Badya (Ittiam)]

A joint Low-complexity Operation Point (LOP) in-loop filter (LOP.2) was proposed in JVET-AF0043 and adopted as the new NNVC LOP anchor. To further reduce the complexity and decoding time, EE1-2.3 proposes a new backbone block architecture for the LOP.2. The proposed residual block fusion (RBF) architecture adjusts the order of the layers in the backbone block architecture and hence the last two 1x1 convolutional layers could be fused into one single 1x1 convolution for further complexity reduction as well as reduction in the number of convolution layers. Compared to the LOP.2, the new RBF design can reduce the complexity of LOP.2 from 17.05 kMac/Pixel to 16.56 kMac/Pixel and reduce the number of convolutional layers by 13%.

Implemented on top of NNVC-7.1 using SADL, the fast stage 3 training results show that the BD-Rate for int16 (RBF LOP + NN INTRA) is {-0.03%, 1.83%, 1.54%} under AI and {-0.25%, 2.62%, 2.11%} under RA compared to NNVC-7.1 (LOP.2 + NN INTRA). Compared to NNVC-7.1 VTM anchor, the BD-Rate for int16 (RBF LOP) is {-4.84%, -8.55%, -8.96%} under AI and {-5.57%, -10.06%, -9.26%} under RA. Experimental results show that the proposed RBF model is slightly better in coding efficiency compared to base LOP.2 and it is coupled with obvious reduction in complexity (number of layers and kMac/Pixel). A BD-Rate deviation of {x.xx%, x.xx %, x.xx %} and {x.xx %, x.xx %, x.xx %} for AI and RA are reported using the model checkpoint trained by the crosschecker.

Was presented Monday 0905-0925 (chaired by JRO)Modified training (batch size, learning rate). When using the same training parameters in LOP2, the gain in luma and loss in chroma stays almost the same.

It was commented that there is loss in compression, and the reduction in complexity (in terms of kMAC/pixel) is relatively low. The reduction of convolutional layers appears more interesting. An operation point with same complexity as LOP2 would be interesting to look at.

It as commented that the chroma branch has less layers, so putting more complexity to chroma processing might resolve the issue of chroma performance drop.

Investigate in next EE along with JVET-AG0155/156.

[JVET-AG0161](https://jvet-experts.org/doc_end_user/current_document.php?id=13717) Crosscheck of EE1-2.3 [D. Rusanovskyy (Qualcomm)] [late]

[JVET-AG0174](https://jvet-experts.org/doc_end_user/current_document.php?id=13730) EE1-1.1: Report on training with HOP architecture change for EE1-0 (variant 1) [Y. Li, C. Lin, J. Li, K. Zhang, L. Zhang (Bytedance), D. Rusanovskyy, Y. Li, M. Karczewicz (Qualcomm), R. Chang, L. Wang, X. Xu, S. Liu (Tencent)]

[JVET-AG0175](https://jvet-experts.org/doc_end_user/current_document.php?id=13731) EE1-1.3: Separate models for HOP filter [Y. Li, C. Lin, J. Li, K. Zhang, L. Zhang (Bytedance), F. Galpin (InterDigital), D. Rusanovskyy (Qualcomm), R. Chang (Tencent)]

### EE1 related contributions: Neural network-based video coding (3)

Contributions in this area were discussed at 2330–0035 on Friday 19 Jan. 2024 (chaired by JRO).

[JVET-AG0155](https://jvet-experts.org/doc_end_user/current_document.php?id=13711) EE1-Related: On Low Complexity Operational Point for In-Loop Filtering [D. Rusanovskyy, Y. Li, M. Karczewicz (Qualcomm)]

The LOP2 filter was adopted to NNVC in result of joint training, report provided in [JVET-AF0043](https://jvet-experts.org/doc_end_user/current_document.php?id=13287), with several additional contributions indicating that further complexity reduction is feasible. In this contribution, LOP2 filter design is revised toward further complexity reduction. Separable convolution was extended toward use in fusion/transition block and LOP architecture is adjusted to accommodate the released complexity reduction budget. Stage 3 training and inference results indicate that proposed design provide acceptable complexity-performance trade-off and is suitable for a very low complexity ILF configuration.

In the LOP2 complexity configuration (targeting 17kMAC/pixel), the proposed method with 9% lower complexity (15.6kMAC/pixel) indicatively provides lower validation cost, comparing to LOP2 anchor. The inference test reportedly provides {0.00%, 1.44%, 0.89%} and {0.08%, 0.98%, 0.75%} BD-rate change in RA and AI configurations, respectively.

In the very low complexity configuration (5kMAC/pixel), the BD-rate changes vs. VTM anchor are reportedly summarized as follows: RA: {-2.3%, -5.3%, -4.6%}; AI: {-2.9%, -5.1%, -5.6%}

It is proposed to investigate the method in the next round of EE1.

Results for very low operation point are not yet integerized.

It was asked what the reason for the larger drop in chroma was? Would need more investigation. However, compared to VTM anchor, the gain in chroma is still better than luma.

Study in EE. Also a version with same complexity (17 kMAC/pix, equivalent to current LOP2) should be studied, to see if it provides gain. This should be the main aspect of this EE.

It was commented that it would also be interesting to see how much of the gain of LOP is retained if run in combination with ECM (AI and RA), but this should better be done outside of the EE, as ist might also benefit from re-training with ECM coded sequences.

The very low operation point should be studied in a subtest of the EE with an integerization.

It was commented by other experts that in general introducing yet another operation point below LOP might not be beneficial if the loss is significant in terms of compression/complexity tradeoff.

[JVET-AG0156](https://jvet-experts.org/doc_end_user/current_document.php?id=13712) EE1-Related: On LOP2 training process [D. Rusanovskyy, Y. Li, M. Karczewicz (Qualcomm)]

The LOP2 filter was adopted to NNVC in result of joint training, report provided in [JVET-AF0043](https://jvet-experts.org/doc_end_user/current_document.php?id=13287), with further complexity reduction being studied in new EE1-2.3. During LOP2 joint training, several tests have been conducted on selecting batch size and learning rate combination. Additionally, a change to the LOP2 training strategy were employed with change to the ILF network architecture in EE1-2.3.

In this contribution, results of LOP2 Stage 3 training with modified strategy (batch size 64 and learning rate 0.0008) are reported. Assessment of the validation costs indicates that a new training strategy allows LOP2 network to be trained faster than the anchor strategy, however, it exhibits larger cost value fluctuations.

Inference testing with NNVC7.1 shows that comparing to LOP2 anchor, the float model trained with a new training strategy provides BD-rate change of {-0.20%, 0.18%, -0.77%} in RA and {-0.02%, -0.08%, - 0.85%} in AI configurations. A quantized model provides BD-rate change of {-0.09%, 0.12%, -0.73%} and {0.09%, - 0.09%, -0.84%}, in RA and AI, respectively.

It is proposed to consider the proposed training strategy for ILF training in LOP category.

It was commented that, as a saturation is observed around epoch 30 (and validation loss in chroma even increasing), it should be considered to switch/drop the learning rate beyond that point (currently, the learnimg rate is dropped around epoch 50, where the validation loss starts becoming lower).

Study in EE as joint effort, i.e. several parties to run training independently.

[JVET-AG0179](https://jvet-experts.org/doc_end_user/current_document.php?id=13735) EE1-1.1-related: HOP filter complexity alignment with wider activation [Y. Li, C. Lin, J. Li, K. Zhang, L. Zhang (Bytedance)]

EE1-1.1 investigates a refined joint filter architecture for HOP by combining key components from several JVET-AF contributions including JVET-AF0102, JVET-AF0103, JVET-AF0182. EE1-1.1 tries to align the complexity of the joint filter architecture with the HOP.2 filter in NNVC-7.1 by increasing the number of residual blocks. This contribution proposes to achieve the complexity alignment in the other way around, i.e. by increasing the number of feature maps.

BD-rate changes on top of NNNC-7.1 HOP anchor (NN intra + HOP.2) are reportedly summarized as below:

RA: -0.16%, -0.38%, -0.69%, AI: -0.21%, -0.40%, -0.12%

kMAC/pixel: 442

It was commented that an advantage compared to EE1-1.1 is the even number of residual blocks, which is beneficial for the alternating operation. Performance is similar to EE1-1.1 that was adopted in the current meeting.

Study in EE. The complexity kMAC/pix should be equivalent to EE1-1.1 (466 kMAC/pix), to identify if the modified architecture has benefit.

### Improvements of NNVC beyond EE1 (4)

Contributions in this area were discussed at 0035–0130 on Saturday 20 Jan. 2024, and at 0830-0905 on Monday 22 Jan. 2024 (chaired by JRO).

[JVET-AG0057](https://jvet-experts.org/doc_end_user/current_document.php?id=13613) [AHG11] Study on lower-complexity NNLF [Z. Xie, Y. Yu, H. Yu, D. Wang (OPPO)]

This contribution reports the experimental results of a lower-complexity NNLF, with computational complexity about 5.0 KMAC/pixel. The training and inference processes are identical to the NNVC-7.1 LOP.2. Compared with VTM anchor, the simulation results of the proposed lower-complexity NNLF are summarized as follows:

Compared with VTM-anchor:

AI: -3.32% (Y) , -4.19% (U), -3.42% (V), 81% (EncT), 1631% (DecT)

RA: -2.89% (Y), -3.83% (U), -2.29% (V), 104% (EncT), 3473% (DecT)

Encoder run time is likely unreliable.

Study in EE along with JVET-AG0155 and JVET-AF0206). Integerization also to be studied. It was suggested that for ultra low operation, it would be interesting to also investigate more aggressive integerization, such as 8/10 bits (clarify if that is possible with SADL).

[JVET-AG0069](https://jvet-experts.org/doc_end_user/current_document.php?id=13625) AhG11: LOP with inputs transformed [D. Liu, J. Ström, M. Damghanian, P. Wennersten (Ericsson)]

This contribution proposes to apply a 2x2 DCT-II transform and reshaping to the inputs of the LOP model and apply an inverse DCT-II and inverse reshaping to the output of the LOP model. An input of spatial size 144x144 is transformed and reshaped into 72x72x4 if it is luma (needed in Rec, Pred, BS, IPB). If chroma is included in the input, chroma samples are reshaped to 72x72x2 without transform and is concatenated with the transformed luma. If the input is QP, it is directly reshaped to a size of 72x72x1. The complexity in kMACs is reduced, since the spatial resolution inside the network is reduced by a factor of four while the final output still has the same number of pixels. To meet the LOP complexity, it is proposed to further increase the number of backbone blocks and channels. Regeneration of training data and training of the proposed model is needed.

It is stated that the proposed model has a complexity of 16.9 kMAC/pixel and 0.2M parameters (LOP.2: 17 kMAC/pixel 0.05M parameters). The performance (including NNIntra) compared to the NNVC-7.1 anchor (NNIntra+LOP) is reported to be:

Float: RA {-0.48%, -3.94%, -3.51%}, AI {-0.17%, -4.16%, -4.55%}, LDB {-0.21%, -0.03%, 0.53%}

Int16: RA {-0.44%, -3.37%, -3.42%}, AI {-0.18%, -4.07%, -4.57%}, LDB {-0.03%, 0.61%, 1.75%}

It is proposed to study the input-transformed model in the next EE.

The four channels are combined in the first layer, such that in the subsequent channels, 2x subsampling hor/ver is given. To improve the performance, rather than running with less parameters, the complexity is brought to a similar level as LOP2, increasing the number of layers. This also increases latency.

It was suggested to consider training of the transform rather than using fixed DCT

It was commented that a different dataset was used for training in RA, which makes a difficult to compare against LOP2.

It was commented that effectively the receptive field is enlarged due to the subsampling

It was asked why the HOP model (and even different complexity) was used for training stages 1 and 2? LOP was not available yet.

It was commeted that rather than increasing the number of layers, the number of channels might be increased to keep the same complexity as LOP2.

Investigate in EE.

It is important that the training would be comparable with the procedure used for current LOP.

BoG (E. Alshina, F. Galpin):

* Design of next EE
* Call for training material
* Review 5.1.5

[JVET-AG0258](https://jvet-experts.org/doc_end_user/current_document.php?id=13831) Crosscheck of JVET-AG0069 (AhG11: LOP with inputs transformed) [Y. Li (Qualcomm)] [late]

[JVET-AG0114](https://jvet-experts.org/doc_end_user/current_document.php?id=13670) AHG11: On new input and backbone enhancement of model for super-resolution [J. Ye, X. Li, Y. Zhu, Q. Liu (HUST), C. Zhou, M. Rafie, Z. Lv (vivo)]

In this contribution, 2 methods are proposed to further improve the coding performance of the super-resolution (SR) model. The SR network is designed with reference to the unified model HOP of in-loop filter, which was defined in JVET-AD0380 [1]. First, the SR models in NNVC use reconstruction and multiple side information including prediction, slice QP, base QP and slice type as the inputs of the network. Reference picture resampling (RPR) picture is added as a new input to the neural network. Second, to preprocess the shallow features and improve the diversity of the features fed into the backbone, enhancement modules are added before the backbone part.

Considering both flexibility and performance in the use of the model, experiments are conducted under two conditions.

1. A unified model is trained using both I- and B-frame training data, and the SR of both I- and B-frames in the inference phase is performed by this model. This model, which is applied to both I-frame B-frame, is called ***IB\_model***.
2. Ttwo separated models are trained using I- and B-frame training data separately, and SR of I- and B-frames in the inference phase is performed by the respective models. These are called ***I\_model*** and ***B\_model***.

For ***IB\_model***, compared with NNVC-7.0-vtm, the results with SADL implementation in float are shown {-5.88%, -3.99%, -2.82%} and {-3.48%, -10.39%, -7.20%} BD-rate savings under AI and RA, respectively.

For ***I\_model*** and ***B\_model***, compared with NNVC-7.0-vtm, the results with SADL implementation in float are shown {-5.96%, -4.15%, -3.07%} and {-3.80%, -10.45%, -7.02%} BD-rate savings under AI and RA, respectively.

The complexity of IB\_model is 452kMAC/pixel compared to 469kMAC/pixel of the SR model in NNVC-7.0.

The complexity of I\_model and B\_model are 447 and 452kMAC/pixel, respectively, compared to 469kMAC/pixel of the SR model in NNVC-7.0.

No results comparing against NNVC7.1 anchor. However, interesting gain relative to the SR approach that was investigated in EE1.

Study in next EE1. Compare against NNVC anchor (with other tools enabled). Also an additional test using an NNVC anchor with conventional RPR enabled should be performed.

Clarify the drop occurring in Daylight Road – possibly wrong encoder decision?

[JVET-AG0129](https://jvet-experts.org/doc_end_user/current_document.php?id=13685) AHG11: Unified CNN-based super resolution [C. Zhou, Z. Lv (vivo)]

This contribution proposes a unified CNN-based super-resolution method, which includes two aspects: 1) using the RPR mechanism to generate high-resolution reconstructions to replace low-resolution reconstructions as input, and 2) introducing scaling ratio as an input to support arbitrary scaling factors. The average BD rate savings for Class A1 and Class A2 are reported as follows:

Aspect #1:

RPR2x: AI {-3.47%, 1.64%, 3.31%} RA {-3.71%, 1.58%%, 1.80%}

RPR1.5x: AI {-3.63%, 0.91%, 1.92%} RA {-3.89%, 2.74%, 3.05%}

Aspect #2:

RPR2x: AI {-3.47%, 2.49%, 4.38%} RA {-3.73%, 2.09%, 2.84%}

RPR1.5x: AI {-3.56%, 1.89%, 2.75%} RA {-3.99%, 3.84%, 4.77%}

No results in other classes.

Postprocessing of RPR with LOP loop filter after upsampling. When combined with NNVC anchor, the LOP LF would be applied in low resolution loop, then RPR, and another LOP LF (as postprocessing) at full resolution.

Study in next EE.

### SADL implementation (3)

Contributions in this area were discussed at XXXX–XXXX on XXday XX Jan. 2024 (chaired by XXX).

Reviewed in BoG JVET-AG0325 – include abstracts and disposition from report.

[JVET-AG0109](https://jvet-experts.org/doc_end_user/current_document.php?id=13665) AhG14 SADL update [F. Galpin, T. Dumas, P. Bordes, E. François (InterDigital)]

[JVET-AG0115](https://jvet-experts.org/doc_end_user/current_document.php?id=13671) AHG14: The extension of SADL library [X. Li, J. Ye, Y. Zhu, Q. Liu (HUST)]

[JVET-AG0219](https://jvet-experts.org/doc_end_user/current_document.php?id=13792) AHG14: The extension of SADL library [Y. Cai, W. Bao, Z. Chen (Wuhan Univ.)] [late]

## AHG6/AHG12: Enhanced compression beyond VVC capability (87+1)

### Summary and BoG reports (1)

Contributions in this area were discussed at 2330–0120 on Wednesday 17 Jan. 2024, at 0500–0710 and at 2100-XXXX on Thursday 18 Jan. 2024 (chaired by JRO).

[JVET-AG0024](https://jvet-experts.org/doc_end_user/current_document.php?id=13791) EE2: Summary report of exploration experiment on enhanced compression beyond VVC capability [V. Seregin, J. Chen, R. Chernyak, K. Naser, J. Ström, F. Wang, M. Winken, X. Xiu, K. Zhang (EE coordinators)]

This document provides a summary report of Exploration Experiment on Enhanced Compression beyond VVC capability. The tests are categorized as intra prediction, inter prediction, transform and coefficient coding, in-loop filtering, and entropy coding.

The software basis for this EE is ECM-10.0, released at <https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tags/ECM-11.0>. ECM-11.0 is used as an anchor in the tests.

Software for EE tests is released in the corresponding branches at <https://vcgit.hhi.fraunhofer.de/ecm/jvet-af-ee2/ECM/-/branches>.

Test results can be found in input JVET contributions, cross-check results are uploaded to <https://vcgit.hhi.fraunhofer.de/ecm/jvet-af-ee2/simulation-results> if cross-check reports are not submitted as they are optional for EE tests.

**List** of **tests**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Tests** | **Tester** | **Cross-checker** |
| **1 Intra prediction** | | | |
| 1.1a | Decoder derived CCP mode | Y.-J. Chang (Qualcomm)  [JVET-AG0154](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0154-v1.zip) | H.-J. Jhu  (Kwai)  [JVET-AG0269](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0269-v1.zip) |
| 1.1b | Test 1.1a with decoder derived CCP fusion modes | Y.-J. Chang (Qualcomm)  [JVET-AG0154](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0154-v1.zip) | J. Lainema  (Nokia)  [JVET-AG0278](https://jvet-experts.org/doc_end_user/current_document.php?id=13851) |
| 1.2 | IntraTMP with merge candidates | K. Naser (InterDigital)  [JVET-AG0151](https://jvet-experts.org/doc_end_user/current_document.php?id=13707) | I.Zupancic (Nokia)  [JVET-AG0245](https://jvet-experts.org/doc_end_user/current_document.php?id=13818) |
| 1.3 | SGPM with IntraTMP and IBC | K. Naser  (InterDigital)  [JVET-AG0152](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0152-v1.zip) | I.Zupancic (Nokia)  [JVET-AG0245](https://jvet-experts.org/doc_end_user/current_document.php?id=13818) |
| 1.4 | IntraTMP extension to DIMD | K. Naser  (InterDigital)  [JVET-AG0146](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0146-v1.zip) | K. Cui  (Qualcomm)  [JVET-AG0170](https://jvet-experts.org/doc_end_user/current_document.php?id=13726) |
| 1.5 | IntraTMP extension to LIC | F. Le Léannec  (InterDigital)  [JVET-AG0136](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0136-v1.zip) | K. Cui  (Qualcomm)  [JVET-AG0171](https://jvet-experts.org/doc_end_user/current_document.php?id=13727) |
| 1.6 | Test 1.4 + Test 1.5 | F. Le Léannec  (InterDigital)  [JVET-AG0137](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0137-v2.zip) | M. Abdoli  (Xiaomi)  [JVET-AG0255](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0255-v1.zip) |
| 1.7a | Test 1.6 + Test 1.3 | K. Naser  (InterDigital)  [JVET-AG0137](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0137-v2.zip) | F. Wang,  L. Zhang  (OPPO)  [JVET-AG0274](https://jvet-experts.org/doc_end_user/current_document.php?id=13847) |
| 1.7b | Test 1.7a + Test 1.2 | F. Le Léannec (InterDigital)  [JVET-AG0137](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0137-v2.zip) | F. Wang,  L. Zhang  (OPPO)  [JVET-AG0274](https://jvet-experts.org/doc_end_user/current_document.php?id=13847) |
| 1.8a | The length of the auto-relocated BVP trace path is 1 (i.e, n=1) | N. Zhang  (Bytedance)  [JVET-AG0091](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0091-v1.zip) | D. Ruiz Coll, (Ofinno)  [JVET-AG0134](https://jvet-experts.org/doc_end_user/current_document.php?id=13690)  Y. Kidani,  (KDDI)  [JVET-AG0225](https://jvet-experts.org/doc_end_user/current_document.php?id=13798) |
| 1.8b | The length of the auto-relocated BVP trace path is 2 (i.e, n=2) | N. Zhang  (Bytedance)  [JVET-AG0091](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0091-v1.zip) | D. Ruiz Coll, (Ofinno)  [JVET-AG0134](https://jvet-experts.org/doc_end_user/current_document.php?id=13690)  Y. Kidani,  (KDDI)  [JVET-AG0225](https://jvet-experts.org/doc_end_user/current_document.php?id=13798) |
| 1.8c | No constraint for the length of the auto-relocated BVP trace path | N. Zhang  (Bytedance)  [JVET-AG0091](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0091-v1.zip) | D. Ruiz Coll, (Ofinno)  [JVET-AG0134](https://jvet-experts.org/doc_end_user/current_document.php?id=13690)  Y. Kidani,  (KDDI)  [JVET-AG0225](https://jvet-experts.org/doc_end_user/current_document.php?id=13798) |
| 1.9 | Intra TMP fusion probing | J.-L. Lin  (Qualcomm)  [JVET-AG0118](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0118-v1.zip) | Z. Deng  (Bytedance)  JVET-AG0236 |
| 1.10 | Bilateral filtering for intra prediction | W. Yin  (Bytedance)  [JVET-AG0123](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0123-v1.zip) | X. Li  (Alibaba)  JVET-AG0248 |
| 1.11a | IntraTMP search area extension with sampling factor proportional to CU distance | D. Ruiz Coll  (Ofinno)  [JVET-AG0131](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0131-v1.zip) | G. Verba  (Qualcomm)  [JVET-AG0229](https://jvet-experts.org/doc_end_user/current_document.php?id=13802) |
| 1.11b | IntraTMP search area extension with sampling factor proportional to search area dimension | K. Naser  (InterDigital)  [JVET-AG0131](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0131-v1.zip) | G. Verba  (Qualcomm)  [JVET-AG0229](https://jvet-experts.org/doc_end_user/current_document.php?id=13802) |
| 1.11c | Test 1.11a + Test 1.11b | D. Ruiz Coll  (Ofinno)  K. Naser  (InterDigital)  [JVET-AG0131](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0131-v1.zip) | G. Verba  (Qualcomm)  [JVET-AG0229](https://jvet-experts.org/doc_end_user/current_document.php?id=13802) |
| 1.12a | IBC with extended reference area up to the picture’s upper side boundary | Y. Kidani  (KDDI)  [JVET-AG0072](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0072-v1.zip) | N. Zhang  (Bytedance)  [JVET-AG0206](https://jvet-experts.org/doc_end_user/current_document.php?id=13762) |
| 1.12b | IBC with four CTU rows instead of two CTU rows in HD resolution or less | Y. Kidani  (KDDI)  [JVET-AG0072](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0072-v1.zip) | N. Zhang  (Bytedance)  [JVET-AG0206](https://jvet-experts.org/doc_end_user/current_document.php?id=13762) |
| 1.13a | Test 1.11c + Test 1.12a | Y. Kidani  (KDDI)  D. Ruiz Coll  (Ofinno)  K. Naser  (InterDigital)  [JVET-AG0199](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0199-v1.zip) | N. Zhang  (Bytedance)  [JVET-AG0207](https://jvet-experts.org/doc_end_user/current_document.php?id=13763) |
| 1.13b | Test 1.11c + Test 1.12b | Y. Kidani  (KDDI)  D. Ruiz Coll  (Ofinno)  K. Naser  (InterDigital)  [JVET-AG0199](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0199-v1.zip) | N. Zhang  (Bytedance)  [JVET-AG0207](https://jvet-experts.org/doc_end_user/current_document.php?id=13763) |
| 1.14a | Encoder run-time reduction methods for extrapolation filter-based intra prediction mode only | L. Xu (OPPO)  [JVET-AG0058](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0058-v1.zip) | X. Li  (Alibaba)  [JVET-AG0250](https://jvet-experts.org/doc_end_user/current_document.php?id=13823)  H.-J. Jhu  (Kwai)  [JVET-AG0265](https://jvet-experts.org/doc_end_user/current_document.php?id=13838) |
| 1.14b | Encoder run-time reduction methods for extrapolation filter-based intra prediction and its merge mode | L. Xu (OPPO)  [JVET-AG0058](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0058-v1.zip) | X. Li  (Alibaba)  [JVET-AG0250](https://jvet-experts.org/doc_end_user/current_document.php?id=13823)  H.-J. Jhu  (Kwai)  [JVET-AG0265](https://jvet-experts.org/doc_end_user/current_document.php?id=13838) |
| 1.15a | CCP merge fusion | H. Huang  (OPPO)  Z. Deng  (Bytedance)  [JVET-AG0059](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0059-v1.zip) | X. Li  (Alibaba)  [JVET-AG0249](https://jvet-experts.org/doc_end_user/current_document.php?id=13822) |
| 1.15b | Inheriting LB-CCP flag in CCP merge mode | H. Huang  (OPPO)  Z. Deng  (Bytedance)  [JVET-AG0059](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0059-v1.zip) | X. Li  (Alibaba)  [JVET-AG0249](https://jvet-experts.org/doc_end_user/current_document.php?id=13822) |
| 1.15c | Test 1.15a + Test 1.15b | H. Huang  (OPPO)  Z. Deng  (Bytedance)  [JVET-AG0059](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0059-v1.zip) | X. Li  (Alibaba)  [JVET-AG0249](https://jvet-experts.org/doc_end_user/current_document.php?id=13822) |
| 1.16 | Slope adjustment for IBC LIC | C. Ma  (Kwai)  [JVET-AG0103](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0103-v1.zip) | Y. Wang  (Bytedance)  [JVET-AG0221](https://jvet-experts.org/doc_end_user/current_document.php?id=13794) |
| 1.17a | IBC GPM with block vector difference | C. Ma  (Kwai)  [JVET-AG0104](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0104-v1.zip) | Y. Wang  (Bytedance)  [JVET-AG0222](https://jvet-experts.org/doc_end_user/current_document.php?id=13795) |
| 1.17b | IBC GPM with split mode reordering | C. Ma  (Kwai)  [JVET-AG0104](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0104-v1.zip) | Y. Wang  (Bytedance)  [JVET-AG0222](https://jvet-experts.org/doc_end_user/current_document.php?id=13795) |
| 1.17c | Test 1.17a + Test 1.17b | C. Ma  (Kwai)  [JVET-AG0104](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0104-v1.zip) | Y. Wang  (Bytedance)  [JVET-AG0222](https://jvet-experts.org/doc_end_user/current_document.php?id=13795) |
| 1.18a | DIMD mode derivation from spatial blocks | J. Huo, J. Fan  (Xidian Univ.)  M. Li  (OPPO)  [JVET-AG0076](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0076-v1.zip) |  |
| 1.18b | DIMD mode derivation with reduced complexity | J. Huo, J. Fan  (Xidian Univ.)  M. Li  (OPPO)  [JVET-AG0076](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0076-v1.zip) |  |
| 1.19 | IBC-LIC model merge mode | L. Zhang  (OPPO) | Y. Wang  (Bytedance)  [JVET-AG0220](https://jvet-experts.org/doc_end_user/current_document.php?id=13793) |
| 1.20a | TIMD fusion with non-angular predictor | P. Andrivon (Ofinno)  [JVET-AG0092](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0092-v1.zip) | M. Abdoli  (Xiaomi) |
| 1.20b | Test 1.20a + TIMD sample-based fusion | P. Andrivon (Ofinno)  [JVET-AG0092](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0092-v1.zip) | S. Blasi  (Nokia)  [JVET-AG0132](https://jvet-experts.org/doc_end_user/current_document.php?id=13688) |
| 1.21 | TIMD fusion reference line determination | C. Zhou (vivo)  [JVET-AG0128](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0128-v1.zip) | J. Chen  (Alibaba) |
| 1.22a | Test 1.20a + Test 1.21 | P. Andrivon (Ofinno)  C. Zhou  (vivo)  [JVET-AG0093](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0093-v1.zip) | M. Abdoli  (Xiaomi) |
| 1.22b | Test 1.20b + Test 1.21 | P. Andrivon (Ofinno)  C. Zhou (vivo)  [JVET-AG0094](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0094-v1.zip) | S. Blasi  (Nokia)  [JVET-AG0133](https://jvet-experts.org/doc_end_user/current_document.php?id=13689) |
| 1.22c | Test 1.20b + Test 1.21 + Test 1.4 | P. Andrivon (Ofinno)  C. Zhou (vivo)  K. Naser  (InterDigital)  [JVET-AG0095](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0095-v1.zip) | H. Qin  (TCL)  [JVET-AG0127](https://jvet-experts.org/doc_end_user/current_document.php?id=13683) |

***Intra prediction***

**Test 1.1: Decoder derived CCP mode (**[**JVET-AG0154**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0154-v1.zip)**)**

In this method, a candidate list of cross-component prediction (CCP) modes is constructed, and to select the best candidate from the list a template cost is calculated to compare the reconstructed samples and the prediction values generated by the evaluated CCP mode. The template is shown in the next figure.

A blue and black rectangle

Description automatically generated

The template adjacent to the current chroma CU

The CCP mode list is constructed from the already existed in ECM modes by single model CCLM, single model CCCM, multi-model CCCM, single model GLCCCM, single model CCCM applied with LBCCP, and multi-model CCCM applied with LBCCP.

In the second aspect of the method, various decoder-derived CCP fusion candidates are added. A fusion candidate is the combination of two CCP modes selected from the existing CCP mode lists reordered by template costs.

Mode flag and a fusion flag are signalled to indicate the mode usage.

Test 1.1a: Decoder derived CCP mode

Test 1.1b: Test 1.1a with decoder derived CCP fusion mode

**Test 1.2: IntraTMP with merge candidates (**[**JVET-AG0151**](https://jvet-experts.org/doc_end_user/current_document.php?id=13707)**)**

In the test, new candidates, consisting of block vectors from neighbouring PUs coded in IBC or IntraTMP mode, are added into IntraTMP search process. There are up to 25 local and non-local PUs that may be checked during the selection of the neighbouring candidates, those candidates are evaluated together with other candidates, and 30 best candidates (same as in ECM) are selected for further refinement.

The neighbouring candidates may come from outside of IntraTMP search area and such outside candidates are prioritized.

**Test 1.3: SGPM with IntraTMP and IBC (**[**JVET-AG0152**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0152-v1.zip)**)**

In the test, spatial GPM (SGPM) mode is extended to additionally consider BVs from IBC and IntraTMP to form a predictor. In the extended SGPM, the predictor may be IBC, IntraTMP or regular intra prediction.

BV are selected as follows:

1. Obtain block vectors of all merge candidates
2. Select the best block vector according to SATD template cost
3. Test up to 6 best block vectors (if available) inside SGPM candidate list construction.



Identifiers intra\_pred\_0 and intra\_pred\_1 correspond to be regular or BV based prediction.

**Test 1.4: IntraTMP extension to DIMD (**[**JVET-AG0146**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0146-v1.zip)**)**

The method consists in adaptively selecting between planar or block vector based prediction obtained from IntraTMP or IBC mode of neighboring blocks for the blending with an angular intra prediction in DIMD (up to five predictors and the planar mode predictor with the weights derived from the histogram of gradients can be used). The selection is based on the SAD template cost, where the planar mode is compared to all block vector based prediction. The mode with the minimum cost is selected.

**Test 1.5: IntraTMP extension to LIC (**[**JVET-AG0136**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0136-v1.zip)**)**

In the test, LIC usage is extended to IntraTMP mode. A CU-level flag is signalled to indicate the use of LIC. For screen content coding, the IntraTMP LIC mode is extended to support top-only and left-only templates to compute LIC model parameters, while for camera-captured coding, only the top-left template is employed. MMLM is also supported for IntraTMP LIC mode in screen content coding, similar to IBC LIC.

When LIC is used for a given CU, the IntraTMP search process employs MRSAD rather than SAD distortion function to search for block vector.

A restricted version of IntraTMP LIC is also tested. The restriction consists in a reduced IntraTMP search process for both LIC and non-LIC IntraTMP coding units. That is, a SAD-based search is performed for non-LIC blocks and a MRSAD-based search is performed for IntraTMP CUs coded in LIC mode.

To limit the amount of BV search process, the sampling of the IntraTMP search area is increased from 3 to 4.

**Tests 1.6-1.7: Combination of IntraTMP related tests (**[**JVET-AG0137**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0137-v2.zip)**)**

Test 1.6: Test 1.4 + Test 1.5 (restricted)

Test 1.7a: Test 1.3 + Test 1.4 + Test 1.5 (restricted)

Test 1.7b: Test 1.2 + Test 1.3 + Test 1.4 + Test 1.5 (restricted)

Test 1.7c: Test 1.2 + Test 1.3 + Test 1.4

**Test 1.8: Auto-relocated block vector prediction (**[**JVET-AF0079**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0079-v1.zip)**)**

In the test, auto-relocated BV is introduced to IBC merge and AMVP candidate list construction. As shown in the next figure, a guiding block vector BV0,1 associated with the current block B0 points to a reference block B1. If B1 has a BV denoted as BV1,2 pointing to a reference block B2, then BV0,2,given by BV0,2 = BV0,1 +BV1,2, is defined as the auto-relocated BVP, guided by BV0,1. Similarly, BV0,n+1 can be derived by

BV0,n+1 =BV0,n+BVn,n+1 = BV0,1+BV1,2 +…+BVn-1,n +BVn,n+1.

A black background with blue lines and squares

Description automatically generated

When deriving BVn,n+1 guided by BV0,n, all five positions shown in the next figure including top-left, top-right, center, bottom-left, and bottom-right positions of Bn are checked to find BVn,n+1.

Auto-relocated BV candidates are inserted after the HBVP candidates. The IBC merge and AMVP candidate list sizes are kept unchanged.

A screen shot of a computer

Description automatically generated

Three tests are conducted varying the length of the trace path, n=1, n=2, and unrestricted.

Test 1.8a: Auto-relocated BV with trace path 1.

Test 1.8b: Auto-relocated BV with trace path 2.

Test 1.8c: Auto-relocated BV with unrestricted trace path.

**Test 1.9: IntraTMP fusion probing (**[**JVET-AG0118**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0118-v1.zip)**)**

In ECM, IntraTMP candidate list is constructed from candidates derived by L-shape, left or above template based on the template matching cost which is calculated as the SAD between the template of the reference block and the template of current block. For IntraTMP fusion mode, multiple IntraTMP candidates are linearly combined according to the fusion weights derived from the template matching costs or derived by the MSE minimization method. A flag intra\_tmp\_fusion\_weight\_typeis signalled to indicate which weight derivation method is used and an index intra\_tmp\_fusion\_idx is signalled to indicate which candidate set is selected for IntraTMP fusion.

In the tested method, a fusion candidate with the minimum probing cost is selected from a fusion candidate list without signalling the intra\_tmp\_fusion\_weight\_typeand intra\_tmp\_fusion\_idx. The probing cost is derived as the SAD between the samples in the probing line of the fused template and the current block’s template. The fusion weights derivation is the same as in ECM-11.0, but the samples in probing line are excluded.



A flag is signalled to indicate the IntraTMP fusion probing mode.

**Test 1.10: Bilateral filtering for intra prediction (**[**JVET-AG0123**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0123-v1.zip)**)**

In this test, the generated intra prediction samples are further filtered by bilateral filter to reduce the noise level. The filter is similar to BIF used in the loop-filtering stage. The filter length is kept as 2.

**Test 1.11: IntraTMP search range extension with adaptive sampling (**[**JVET-AG0131**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0131-v1.zip)**)**

In ECM, the IBC search region comprises the reconstructed samples of a fixed number of CTUs based on the CTU dimension. This region encompasses all left neighbor CTUs within the same row as the current CTU and extends to two CTU rows above the current CTU.



In the test, IntraTMP search region is extended to align with the one of IBC. To cover the whole IBC reference region, the number of IntraTMP regions is increased by one compared to ECM-11, and exclusively, the boundaries of the farthest original regions (R1, R2, and R3) have been modified to cover the whole IBC search region, as it is depicted in next figure.



The ‘*regular search area’* maintains the exact dimensions as in ECM-11 constrained by the *SearchRangeWidth* and *SearchRangeHeight* parameters, and the same sample rate of 3.

The searching process in the further away regions (R1 and R7) is subject to pruning if the normalized TM cost of the last BV candidate in the sparse list is less than a threshold, which depends on the current block's horizontal and vertical size (W, H). Furthermore, the search regions R1 and R6 are scanned in the inverse direction, from left to right, checking before the most probable candidates, and the scanning pattern is shifted by one pixel per row. Lastly, for block sizes of 64 samples, the factor parameter ‘a’ is set to 4, preventing the regular search region from exceeding the IBC search region.

Two methods are tested adapting the sampling rate beyond the regular search area to reduce the complexity of the search extension.

In Test 1.11a, sampling rate is adjusted based on stepwise function stored as lookup table based on the block distance ranging from 3 within the regular search region (a\*W, a\*H) to a maximum sampling factor of 80.

In Test 1.11b, sampling rate is based on IntraTMP regions dimension, where a new subsampling factor is determined for each region () according to the region dimensions as follows:

where F is the constant that controls the gain/complexity trade-off. The new sampling rate factors are only applied at the extended region locations.

Test 1.11c is a combination of Test 1.11a and Test 1.11b:

* The multifactor parameter 'a' for block sizes of 8 and 4 pixels is unified to 5.
* The sampling factor rate for both methods is calculated as described in Test 1.11b.
* The most significant sampling factor is selected for the current location:

**Test 1.12: IBC with upward-extended reference area (**[**JVET-AG0072**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0072-v1.zip)**)**

In the ECM, the current IBC reference area is shown in the next figure.



In Test 1.12a, IBC reference area is extended to the whole causal area of a picture (to the top picture boundary).

In Test 1.12b, IBC search area is increased from 2 to 4 CTU rows above the current CTU for HD resolution or less.

**Test 1.13: Combination of IBC and IntraTMP search area related tests (**[**JVET-AG0199**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0199-v1.zip)**)**

Test 1.13a: Test 1.11c + Test 1.12a.

Test 1.13b: Test 1.11c + Test 1.12b.

**Test 1.14: Extrapolation filter-based intra prediction mode (**[**JVET-AG0058**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0058-v1.zip)**)**

The extrapolation filter-based intra prediction mode consists of three steps:

1. The extrapolation filter coefficients are derived from a neighbouring reconstructed area of the current block or filter shape and coefficients are inherited from a previous block coded with this mode. For the latter, a mode flag is signalled, and candidate list is constructed using spatial and non-adjacent blocks coded with this mode, initial list of 12 candidates is reduced to 6 after reordering based on SAD cost measured on the L-shape template of size 1. An index is further signalled to identify a candidate from the list.
2. The extrapolation process generates predicted signals from the top-left to bottom-right corner within the current block.
3. An intra prediction angle using DIMD process is derived by analyzing the gradient of the predicted block, and then the corresponding intra mode is used to select MTS, NSPT, and LFNST kernels for transform.

The mode is restricted to blocks with sizes not greater than 32x32 and luma component only.

Three filer shapes are used in this method, a choice of reconstructed area and filter shape shown in the next figure is signalled.

A black and white grid

Description automatically generated

A black screen with white squares

Description automatically generated

The selected filter shape moves in the selected reconstructed area with a one-sample step either horizontally or vertically to collect input samples and output samples, then the filter coefficients are derived using CCCM solver.

The merge mode is also introduced, where the filter shape and the filter coefficients are inherited from the previous decoded blocks that are coded with the tested extrapolation filter-based intra prediction mode or the merge mode based on this mode. In the merge mode, the positions and inclusion order of the spatial adjacent, temporal, non-adjacent, shifted temporal, and history candidates are the same as those defined in ECM-11.0 for the CCP merge prediction candidates. In addition, the merge candidates are reordered by comparing the SAD cost on an L-shape template with column width and row height equal to 1. In the SAD calculation, predictions of the template area by the mode filters are generated only from reconstructed (neighbouring and template) samples, allowing the filters to be applied in parallel rather than sequentially.

In the current block, the recursive predictor is derived from the top-left to the bottom-right position by a diagonal prediction order as shown in the next figure, where the predicted results of the previous diagonal are used.

A diagram of a graph

Description automatically generated

The predicted samples are calculated as follows,

where is the predicted value at (x, y) in the current block, is the coefficient of the selected filter, the index of the coefficients is from 0 to 14, is a reconstructed or a predicted value used for the current position’s prediction, and are the position offsets to the current position along x and y directions, respectively.

At encoder side, SATD is used to compare the costs for the added modes to update the list for full RDO, the number of total RDs is not increased. Additionally, the encoder will conditionally reduce one mode with the worst SATD cost in the full RDO cost when extrapolation-based prediction mode is already included in the full RDO list.

Test 1.14a: Encoder run-time reduction methods for extrapolation filter-based intra prediction mode only

Test 1.14b: Encoder run-time reduction methods for extrapolation filter-based intra prediction and its merge mode

**Test 1.15: Enhancements on CCP merge for chroma intra coding (**[**JVET-AG0059**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0059-v1.zip)**)**

In the CCP merge mode of ECM, a candidate list is created, which includes CCP models collected from the previous blocks coded by CCLM, MMLM, CCCM, GLM, chroma fusion, and CCP merge modes. Then one candidate in the list is selected as the CCP model for the current block.

In the test, two aspects are evaluated.

Aspect #1: CCP-merge fusion method is introduced where a fused prediction is generated by a weighted sum of the CCP-merge prediction and either the MM-CCCM prediction or the DIMD prediction. A CCP-merge fusion flag is signalled conditionally under the CCP-merge flag, to indicate whether the fusion mode is applied. A CCP-merge fusion type flag is further signalled if the CCP-merge fusion flag is true, to indicate whether the MM-CCCM prediction or the DIMD prediction is selected and fused with the CCP-merge prediction.

Aspect #2: The local-boosting CCP mode flag is inherited from a CCP candidate in the CCP merge candidate list.

Test 1.15a: CCP merge fusion

Test 1.15b: Inheriting LB-CCP flag in CCP merge mode

Test 1.15c: Test 1.15a + Test 1.15b

**Test 1.16: IBC LIC with slope adjustment (**[**JVET-AG0103**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0103-v1.zip)**)**

In this test, the slope adjustment of IBC LIC is introduced, where an offset parameter chosen from the set {-1/2, -1/4, -1/8, -1/16, 1/16, 1/8, 1/4, 1/2} is signalled to update the slope parameter of IBC LIC. The proposed method is only applied in IBC AMVP mode and luma component. In addition, the slope adjustment is not applied to the CUs that apply multi-mode IBC LIC.

**Test 1.17:  IBC GPM with block vector difference and split mode reordering (**[**JVET-AG0104**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0104-v1.zip)**)**

In Test 1.17a, IBC GPM with block vector difference is introduced, where one additional BVD, which is selected from the set {1/4-pel, 1/2-pel, 1-pel, 2-pel, 4-pel, 8-pel, 16-pel, 32-pel}, is allowed to be added on top of the BV of one IBC GPM partition which is based on regular IBC merge. A flag is signalled for each IBC partition to indicate whether additional BVD is applied. When the flag is true, the corresponding BVD index is further signalled.

In Test 1.17b, the split modes of IBC GPM with template matching are reordered, the reordering is only applied to the first set of IBC GPM split modes.

Test 1.17c: Test 1.17a + Test 1.17b.

**Test 1.18: DIMD mode derivation from spatial blocks (**[**JVET-AG0076**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0076-v1.zip)**)**

In this test, DIMD mode is derived by spatial blocks including non-adjacent spatial candidates. When neighbouring blocks are coded with DIMD mode, the DIMD histograms are combined to form a new DIMD merge histogram for the current block. DIMD merge modes and weights are computed based on this merged histogram.

As shown in the below figure, the distances between non-adjacent candidates and current block are defined based on the width and height of current coding block. When using DIMD merge, the DIMD information extracted from neighbouring blocks and non-adjacent spatial blocks is used to compute the intra prediction for the current block.



In Test 1.18a, when more than one neighbouring block encoded with DIMD or DIMD merge are available, all the spatial blocks are used to generate the DIMD merge mode.

In Test 1.18b, when more than one neighbouring block encoded with DIMD or DIMD merge are available, only the non-adjacent spatial candidates are used to derive the DIMD merge mode.

**Test 1.19: IBC-LIC model merge mode (**[**JVET-AG0060**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0060-v1.zip)**)**

In the test, IBC-LIC model parameters are inherited from spatial adjacent and non-adjacent neighbours, history candidates, or default models as follows.

1. Construct a model candidate list which consists of model parameters from spatial adjacent and non-adjacent neighbors, history candidates, and default models. The size of the candidate list is 12.

If the list is not full, a default model and the scaled models are then added to the list. To avoid redundant models, a pruning operation is also applied.

1. Select an IBC-LIC model from the candidate list and signal its index.

The IBC-LIC model merge mode can be applied to both IBC-merge mode and IBC-AMVP mode. In IBC-merge mode, the selected LIC model is applied to the reference template during merge list reordering. Then the IBC merge list is further sorted by the inherited IBC LIC flags. In the IBC-AMVP mode, the selected LIC model is applied to the reference template during BVD prediction.

A flag is signalled to indicate whether the IBC-LIC merge mode is applied or not. If this flag is true, an index is further signaled to indicate which candidate model is used by the current block.

**Test 1.20: TIMD fusion with non-angular predictor (**[**JVET-AG0092**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0092-v1.zip)**)**

In TIMD of ECM, two intra modes providing the smallest template cost are selected from the list of candidates and blended in a fusion process according to their respective template costs.

In Test 1.20a, a third non-angular intra prediction mode (among DC and planar) with the smallest SATD cost is added in TIMD fusion process, the fusion weights are computed based on the SATD costs.

In Test 1.20b, additionally location-dependent sample-based blending of the DIMD fusion process is re-used but the location-dependent criterion applying to amplitudes of the selected predictors is replaced by a SATD cost-based criteria. The location-dependent criterion is determined from a ratio of the normalized SATD of the selected TIMD predictors computed in ABOVE and LEFT templates area.

**Test 1.21: TIMD fusion reference line determination (**[**JVET-AG0128**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0128-v1.zip)**)**

For TIMD mode with blending in ECM-11.0, the first mode uses reference line . While for the second mode, whether to use reference line or +1 depends on the following conditions:

* If all of the following conditions are true, is chosen.
* the current block is not ISP block
* the second mode is angular prediction mode
* the second mode does not represent non-fractional angles
* Otherwise, is chosen.

In this test, for the second mode, the modified (underscored) conditions of reference line determination are as follows:

* If all of the following conditions are true, is chosen.
* the current block is not ISP block
* both the first mode and second mode are angular prediction mode
* all of the following conditions are false
* abs(predModeIntra1 – predModeIntra2) is greater than *Threshold*. The value of *Threshold* is set to 8 or 4.
* (predModeIntra1 - EXT\_HOR\_IDX) \* (predModeIntra2 - EXT\_HOR\_IDX) is less than 0.
* (predModeIntra1 - EXT\_VER\_IDX) \* (predModeIntra2 - EXT\_VER\_IDX) is less than 0.
* Otherwise, is chosen.

**Test 1.22: Combination test of TIMD related tests**

Test 1.22a: Test 1.20a + Test 1.21 ([JVET-AG0093](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0093-v1.zip))

Test 1.22b: Test 1.20b + Test 1.21 ([JVET-AG0094](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0094-v1.zip))

Test 1.22c: Test 1.20b + Test 1.21 + Test 1.4 ([JVET-AG0095](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0095-v1.zip))

* TIMD/DIMD intra prediction modes may select IntraTMP/IBC BV (Test 1.22b + Test 1.4)
* a non-angular intra prediction mode (among DC, Planar or the best IntraTMP/IBC BV for TIMD/DIMD modes) is conditionally introduced (Test 1.20a + Test 1.4)
* location sample-based blending for TIMD is used similarly to DIMD (Test 1.20b and Test 1.22b)
* reference lines for TIMD are adjusted (Test 1.21 and Test 1.22b)



1.1: It was asked what the benefit of 1.1a would be, why not the fusion mode from 1.1b could be used without 1.1a? A proponent commented that 1.1a provides gain for luma, and 1.1b adds up gain for chroma. It was asserted that the fusion mode is a superset and cannot be operated standalone. 1.1b provides a reasonable tradeoff, runtimes also confirmed by crosscheck.

Decision: Adopt JVET-AG0154 test 1.1b.

1.2…1.7 are related to each other. It is observed that in the combinations the gain is in most cases more than additive from the different standalone tools (1.2…1.5). A possible explanation may be that all the tools use information from neighboring blocks, and it might be the case that a more frequent usage of intra TMP provides this benefit.

The test 1.7c was not originally planned, and RA results were not made available, and does not have a cross-check. From AI results, 1.7b and 1.7c are providing most attractive tradeoff, where 1.7b has higher gain (0.21% in AI, 0.12% in RA, but also has an encoder runtime increase of 2% in AI. All results except 1.7c are confirmed by crosschecks.

Decision: Adopt JVET-AG0137 test 1.7b.

1.8x is introducing additional candidates. 1.8a is most restricting the number of possible BVs to be traced and confirmed to be straightforward to implement by crosscheckers. In terms of gain, the different versions are not much different (small gain in AI, no gain in RA).

It was asked if effectively the search range / reference area of IBC was increased? It was confirmed by the proponent and a crosschecker that this was not the case.

Decision: Adopt JVET-AG0091 test 1.8a

1.9 has reasonable tradeoff in terms of encoder runtime vs. benefit, but probably introduces some additional processing in intra TMP: According to proponents, the decoder runtime increase is mostly caused by the more frequent usage of intra TMP mode itself rather than the additional processing steps. This is an additional mode which needs to be signalled. According to crosschecker, the mode should be simpler to process than the current Intra TMP mode.

Except for crosschecker, no support was expressed by other experts.

1.10 applies a bilateral filter on the intra prediction blocks (not in DC mode, and somewhat different in case of PDPC). Though there is almost no gain for RA, 0.12% luma gain in AI vs. 1% encoder run time increase would be a reasonable tradeoff. Crosschecker however reports an increase of 2-3% encoding run time. May not be accurate, though.

Gain is higher than reported in last meeting, due to some modifications of the filter.

It was commented that encoding run time might be reduced if the filtering would be omitted in the pre-checks for the RD decision.

It was commented that the bilateral filter when proposed to VVC was originally operating for intra priediction, but was moved into the loop filter stage for complexity considerations.

It was requested to clarify deviation of encoder run time between proponent and crosschecker. It was suggested that 1% would be reasonable tradeoff for adoption. It was later confirmed (Thu 25 Jan) by one cross-checker that in AI run times of encoder increase by 3.4%, decoder by 2.5%. Another cross-checker reports 1% encoding time (with partial results). Proponents explain that it may be dependent on the cluster configuration. Such dependency on simulation environment might indicate that the additional processing is not trivial. Further investigate in EE. It was recommended to test also a configuration where BF is not applied to planar.

1.11 … 1.13 are extending search ranges for intra TMP and IBC (1.13 is combination). Relevant gain is only observed for screen content (mainly class TGM).

1.11 (intra TMP) increases the encoder and decoder run time, whereas 1.13 does not. According to results, most of the gain of 1.13 seems to come from 1.12 (IBC).

The argument is made that the combination makes IBC and intra TMP more aligned.

Except for crosschecker, no support was expressed by other experts.

1.14 investigates methods of reducing encoder run time for extrapolation based intra prediction (which had been assessed of being too complex at encoder in past EEs). Tradeoff is acceptable, where 1.14b gives better gain (0.2% for luma AI) with slightly higher encoder run time. Also RA gain is reasonable

Decision: Adopt JVET-AG0058 test 1.14b

1.15 provides gain mostly in chroma, which comes mostly from 1.15a, but 1.15b may save some computation. It was asked how this relates to the proposal 1.1. According to proponents, the effect should be independent.

Some support expressed by non-proponents to go for the combination

Decision: adopt JVET-AG0059 test 1.15c

1.16 has small gain for screen content (<0.2%), not applied for camera content, introducing a somewhat new mode in LIC.

Tools with such small gain only for screen content are not worthwhile to consider. This particular proposal is also not a straightforward extension of the existing LIC.

1.17 is only relevant for screen content, where 1.17a has slightly higher gain (0.3%) but encoder run time increase, also in the combination it is still <0.4% for TGM, with >1% encoder run time.

Tools with such small gain only for screen content are not worthwhile to consider. This particular proposal is also not a straightforward extension of the existing LIC.

Except for crosschecker, no support was expressed by other experts.

1.18 gives 0.1% gain in AI, increasing encoder run time by >4%. No results on RA. No reasonable tradeoff.

1.19 is only beneficial for screen content, where the gain comes mainly from 2 sequences in class F (almost no gain in TGM). Run times of both encoder and decoder are increased, as there is obviously additional processing necessary. For camera content, small loss it observed, while still increasing run time.

Tools with such small gain only for screen content are not worthwhile to consider. This particular proposal (if disabled camera captured content) would also introduce a specific IBC LIC for SC classes, which is undesirable.

No action.

Tests 1.20 … are related. 1.20b has some additional processing, while giving almost no additional gain in the combination 1.22b compared to 1.22a. 1.22c also shows benefit in combination with 1.4 (which was already adopted).

However, the sample based fusion is already used for non-angular modes.

Reasonable tradeoff for AI, but almost no gain for RA. This may be due to the fact that in RA there are less adjacent intra coded blocks?

Decision: Adopt JVET-AG0094 test 1.22b.

***Inter prediction***

**Tests 2.1, 2.2, 2.5, 2.6a, 2.6e, and 2.6f on LIC improvement (**[**JVET-AG0176**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0176-v2.zip)**)**

In the tests, 3 aspects are introduced:

1. To signal LIC flag for inter-prediction merge modes, instead of inheriting the flag value from a merge candidate. The flag is signalled to indicate if the original inherited LIC flag or the reverse LIC flag value is used for a merge candidate. The flag is signalled for regular, affine, and TM merge modes.
2. To enable LIC with PU level BDMVR and BDOF.
3. Non-local illumination compensation, where the linear model is derived from the previously coded inter CUs by minimizing the difference between their reconstruction and prediction samples. When constructing the merge lists, up to 16 and 6 non-local candidates (obtained from both spatial adjacent and non-adjacent positions) are inserted to the lists of regular and subblock merge respectively and reordered with the existing merge candidates. The lengths of the output merge lists are kept unchanged. The same pattern used for non-adjacent merge mode is reused to locate the non-adjacent positions in the scheme.

Test 2.1: LIC flag signalling for inter prediction merge modes.

Test 2.2: Enable PU level BDMVR and BDOF for bi-predicted LIC.

Test 2.5: Non-local illumination compensation.

Test 2.6a: Test 2.1 + Test 2.2

Test 2.6e: Test 2.2 + Test 2.5

Test 2.6f: Test 2.1 + Test 2.2 + Test 2.5

**Tests 2.3, 2.4, 2.6b LIC with multiple templates and slope adjustment (**[**JVET-AG0099**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0099-v1.zip)**)**

In Test 2.3, LIC with multiple templates is introduced, where besides using a template comprising neighbouring samples both top to and left to the current block, two new LIC modes using top-only or left-only template are added to derive the parameters of LIC. LIC with multiple templates is only applied to AMVP uni-predicted blocks.

In Test 2.4, the slope adjustment like in CCLM is introduced to LIC, in which an adjustment parameter is used to modify parameters of LIC. The adjustment parameter is signalled for AMVP mode. LIC with slope adjustment is not applied to bi-prediction.

Test 2.6b: Test 2.2 + Test 2.3

**Tests 2.6g, 2.6h, 2.6i, 2.6j combination of tests on LIC improvement (**[**JVET-AG0276**](https://jvet-experts.org/doc_end_user/current_document.php?id=13849)**)**

Test 2.6g: Test 2.6e + Test 2.3

Test 2.6h: Test 2.6e + Test 2.4

Test 2.6i: Test 2.6f + Test 2.3

Test 2.6j: Test 2.6f + Test 2.4

**Test 2.7: AMVP with SbTMVP mode (**[**JVET-AG0098**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0098-v1.zip)**)**

In AMVP with SbTMVP mode, a CU is predicted in a similar way to SbTMVP in merge mode (CU is split into 4x4 subblocks, and each subblock derives its own motion from a corresponding subblock in the collocated picture) except that the motion shift is signalled in the bitstream instead of being derived from neighbouring blocks. The motion shift is obtained using MVP with a signalled MVD, and the number of MVDs is determined according to the percentage of the area of the blocks coded in the the AMVP with SbTMVP mode in the previous coded picture as follows:

* If the current picture is the first coded picture in a temporal layer, the number of MVD is set to 8, as shown in figure (a).
* Otherwise, if the percentage of the area of the proposed mode is smaller than threshold1, the number of MVD is set to 4, as shown in figure (b).
* Otherwise, if the percentage of the area of the proposed mode is smaller than threshold2, the number of MVD is set to 8, as shown in figure (a).
* Otherwise, the number of MVD is set to 12, as shown in figure (c).

|  |  |  |
| --- | --- | --- |
| A grid with a red dot in it  Description automatically generated | A grid with a red dot in center  Description automatically generated | A grid with a red dot in center  Description automatically generated |
| (a) | (b) | (c) |

In Test 2.7a, threshold1 and threshold2 are set to 3% and 6%, respectively.

In Test 2.7b, threshold1 and threshold2 are set to 4% and 7%, respectively.

**Test 2.8: On DMVR Extensions (**[**JVET-AG0067**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0067-v1.zip)**)**

In the test, 3 aspects are introduced.

In Test 2.8a, DMVR is extended to non-equal POC distance.

In Test 2.8b, 16x16 BDOF DMVR stage is added.

In Test 2.8c, mean removed equations are used to derive BDOF MV refinement parameters as follows.

(åGx.Gx+R1) \* vx + åGx.Gy \* vy = ådI . Gx. è (åGx.Gx+R1) \* vx + åGx.Gy \* vy = ådI . Gx - dM . åGx

åGx.Gy \* vx + (åGy.Gy+R1) \* vy= ådI . Gy è åGx.Gy \* vx + (åGy.Gy+R1) \* vy = ådI . Gy - dM . åGy

Test 2.8d: Test 2.8a + Test 2.8b + Test 2.8c

**Test 2.9: CIIP with subblock-based motion compensation (**[**JVET-AG0135**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0135-v1.zip)**)**

In the test, CIIP with subblock-based motion compensation is presented, where the inter prediction signal of CIIP is generated based on affine or SbTMVP motion compensation. A subblock-based merge candidate is used to generate the inter signal of CIIP, where the same subblock-based merge candidate list used by affine and SbTMVP is utilized.

When CIIP flag is true, a subbock-based CIIP flag is signalled. If subblock-based CIIP flag is true, an index indicating specific candidate in the subblock-based merge list is signalled, and TIMD is used to generate intra signal by default thus no CIIP-PDPC flag is signalled any more.

**Test 2.10: GPM with affine prediction (**[**JVET-AG0164**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0164-v2.zip)**)**

In ECM, GPM mode splits a CU into two partitions separated by a straight splitting line. The motion compensation for each GPM partition is performed to the whole CU, and a sample-based weighted blending process is used to generate the final prediction. Each partition of a GPM CU is merge-predicted with its own merge candidate index. In ECM, affine prediction is disallowed to predict a GPM partition.

In the test, GPM is extended to affine motion compensation, so a GPM partition can be predicted by affine motion compensation, non-affine motion compensation or intra prediction. When affine MC is applied, a uni-prediction affine merge candidate list is constructed from the subblock-based merge candidate list after discarding sub-TMVP candidates, similar to the uni-prediction merge candidate list construction for GPM in VVC.

A gpm\_affine\_flag is signalled for each GPM partition to indicate whether affine MC is applied for the GPM partition.

**Test 2.11: Regression-based GPM blending (**[**JVET-AG0112**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0112-v1.zip)**)**

In the test, two changes are introduced to GPM mode: an additional GPM blending mode and GPM partition merge index signalling modification.

In GPM implicit blending mode, two integer blending matrices (*W0* and *W1*) are derived from the template (1 line above, 1 column left). The blending matrices are modelled as an affine linear function of the sample positions (x,y) in the current CU:

*W0(x,y) = a.x + b.y + c* and *W1(x,y) = 1 - W0(x,y)*

The parameters (a,b,c) are derived from the reference template using the same solver (MSE minimization) as the one used for CCCM, GLM or GL-CCCM. A list of pair of candidates is built from the regular GPM candidates and re-ordered with the template cost.

The GPM implicit mode is signalled by a CU-level flag (*gpm\_implicit\_flag*). If *gpm\_implicit\_flag* is true, a *merge\_idx* is coded to signal the pair of GPM candidates to be used. If *gpm\_implicit\_flag* is false, the regular GPM syntax elements are signalled.

Three variants have been tested:

Test 2.11a (same as described above): a list of pair of candidates is built from the regular GPM candidates and re-ordered with the template cost. A single *merge\_idx* is coded to signal the pair of GPM candidates to be used.

Test 2.11b: proposed GPM blending but two *merge\_idx* are coded (one for each partition) without re-ordering, same as regular GPM syntax.

Test 2.11c: the regular GPM blending is used but a single *merge\_idx* is coded to signal the pair of GPM candidates and the split to be used. Both pairs of candidates and split are re-ordered with the template cost.

**Test 2.11: Combination of Test 2.9, Test 2.10 and Test 2.11a (**[**JVET-AG0142**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0142-v1.zip)**)**

Test 2.11d: Test 2.10 + Test 2.11a

Test 2.11e: Test 2.9 + Test 2.10

Test 2.11f: Test 2.9 + Test 2.10 + Test 2.11a

**Test 2.12: Adjusting out-of-boundary prediction samples (**[**JVET-AG0097**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0097-v1.zip)**)**

In ECM, in bi-directional motion compensation the out of boundary (OOB) prediction samples are discarded and only the non-OOB predictors, when available, are used to generate the final predictor.

If is OOB and is non-OOB

else if is non-OOB and is OOB

else

In the test, the final prediction samples for luma are generated using scales and offsets as follows,

if is OOB and is non-OOB

else if is non-OOB and is OOB

else

.

The scales and offsets are derived over the non-OOB parts of and using least squares minimization,

and

where are the non-OOB samples of and . Scales and offsets are derived using sample covariance and sample variance at integer precision. This modification is applied to luma only, and chroma uses ECM process.

Following results are RA and LB.



2.1 … 2.6 are all related to LIC modifications. All combination tests relating to 2.3 and 2.4 were not finished initially, see further comments below (test 2.6j).

Consideration about 2.1, 2.2 and 2.5 and their combinations. 2.6e and 2.6f look attractive, where 2.6e has a better tradeoff (though slightly less gain), increasing encoder run time only by 1.2% rather than 2.4%

Decision: Adopt JVET-AG0176 test 2.6e

After completion of the combinations with 2.3 and 2.4, the additional benefit of test 2.6f was investigated. In an update (v3) of the EE summary, test 2.6j indicates (in comparison to test 2.6h) that 2.6f has a benefit of 0.08% in the combination of all tools (except for 2.3) gains are additive. Decision: Adopt JVET-AG0276 test 2.6j.

2.7 has reasonable tradeoff, gain in RA and even more in LB.

Decision: Adopt JVET-AG0098 test 2.7b

2.8 includes three different aspects, of which 2.8b has slightly worse tradeoff (largest encoding time increase) compared to a/c. Combination 2.8d indicates that gains are (more than) additive.

Decision: Adopt JVET-AG0067 test 2.8d

2.9 … 2.11 are related to each other. Combination 2.11f indicates that gains of 2.9, 2.10 and 2.11a are somewhat additive, overall a reasonable tradeoff (0.28% in RA, with 2.3% encoder, similar in LB).

Decision: Adopt JVET-AG0142 test 2.11f

2.12 shows a small benefit in coding performance, marginal increase in runtime, which can be expected as it is only applied to few samples at picture boundary. In class D, where it is applied relatively more frequently, the gain is 0.08%, but run time still kept almost constant.

No support by non-proponents – no action.

***Transform and coefficient coding***

**Test 3.1 CABAC inter/intra model switch for residual coding (**[**JVET-AG0143**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0143-v1.zip)**)**

In the current context design, residual information (e.g. residual coefficients, QtCbf, SigCoeffGroup etc.) are coded using context models depending on the slice type, but not on the prediction mode of the current CU.

It is asserted that it is better to use context models depending on the prediction rather than slice type.

In the tests, for intra coded CUs, intra context models are used independently of the slice type. Conversely, CUs which are coded with a mode similar to inter prediction (IBC and IntraTMP), inter context models are used to code residual information independently of the slice type.

The following tests have been conducted.

Test 3.1a: Use intra context models to encode residual information of intra coded CU in inter slices.

Test 3.1b: Test 3.1a and use inter context models to encode residual information of IBC or IntraTMP coded CUs in intra slices.

Test 3.1c: Test 3.1b, but QtCbf flag is excluded from the syntax elements for which a model switch is allowed.

Test 3.1d: as in Tests 3.1b and 3.1c, but all CABAC contexts are retrained.

**Test 3.2 Transform coefficient coding (**[**JVET-AG0100**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0100-v1.zip)**)**

In ECM and VVC residual coding, regular coded absolute coefficient levels are represented by sig\_coeff\_flag(sig), abs\_level\_gtx\_flag[0](gt1), par\_level\_flag(par),abs\_level\_gtx\_flag[1](gt3) syntax elements along with bypass coded abs\_remainder(remLevel). Associated coefficient level is reconstructed in ECM in the following way:

abs\_coeff\_level = sig + gt1 + par + 2\*(gt3 + remLevel).

In the test, the number gtX flags coded in the regular mode of the arithmetic coding using adaptive context models before remainder coding is changed. There are N+1 flags coded in regular mode: sig\_coeff\_flag, gt1\_flag, gt2\_flag, gt3\_flag, …, gtN\_flag. Parity flag, par\_level\_flag is coded in bypass mode after the gtN\_flag. In the experiments N was set to 7. Absolute coefficient level is reconstructed as follows:

abs\_coeff\_level = sig +gt1+ gt2 +gt3 +… + gtN + par + 2\*remLevel,

The grouping of regular coded bins and by bypass coded bins is maintained.

The flags gt1, gt2, gt3 use separate contexts, flags gt4, gt5, gt6 and gt7 share the same context.

In Test 3.2a the new added contexts are not retrained, in Test 3.2b contexts related to coefficient coding are retrained. In Test 3.2c, all contexts are retrained.

Additional test noted as Test 3.b\* was performed with the reduced bin budget.

**Test 3.3: Utilizing LFNST/NSPT for inter coding (**[**JVET-AG0061**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0061-v1.zip)**)**

In the test, the utilization of LFNST/NSPT for inter coding is evaluated. For inter coded block, an intra prediction mode is first derived according to the inter prediction block with a DIMD-like process applied to the prediction. Then the derived intra prediction mode is used to select an LFNST/NSPT transform set and the transform can be processed with the selected LFNST/NSPT kernel, like in the intra coding process.

The signalling of inter LFNST/NSPT index is different from that of intra LFNST/NSPT index. The intra LFNST/NSPT index binarization employs two context coded bins for each symbol, while truncated unary code with different context models for inter LFNST/NSPT index coding is used in the test.

Encoding fast algorithm like inter MTS is applied to reduce the encoding time.

Results AI/RA/LB



(check table copy/paste, and include LP where available)

Test 3.1 cannot be expected having impact on run time, these should be asserted to be unreliable (cross check indicates 100%). 3.1c gives gain particularly in RA and LB

Decision: Adopt JVET-AG0143 test 3.1c

3.2c is same as 3.2b with another anchor – just intends to show that the gain is retained in combination with 5.1. It was commented that 3.2b increases the bin budget relative to current ECM, whereas 3.2b\* is keeping the same bin budget as ECM, and performance is almost identical. Generally, both 3.2b proposals show good tradeoff.Decision: Adopt JVET-AG0100 test 3.2b\*

3.3 applies LFNST/NSPT to all inter coded blocks. The ECM/VTM11 anchor of CTC is using VVC LFNST for AI and RA (only intra coded blocks), not for LB. For the results, the anchor also enables VVC LFNST for LB. The gains are due to the modifications of LFNST in ECM (and using for all inter coded blocks), and NSPT. One crosschecker reports that they observed higher increase on encoding run time, but another crosschecker found similar encoding time as reported by proponents.

Compared to CTC in LB, the gain would be around 1.5% for LB, with 10% encoding runtime increase. Some concern is expressed about this amount of encoding time, considering that LB is already critical in run time.

Decision: Adopt JVET-AG0061 Test 3.3, disabled in CTC for LB/LP.

***In-loop filtering***

**Test 4.1: Adaptive clipping with signalled lower and upper bounds (**[**JVET-AG0145**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0145-v1.zip)**)**

In ECM, a bit-depth based clipping defined as [0, 2^bitdepth-1] is applied in filtering, sampling, interpolation, weighted prediction, weighted combination, reconstruction, and other stages, to ensure that the generated prediction and reconstruction samples remain in the defined dynamic range.

In the test, the range [min, max] of the clipping is signalled as delta values in picture header. For I pictures, the min/max is signalled as delta values compared to [64, 940], for other pictures, the delta values compared to the min/max values of the collocated picture are signalled.

When clipping is performed in LMCS domain, the min/max values are derived by applying forward LMCS look-up table to the signalled min/max values.

At encoder, the min/max values are obtained by scanning the sample values in the original picture. If motion compensated temporal filter (MCTF) is applied to a picture, the min/max values are obtained by scanning the sample values in the MCTF pre-filtered picture. The min/max values are used as the lower and upper bounds for clipping at the reconstruction stage (when adding the residual signal to the prediction) and before saving the reconstructed picture into the decoded picture buffer.

Test 4.1a: Adaptive clipping with signalled min/max values from MCTF prefiltered picture.

Test 4.1b: Adaptive clipping with signalled min/max values from original picture.

**Test 4.2: Fixed filter for chroma ALF (**[**JVET-AG0157**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0157-v1.zip)**)**

In the tests, fixed filter for chroma ALF is introduced.

In Test 4.2a, like the classifiers for luma component, a classifier based on Laplacian values and variance is applied to a 2x2 chroma block. The derived class index is then used to select a fixed filter from a chroma filter set, which is trained from BVI video set. A chroma fixed filter is applied to chroma ALF input samples in a 13x13 diamond shape and DBF input samples in a 7x7 diamond shape.

In Test 4.2b, luma ALF fixed filter is reused. The first luma classifier is applied to each 2x2 chroma block, then the derived class index is used to select a fixed filter from the luma fixed filter set related to this classifier. A fixed filter is applied to chroma ALF input sample in a 9x9 diamond shape and DBF input samples in a 9x9 diamond shape.

Finally, in a signalled chroma filter, 5x5 crossing extra taps are introduced, which are applied to the fixed filter output.

**Test 4.3: Adaptive precision for luma ALF coefficients (**[**JVET-AG0158**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0158-v1.zip)**)**

In ALF of ECM-11.0, 7 bits are used to represent the fractional part of a coefficient of a signalled ALF filter.

In the test, the number of bits used to represent the fractional part of a luma coefficient is adaptive from 5 to 8, inclusively. For each luma filter set, which contains up to 25 filters, a 2-bit syntax element is signalled in APS to indicate the number of bits used for the coefficients in this set. The value range of a coefficient is not changed.



4.1 is a straightforward change and shows reasonable tradeoff (4.1a better).

It was asked if clipping also in prediction would provide benefit? According to previous investigations, proponents commented that they did not observe additional gain.

Decision: Adopt JVET-AG0145 test 4.1a

About 4.2, it was asked how many new filters are introduced in 4.2a? 8 sets with 512 filters each, each with >50 coefficients. 50 mult per chroma samples. 4.2b is less complex (using 9x9 cross filter shape, only first is used), but still has interesting gain.

Decision: Adopt JVET-AG0157 test 4.2b

4.3 is simple change with reasonable tradeoff.

It was asked what the impact of encoder optimization that was applied in the proposal was? According to proponents, it is minor, and additional results applying the same optimization to anchor are shown in the contribution. It was also asserted to be numerically stable (float precision in optimization).

It was also reported that a bug was found in the optimization of ALF, which might also need correction in VTM. Proponents will submit a ticket on this.

Decision: Adopt JVET-AG0158 test 4.3

***Entropy coding***

**Tests 5.1-5.2: CABAC context initialization retraining and slice type based window offsets (**[**JVET-AG0196**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0196-v1.zip)**)**

In ECM, CABAC has the following parameters: probability states (6 bits), short and long window sizes (4 bits), weights (6 bits), and offsets for short (8 bits) and long (8 bits) windows. In total, (6 bits + 4 bits + 6 bits) \* 3 slices + 8 bits \* 2 windows = 64 bits is required per context.

In Test 5.1, initialization parameters of all CABAC contexts for all slices are retrained.

In Test 5.2, windows offsets defined per slice type are added and contexts are retrained. It requires 8 bits \* 2 windows \* 2 slices = 32 bits additional memory per context.

Context training software is provided located at source/App/CabacTraining including readme file describing how to use it.

**Test 5.3: Spatial CABAC tuning (**[**JVET-AG0117**](https://jvet-experts.org/doc_end_user/documents/33_Teleconference/wg11/JVET-AG0117-v1.zip)**)**

In ECM, CTUs are scanned in raster scan order, while CUs inside CTUs are scanned in nested tree order. When starting to process the current CTU (CTUC) the processing “jumps” from the bottom-right CU (yellow CU 8) to the top-left CU of the current CTU (blue CU 0) as shown in the next figure. At that stage, the context parameters of the CABAC engine can be expected to be fine tuned for the bottom-right area of the previous CTU instead of the ideal case where the context parameters would be fine tuned for the actual next CU to be processed.



In the tested method, syntax elements bins of the bottom CUs of the above CTU (the green CUs in the figure) are used to tune the CABAC contexts when starting to process a new CTU.

In the implementation, bins for 128 contexts are buffered for each CTU of a CTU line. The number of bin values that can be stored for each of those contexts is limited to 4 in this experiment. Thus, a maximum of 128 \* 4 bin values equalling to 512 bin values may be stored for each CTU in one CTU row.

When starting to encode or decode a new CTU, the stored bins from the above CTU are retrieved and fed into the context update mechanism of the shorter window probability estimator (with “window 0”). The bins are fed to context update process in the reverse order with respect to their coding order to simulate encoding or decoding that approaches the top-left corner of the new CTU from the direction of bottom-right corner of the CTU above.

Decoded bins are stored in a buffer, an element of such buffer should hold a 10-bit context ID (for the current ECM where the number of contexts is below 1024), a 4-bit bin storage containing the bin values for that context and a 2-bit counter indicating how many bins are stored for that context summarized in the next table.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Test | Bits / element | Number of elements | Total bits / CTU | Total bytes / CTU |
| Test 5.3 | 10+4+2 = 16 | 128 | 2048 | 256 |

The total memory requirement then depends on the number of CTUs in a CTU row. Considering the CTC settings for different classes of content and the memory elements outlined above, the total amount of memory needed for the Tests 5.3a and 5.3b are as follows:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Class | CTUs / | Bytes / | Bytes / | Bytes for | Bytes |
|  | row | CTU | CTU row | active CTU | Total |
| A1 | 15 | 256 | 3840 | 747 | 4587 |
| A2 | 15 | 256 | 3840 | 747 | 4587 |
| B | 15 | 256 | 3840 | 747 | 4587 |
| C | 7 | 256 | 1792 | 747 | 2539 |
| D | 4 | 256 | 1024 | 747 | 1771 |
| E | 10 | 256 | 2560 | 747 | 3307 |
| F | 15 | 256 | 3840 | 747 | 4587 |
| TGM | 15 | 256 | 3840 | 747 | 4587 |

In Test 5.3b, context initialization from Test 5.1 is used.



5.1 is providing CABAC retraining for all slice types (including inter slices, for which it had not been available before)

5.2 is further retraining the slice-type dependent offsets.

Both approaches are assessed to be mature. Script should be included into ECM, and be used to retrain the entropy coding regularly (due to the effort in computation, may not be possible in each meeting cycle).

Decision(SW): Adopt JVET-AG0196 tests 5.1 and 5.2. A documentation of the script (and data to be used) should also be included in JVET-AG2025 (ECM).

5.3 is confirmed by crosscheckers to be appropriately implemented, also gives gain on top of re-training (5.3b).

Decision: Adopt JVET-AG0117 test 5.3a

### EE2 contributions: Enhanced compression beyond VVC capability (41)

There was no presentation or discussion about specific proposals in this category.

For actions decided to be taken, see section 5.2.1, unless otherwise noted.

[JVET-AG0058](https://jvet-experts.org/doc_end_user/current_document.php?id=13614) EE2-1.14: An extrapolation filter-based intra prediction mode [L. Xu, Y. Yu, H. Yu, J. Gan, D. Wang (OPPO)]

[JVET-AG0250](https://jvet-experts.org/doc_end_user/current_document.php?id=13823) Crosscheck of JVET-AG0058 (EE2-1.14: An extrapolation filter-based intra prediction mode) [X. Li (Alibaba)] [late]

[JVET-AG0265](https://jvet-experts.org/doc_end_user/current_document.php?id=13838) Crosscheck of JVET-AG0058 (EE2-1.14: An extrapolation filter-based intra prediction mode) [H.-J. Jhu (Kwai)] [late]

[JVET-AG0059](https://jvet-experts.org/doc_end_user/current_document.php?id=13615) EE2-1.15: Enhancements on CCP merge for chroma intra coding [H. Huang, Y. Yu, H. Yu, D. Wang (OPPO), Z. Deng, K. Zhang, L. Zhang (Bytedance)]

[JVET-AG0249](https://jvet-experts.org/doc_end_user/current_document.php?id=13822) Crosscheck of JVET-AG0059 (EE2-1.15: Enhancements on CCP merge for chroma intra coding) [X. Li (Alibaba)] [late]

[JVET-AG0060](https://jvet-experts.org/doc_end_user/current_document.php?id=13616) EE2-1.19: IBC-LIC Model Merge mode [L. Zhang, Y. Yu, H. Yu, D. Wang (OPPO)]

[JVET-AG0220](https://jvet-experts.org/doc_end_user/current_document.php?id=13793) Crosscheck of JVET-AG0060 (EE2-1.19: IBC-LIC Model Merge mode) [Y. Wang (Bytedance)] [late]

[JVET-AG0061](https://jvet-experts.org/doc_end_user/current_document.php?id=13617) EE2-3.3: Utilizing LFNST/NSPT for inter coding [F. Wang, J. Gan, Y. Yu, H. Yu, D. Wang (OPPO)]

[JVET-AG0211](https://jvet-experts.org/doc_end_user/current_document.php?id=13767) Crosscheck of JVET-AG0061 (EE2-3.3: Utilizing LFNST/NSPT for inter coding) [M. Koo (LGE)] [late]

[JVET-AG0067](https://jvet-experts.org/doc_end_user/current_document.php?id=13623) EE2-2.8: On DMVR Extensions [M. Salehifar, Y. He, K. Zhang, L. Zhang (Bytedance)]

[JVET-AG0267](https://jvet-experts.org/doc_end_user/current_document.php?id=13840) Crosscheck of JVET-AG0067 (EE2-2.8: On DMVR Extensions) [H.-J. Jhu, X. Xiu (Kwai)] [late]

[JVET-AG0301](https://jvet-experts.org/doc_end_user/current_document.php?id=13874) Crosscheck of JVET-AG0067 (EE2: test 2.8a - DMVR extension) [P. Le Guyadec (InterDigital)] [late]

[JVET-AG0259](https://jvet-experts.org/doc_end_user/current_document.php?id=13832) Crosscheck report of JVET-AG0061: EE2-3.3: Utilizing LFNST/NSPT for inter coding [K. Naser, S. Puri (InterDigital)] [late]

[JVET-AG0072](https://jvet-experts.org/doc_end_user/current_document.php?id=13628) EE2-1.12: IBC with upward-extended reference area [Y. Kidani, H. Kato, K. Kawamura (KDDI)]

[JVET-AG0206](https://jvet-experts.org/doc_end_user/current_document.php?id=13762) Crosscheck of JVET-AG0072 (EE2-1.12: IBC with upward-extended reference area) [N. Zhang (Bytedance)] [late]

[JVET-AG0076](https://jvet-experts.org/doc_end_user/current_document.php?id=13632) EE2-1.18 DIMD mode derivation from spatial blocks [J. Huo, J. Fan, Z. Zhang, Y. Ma, F. Yang (Xidian Univ.), M. Li (OPPO)]

[JVET-AG0251](https://jvet-experts.org/doc_end_user/current_document.php?id=13824) Crosscheck of JVET-AG0076 (EE2-1.18 DIMD mode derivation from spatial blocks) [X. Li (Alibaba)] [late]

[JVET-AG0091](https://jvet-experts.org/doc_end_user/current_document.php?id=13647) EE2-1.8: Auto-relocated block vector prediction [N. Zhang, K. Zhang, L. Zhang (Bytedance)]

[JVET-AG0134](https://jvet-experts.org/doc_end_user/current_document.php?id=13690) Crosscheck of JVET-AG0091 (EE2-1.8: Auto-relocated block vector prediction) [D. Ruiz Coll (Ofinno)] [late]

[JVET-AG0225](https://jvet-experts.org/doc_end_user/current_document.php?id=13798) Crosscheck of JVET-AG0091 (EE2-1.8: Auto-relocated block vector prediction) [Y. Kidani, K. Kawamura (KDDI)] [late]

[JVET-AG0092](https://jvet-experts.org/doc_end_user/current_document.php?id=13648) EE2-1.20: TIMD fusion with non-angular predictor [P. Andrivon, M. Blestel (Ofinno)]

[JVET-AG0252](https://jvet-experts.org/doc_end_user/current_document.php?id=13825) Crosscheck of JVET-AE0092 (EE2-1.20: TIMD fusion with non-angular predictor) [M. Abdoli, R. G. Youvalari (Xiaomi)] [late]

[JVET-AG0132](https://jvet-experts.org/doc_end_user/current_document.php?id=13688) Crosscheck of EE2-Test 1.20b: TIMD fusion with non angular predictor and sample-based fusion [S. Blasi (Nokia)] [late]

[JVET-AG0093](https://jvet-experts.org/doc_end_user/current_document.php?id=13649) EE2-1.22a: Combination of EE2-1.20a and EE2-1.21 [P. Andrivon, M. Blestel (Ofinno), [C. Zhou](mailto:chuan.zhou@vivo.com), Z. Lv (vivo)]

[JVET-AG0253](https://jvet-experts.org/doc_end_user/current_document.php?id=13826) Crosscheck of JVET-AG0093 (EE2-1.22a: Combination of EE2-1.20a and EE2-1.21) [M. Abdoli, R. G. Youvalari (Xiaomi)] [late]

[JVET-AG0094](https://jvet-experts.org/doc_end_user/current_document.php?id=13650) EE2-1.22b: Combination of EE2-1.20b and EE2-1.21 [P. Andrivon, M. Blestel (Ofinno), C. Zhou, Z. Lv (vivo)]

[JVET-AG0133](https://jvet-experts.org/doc_end_user/current_document.php?id=13689) Crosscheck of JVET-AG0094 (EE2-1.22b: Combination of EE2-1.20b and EE2-1.21) [S. Blasi (Nokia)] [late]

[JVET-AG0095](https://jvet-experts.org/doc_end_user/current_document.php?id=13651) EE2-1.22c: Combination of EE2-1.22b and EE2-1.4 [P. Andrivon, M. Blestel (Ofinno), C. Zhou, Z. Lv (vivo), K. Naser, F. Le Léannec (InterDigital)]

[JVET-AG0127](https://jvet-experts.org/doc_end_user/current_document.php?id=13683) Crosscheck of JVET-AG0095 (EE2-1.22c: Combination of EE2-1.22b and EE2-1.4) [H. Qin (TCL)] [late]

[JVET-AG0097](https://jvet-experts.org/doc_end_user/current_document.php?id=13653) EE2-2.12: Adjusting out-of-boundary prediction samples [P. Astola, J. Lainema (Nokia)]

[JVET-AG0205](https://jvet-experts.org/doc_end_user/current_document.php?id=13761) Crosscheck of JVET-AG0097 (EE2-2.12: Adjusting out-of-boundary prediction samples) [Y. Wang (Bytedance)] [late]

[JVET-AG0098](https://jvet-experts.org/doc_end_user/current_document.php?id=13654) EE2-2.7: AMVP with SbTMVP mode [R.-L. Liao, J. Chen, Y. Ye, X. Li (Alibaba)]

[JVET-AG0234](https://jvet-experts.org/doc_end_user/current_document.php?id=13807) Crosscheck of JVET-AG0098 (EE2-2.7: AMVP with SbTMVP mode) [L. Zhao (Bytedance)] [late]

[JVET-AG0288](https://jvet-experts.org/doc_end_user/current_document.php?id=13861) Crosscheck of JVET-AG0098 (EE2-2.7: AMVP with SbTMVP mode) [F. Pu (Dolby)] [late]

[JVET-AG0099](https://jvet-experts.org/doc_end_user/current_document.php?id=13655) EE2 Test 2.3, 2.4, 2.6b: LIC with multiple templates and slope adjustment [Y. Wang, K. Zhang, Y. He, H. Liu, L. Zhang (Bytedance), X. Xiu, C. Ma, N. Yan, H.-J. Jhu, C.-W. Kuo, W. Chen, X. Wang (Kwai)]

[JVET-AG0214](https://jvet-experts.org/doc_end_user/current_document.php?id=13770) Crosscheck of JVET-AG0099 (EE2-2.4: LIC with slope adjustment) [P. Astola (Nokia)] [late]

[JVET-AG0262](https://jvet-experts.org/doc_end_user/current_document.php?id=13835) Crosscheck of JVET-AG0099 (EE2 Test 2.3, 2.6b: LIC with multiple templates and slope adjustment) [L. Zhang (OPPO)] [late]

[JVET-AG0100](https://jvet-experts.org/doc_end_user/current_document.php?id=13656) EE2-3.2: Transform coefficient coding [P. Nikitin, M. Coban, M. Karczewicz, P. Garus, V. Seregin (Qualcomm)]

[JVET-AG0159](https://jvet-experts.org/doc_end_user/current_document.php?id=13715) Crosscheck of JVET-AG0100 EE2-3.2: Transform coefficient coding [F. Lo Bianco, F. Galpin (Interdigital)] [late]

[JVET-AG0103](https://jvet-experts.org/doc_end_user/current_document.php?id=13659) EE2-1.16: IBC LIC with slope adjustment [C. Ma, X. Xiu, W. Chen, H.-J. Jhu, C.-W. Kuo, N. Yan, X. Wang (Kwai)]

[JVET-AG0221](https://jvet-experts.org/doc_end_user/current_document.php?id=13794) Crosscheck of JVET-AG0103 (EE2-1.16: IBC LIC with slope adjustment) [Y. Wang (Bytedance)] [late]

[JVET-AG0104](https://jvet-experts.org/doc_end_user/current_document.php?id=13660) EE2-1.17: IBC GPM with block vector difference and split mode reordering [C. Ma, X. Xiu, W. Chen, H.-J. Jhu, C.-W. Kuo, N. Yan, X. Wang (Kwai)]

[JVET-AG0222](https://jvet-experts.org/doc_end_user/current_document.php?id=13795) Crosscheck of JVET-AG0104 (EE2-1.17: IBC GPM with block vector difference and split mode reordering) [Y. Wang (Bytedance)] [late]

[JVET-AG0112](https://jvet-experts.org/doc_end_user/current_document.php?id=13668) EE2-2.11: Regression-based GPM blending (tests a,b,c) [P. Bordes, K. Reuzé, F. Galpin, F. Urban, K. Naser, F. Le Léannec, E. François (InterDigital)]

[JVET-AG0285](https://jvet-experts.org/doc_end_user/current_document.php?id=13858) Crosscheck of Test-2.11a from JVET-AG0112 (EE2-2.11: Regression-based GPM blending (tests a,b,c)) [L. Zhao (Bytedance)] [late]

[JVET-AG0117](https://jvet-experts.org/doc_end_user/current_document.php?id=13673) EE2-5.3: Spatial CABAC tuning [J. Lainema, A. Aminlou, P. Astola, D. B. Sansli (Nokia)]

[JVET-AG0119](https://jvet-experts.org/doc_end_user/current_document.php?id=13675) Crosscheck of EE2-Test 5.1 and 5.3 (JVET-AG0117 Spatial CABAC tuning in combination with retrained context initialization for inter slices) [K. Andersson (Ericsson)]

[JVET-AG0124](https://jvet-experts.org/doc_end_user/current_document.php?id=13680) Crosscheck of EE2-Test 5.3 (Spatial CABAC tuning) [P. Nikitin (Qualcomm)] [late]

[JVET-AG0118](https://jvet-experts.org/doc_end_user/current_document.php?id=13674) EE2-1.9: Intra TMP fusion probing [J.-L. Lin, P.-H. Lin, H. Wang, Y.-J. Chang, Z. Zhang, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-AG0236](https://jvet-experts.org/doc_end_user/current_document.php?id=13809) Crosscheck of JVET-AG0118 (EE2-1.9: Intra TMP fusion probing) Z. Deng (Bytedance) [late]

[JVET-AG0123](https://jvet-experts.org/doc_end_user/current_document.php?id=13679) EE2-1.10: Bilateral Filtering for Intra Prediction [W. Yin, K. Zhang, Y. Wang, Z. Deng, L. Zhao, N. Zhang, M. Salehifar, L. Zhang (Bytedance)]

[JVET-AG0248](https://jvet-experts.org/doc_end_user/current_document.php?id=13821) Crosscheck of JVET-AG0123 (EE2-1.10: Bilateral Filtering for Intra Prediction) [X. Li (Alibaba)] [late]

[JVET-AG0318](https://jvet-experts.org/doc_end_user/current_document.php?id=13891) Crosscheck of JVET-AG0123 (EE2-1.10: Bilateral Filtering for Intra Prediction) [V. Shchukin (Ericsson)] [late]

[JVET-AG0320](https://jvet-experts.org/doc_end_user/current_document.php?id=13893) Crosscheck of JVET-AG0123 (EE2-1.10: Bilateral Filtering for Intra Prediction) [N. Hu (Qualcomm)] [late]

[JVET-AG0128](https://jvet-experts.org/doc_end_user/current_document.php?id=13684) EE2-1.21: TIMD fusion reference line determination [C. Zhou, Z. Lv (vivo)]

[JVET-AG0275](https://jvet-experts.org/doc_end_user/current_document.php?id=13848) Crosscheck of JVET-AG0128 (EE2-1.21: TIMD fusion reference line determination) [J. Chen (Alibaba)] [late]

[JVET-AG0131](https://jvet-experts.org/doc_end_user/current_document.php?id=13687) EE2-1.11: IntraTMP search range extension with adaptive sampling [D. Ruiz Coll (Ofinno), K. Naser, P. Bordes, F. Le Léannec, F. Galpin (InterDigital)]

[JVET-AG0229](https://jvet-experts.org/doc_end_user/current_document.php?id=13802) Crosscheck of JVET-AG0131 (EE2-1.11: IntraTMP search range extension with adaptive sampling) [G. Verba (Qualcomm)] [late]

[JVET-AG0135](https://jvet-experts.org/doc_end_user/current_document.php?id=13691) EE2-2.9: CIIP with subblock-based motion compensation [L. Zhao, K. Zhang, L. Zhang (Bytedance)]

[JVET-AG0271](https://jvet-experts.org/doc_end_user/current_document.php?id=13844) Crosscheck of JVET-AG0135 (EE2-2.9: CIIP with subblock-based motion compensation) [W. Chen (Kwai)] [late]

[JVET-AG0136](https://jvet-experts.org/doc_end_user/current_document.php?id=13692) EE2-1.5: IntraTMP extension to LIC [F. Le Léannec, T. Dumas, K. Naser, Y. Chen, M. Radosavljević (InterDigital)]

[JVET-AG0171](https://jvet-experts.org/doc_end_user/current_document.php?id=13727) Crosscheck of JVET-AG0136 (EE2-1.5: IntraTMP extension to LIC) [K. Cui (Qualcomm)] [late]

[JVET-AG0137](https://jvet-experts.org/doc_end_user/current_document.php?id=13693) EE2-1.6/1.7a/1.7b: combinations of tests 1.2, 1.3, 1.4 and 1.5 [F. Le Léannec, K. Naser, T. Dumas, Y. Chen, M. Radosavljević, T. Poirier (InterDigital)

[JVET-AG0255](https://jvet-experts.org/doc_end_user/current_document.php?id=13828) Crosscheck of JVET-AG0137 (EE2-1.6 combinations of tests 1.4 and 1.5) [R. G. Youvalari, M. Abdoli (Xiaomi)] [late]

[JVET-AG0274](https://jvet-experts.org/doc_end_user/current_document.php?id=13847) Crosscheck of JVET-AG0137 EE2-1.7a/b [F. Wang, L. Zhang (OPPO)] [late]

[JVET-AG0142](https://jvet-experts.org/doc_end_user/current_document.php?id=13698) EE2-2.11d/e/f: Combination of Test 2.9, Test 2.10 and Test 2.11a [L. Zhao, K. Zhang, L. Zhang (Bytedance), P. Bordes, K. Reuzé (InterDigital)]

[JVET-AG0223](https://jvet-experts.org/doc_end_user/current_document.php?id=13796) Crosscheck of EE2-2.11d from JVET-AG0142 (EE2-2.11d/e/f: Combination of Test 2.9, Test 2.10 and Test 2.11a) [M. Blestel (Ofinno)] [late]

[JVET-AG0297](https://jvet-experts.org/doc_end_user/current_document.php?id=13870) Crosscheck of JVET-AG0142 (EE2-2.11d/e/f: Combination of Test 2.9, Test 2.10 and Test 2.11a) [C. Ma (Kwai)] [late]

[JVET-AG0143](https://jvet-experts.org/doc_end_user/current_document.php?id=13699) EE2-3.1 CABAC inter/intra model switch for residual coding [F. Lo Bianco, F. Galpin, C. Salmon-Legagneur (InterDigital)]

[JVET-AG0147](https://jvet-experts.org/doc_end_user/current_document.php?id=13703) Crosscheck of EE2-3.1 CABAC inter/intra model switch for residual coding [P. Nikitin (Qualcomm)] [late]

[JVET-AG0145](https://jvet-experts.org/doc_end_user/current_document.php?id=13701) EE2-4.1: Adaptive clipping with signalled lower and upper bounds [K. Cui, Z. Zhang, H. Huang, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-AG0246](https://jvet-experts.org/doc_end_user/current_document.php?id=13819) Cross-check of JVET-AG0145 (EE2-4.1: Adaptive clipping with signalled lower and upper bounds) [F. Le Léannec (InterDigital)] [late]

[JVET-AG0146](https://jvet-experts.org/doc_end_user/current_document.php?id=13702) EE2-1.4: IntraTMP extension to DIMD [K. Naser, T. Poirier, F. Le Léannec, T. Dumas, Y. Chen, M. Radosavljević (InterDigital)]

[JVET-AG0170](https://jvet-experts.org/doc_end_user/current_document.php?id=13726) Crosscheck of JVET-AG0146 (EE2-1.4: IntraTMP extension to DIMD) [K. Cui (Qualcomm)] [late]

[JVET-AG0151](https://jvet-experts.org/doc_end_user/current_document.php?id=13707) EE2-1.2: IntraTMP with Merge Candidates [K. Naser, M. Radosavljević, F. Le Léannec, Y. Chen, T. Dumas, T. Poirier (InterDigital)]

[JVET-AG0244](https://jvet-experts.org/doc_end_user/current_document.php?id=13817) Crosscheck of JVET-AG0151 (EE2-1.2: IntraTMP with Merge Candidates) [I. Zupancic (Nokia)] [late]

[JVET-AG0152](https://jvet-experts.org/doc_end_user/current_document.php?id=13708) EE2-1.3: SGPM with IntraTMP and IBC [K. Naser, Y. Chen, F. Le Léannec (InterDigital)]

[JVET-AG0245](https://jvet-experts.org/doc_end_user/current_document.php?id=13818) Crosscheck of JVET-AG0152 (EE2-1.3: SGPM with IntraTMP and IBC) [I. Zupancic (Nokia)] [late]

[JVET-AG0154](https://jvet-experts.org/doc_end_user/current_document.php?id=13710) EE2-1.1: Decoder Derived Cross-Component Prediction [Y.-J. Chang, P.-H. Lin, V. Seregin, J.-L. Lin, M. Karczewicz (Qualcomm)]

[JVET-AG0269](https://jvet-experts.org/doc_end_user/current_document.php?id=13842) Crosscheck of JVET-AG0154 (EE2-1.1: Decoder Derived Cross-Component Prediction) [H.-J. Jhu (Kwai)] [late]

[JVET-AG0278](https://jvet-experts.org/doc_end_user/current_document.php?id=13851) Crosscheck of EE2-1.1b (Decoder Derived Cross-Component Prediction) [J. Lainema (Nokia)] [late]

[JVET-AG0157](https://jvet-experts.org/doc_end_user/current_document.php?id=13713) EE2-4.2: Fixed filter for Chroma ALF [N. Hu, M. Karczewicz, H. Wang, V. Seregin (Qualcomm)]

[JVET-AG0272](https://jvet-experts.org/doc_end_user/current_document.php?id=13845) Crosscheck of JVET-AG0157 (EE2-4.2: Fixed filter for Chroma ALF) [C.-W. Kuo, H.-J. Jhu (Kwai)] [late]

[JVET-AG0158](https://jvet-experts.org/doc_end_user/current_document.php?id=13714) EE2.4-3: Adaptive precision for luma ALF coefficients [N. Hu, M. Karczewicz, H. Wang, V. Seregin (Qualcomm), W. Yin, K. Zhang, L. Zhang (Bytedance)]

[JVET-AG0279](https://jvet-experts.org/doc_end_user/current_document.php?id=13852) Crosscheck of JVET-AG0158 (EE2.4-3: Adaptive precision for luma ALF coefficients) [N. Song (OPPO)] [late]

[JVET-AG0164](https://jvet-experts.org/doc_end_user/current_document.php?id=13720) EE2-2.10: GPM with affine prediction [K. Zhang, Z. Deng, L. Zhang (Bytedance)]

[JVET-AG0247](https://jvet-experts.org/doc_end_user/current_document.php?id=13820) Crosscheck of JVET-AG0164 (EE2-2.10 GPM with affine prediction) [P. Bordes (InterDigital)] [late]

[JVET-AG0176](https://jvet-experts.org/doc_end_user/current_document.php?id=13732) EE2: Test 2.1, 2.2, 2.5, 2.6a, 2.6e and 2.6f on LIC improvement [Y. Zhang, C.- C. Chen, H. Huang, Z. Zhang, V. Seregin, H. Wang, M. Karczewicz (Qualcomm), X. Xiu, C. Ma, N. Yan, H. -J. Jhu, C.-W. Kuo, W. Chen, X. Wang (Kwai)]

[JVET-AG0177](https://jvet-experts.org/doc_end_user/current_document.php?id=13733) Crosscheck JVET-AG0176 EE2 2.2 [A. Robert, F. Galpin (InterDigital)] [late]

[JVET-AG0270](https://jvet-experts.org/doc_end_user/current_document.php?id=13843) Crosscheck EE2-2.1, 2.2 and 2.6a (LIC improvements) [X. Xiu (Kwai)] [late]

[JVET-AG0260](https://jvet-experts.org/doc_end_user/current_document.php?id=13833) Crosscheck report of JVET-AG0176: EE2: 2.6e and 2.6f on LIC improvement [Y. Yu, Z. Xie (OPPO)] [late]

[JVET-AG0283](https://jvet-experts.org/doc_end_user/current_document.php?id=13856) Crosscheck of JVET-AG0176 EE2: Test 2.1, 2.2, 2.5, 2.6a, 2.6e and 2.6f on LIC improvement (cross-check for 2.1 and 2.6a) [K. Reuzé (InterDigital)] [late]

[JVET-AG0286](https://jvet-experts.org/doc_end_user/current_document.php?id=13859) Crosscheck of EE2-2.5 (LIC improvements) [Y. Zhang (Qualcomm)] [late]

[JVET-AG0196](https://jvet-experts.org/doc_end_user/current_document.php?id=13752) EE2-5.1/2: CABAC context initialization retraining and slice type based window offsets [F. Galpin, F. Lo Bianco, C. Salmon-Legagneur, K. Naser (InterDigital), V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-AG0210](https://jvet-experts.org/doc_end_user/current_document.php?id=13766) Crosscheck of EE2-5.1 and EE2-5.2 (CABAC context initialization retraining and slice type based window offsets) [D. Bugdayci Sansli, J. Lainema (Nokia)]

[JVET-AG0199](https://jvet-experts.org/doc_end_user/current_document.php?id=13755) EE2-1.13: Combination tests of IntraTMP search range extension with adaptive sampling and IBC with upward-extended area [Y. Kidani, H. Kato, K. Kawamura (KDDI), D. Ruiz Coll (Ofinno), K. Naser, P. Bordes, F. Le Léannec, F. Galpin (InterDigital)]

[JVET-AG0207](https://jvet-experts.org/doc_end_user/current_document.php?id=13763) Crosscheck of JVET-AG0199 (EE2-1.13: Combination tests of IntraTMP search range extension with adaptive sampling and IBC with upward-extended area) [N. Zhang (Bytedance)] [late]

[JVET-AG0276](https://jvet-experts.org/doc_end_user/current_document.php?id=13849) EE2 Test 2.6g, 2.6h, 2.6i, 2.6j: Combination of tests on LIC improvement [Y. Wang, K. Zhang, Y. He, H. Liu, L. Zhang (Bytedance), Y. Zhang, C.- C. Chen, H. Huang, Z. Zhang, V. Seregin, H. Wang, M. Karczewicz (Qualcomm), X. Xiu, C. Ma, N. Yan, H.-J. Jhu, C.-W. Kuo, W. Chen, X. Wang (Kwai)] [late]

[JVET-AG0292](https://jvet-experts.org/doc_end_user/current_document.php?id=13865) Crosscheck of JVET-AG0276 (EE2-2.6: Test 2.6i and Test 2.6j) [Z. Lv (vivo)] [late]

[JVET-AG0296](https://jvet-experts.org/doc_end_user/current_document.php?id=13869) Crosscheck of JVET-AG0276 (EE2-2.6: Test 2.6g and Test 2.6h) [L. Zhang (OPPO)] [late]

### EE2 related contributions (12)

Contributions in this area were discussed at 0730–0915 on Thursday 18 Jan. 2024 and 0505–0705 and 0720-0735 on Friday 19 Jan. (chaired by Y. Ye), at 0045+1-0055+1 on Wednesday 24 Jan. 2024 (chaired by Y. Ye).

[JVET-AG0063](https://jvet-experts.org/doc_end_user/current_document.php?id=13619) EE2-1.2 related: AR-BVP for IntraTMP merge candidates [L. Zhang, Y. Yu, F. Wang, H. Yu, D. Wang (OPPO)]

IntraTMP with merge candidates is investigated in EE2-1.2. This contribution proposes to add auto relocated block vector prediction (AR-BVP) to the IntraTMP merge candidate list to further improve the coding efficiency. Specifically, the AR-BVP candidates are inserted after the spatial candidates during the construction of IntraTMP merge candidate list.

On top of ECM11.0, simulation results of the proposed method + EE2-1.2 are reported as below:

The proposed method + EE2-1.2 over ECM-11.0 :

AI: { -0.13% Y, -0.10% U, -0.09% V, 101.0% EncT, 105.0% DecT } for Overall

{ -0.39% Y, -0.47% U, -0.10% V, 101.9% EncT, 105.2% DecT } for class F

{ -1.07% Y, -1.02% U, -0.99% V, 102.2% EncT, 109.4% DecT } for class TGM

The proposed method + EE2-1.2 over EE2-1.2 :

AI: { -0.07% Y, -0.06% U, -0.05% V, 99.7% EncT, 102.0% DecT } for Overall

{ -0.31% Y, -0.24% U, -0.24% V, 99.9% EncT, 101.5% DecT } for class F

{ -0.89% Y, -0.83% U, -0.80% V, 101.0% EncT, 106.5% DecT } for class TGM

RA results not complete when this contribution was presented.

This proposal adds template matching based sorting for the IntraTMP merge candidates, which is currently not part of ECM or EE2-1.2 (which was already adopted).

This proposal takes up to 50 candidates, and keeps 10 candidates after template matching-based sorting, the numbers of maximum candidates and kept candidates are the same as in EE2-1.2. It was commented that it would be desirable to test the performance with a smaller number of maximum candidates (i.e. less than 50) because this impacts the template matching-based sorting.

It was commented that the new aspects, i.e. template matching-based sorting and the addition of AR-BVP candidates, should be tested separately.

For AR-BVP that was tested in EE2-1.8, three variations were tested, i.e. trace path = 1, 2, and unlimited. The adopted variation was trace path = 1 given larger value of trace path did not provide much additional gain.

It was commented by one expert that given EE2-1.2 and EE2-1.8 were already adopted, this proposal seems like a natural extension of these two EE tests. There is some gain (0.07% in AI) for natural content, and more gain in screen content (0.31% for class F, 0.89% for class TGM, both in AI). The performance vs. complexity tradeoff seems to be reasonable.

JVET-AG0080 is a related contribution, with very similar method. See notes under JVET-AG0080.

[JVET-AG0256](https://jvet-experts.org/doc_end_user/current_document.php?id=13829) Crosscheck report of JVET-AG0063: EE2-1.2 related: AR-BVP for IntraTMP merge candidates [K. Naser (InterDigital)] [late] [miss]

[JVET-AG0066](https://jvet-experts.org/doc_end_user/current_document.php?id=13622) EE2-3.2 related: On Regular Residual Coding [Y. Yu, L. Xu, J. Gan, H. Yu, L. Zhang, H. Huang, F. Wang, Z. Xie, N. Song, D. Wang (OPPO)]

EE2-3.2 extended the number of the greater than X flags and moved the parity flag after the greater than X flags in the coding order. The greater than X flags and the parity flag are coded in the first coding pass in the regular residual coding process. However, the parity flag is coded as a bypass bin in the EE2-3.2. As a result, the context coded bins and a bypass-coded bin are interleaved in the first coding pass. This contribution proposes that the parity flag is coded as a context bin so that all bins in the first coding pass are coded as context coded bins. On top of EE2-3.2, the simulation results are reported as follows. There is no coding performance change compared to EE2-3.2 while the proposed method is friendly for hardware implementation.

Over ECM-11.0

AI: -0.20% (Y), -0.17% (U), -0.13% (V) for natural content

-0.07% (Y), -0.05% (U), 0.26% (V) for class F

-0.05% (Y), -0.03% (U), -0.05% (V) for class TGM

Over EE2-3.2b

AI: 0.00% (Y), 0.02% (U), 0.08% (V) for natural content

0.02% (Y), 0.09% (U), 0.18% (V) for class F

-0.03% (Y), -0.02% (U), 0.01% (V) for class TGMIt was commented that the results reported above use EE2-3.2b as anchor, and the version adopted was EE2-3.2b\* (the same design as EE2-3.2b without increasing bin budget compared to ECM-11). However, the performance difference between EE2-3.2b and EE2-3.2b\* is asserted to be really trivial.

Compared to EE2-3.2b\*, the only change proposed here in this proposal is to change parity flag from bypass coding to context coding. The proponent suggests that this prevents the interleaving of bypass coded bins and context coded bins.

It was commented that the performance difference is trivial, which means that context coding of parity flag doesn’t bring benefit in terms of coding performance.

It was commented that at this stage of exploration such detailed level of refinement of design is not crucial. Instead, our goal should be performance gain.

No action at this time.

[JVET-AG0266](https://jvet-experts.org/doc_end_user/current_document.php?id=13839) Crosscheck of JVET-AG0066 (EE2-3.2 related: On Regular Residual Coding) [H.-J. Jhu (Kwai)] [late]

[JVET-AG0068](https://jvet-experts.org/doc_end_user/current_document.php?id=13624) EE2-3.2 related: On CABAC bin budget [J. Gan, Y. Yu, H. Yu (OPPO)]

This contribution proposes two aspects related to the implementation of JVET-AG0100 (EE2-3.2) with CABAC bin budget management. The first aspect proposes to change the CABAC bin check to match the increased number of flags proposed by JVET-AG0100 and brings minor loss of {0.00%, 0.04%, 0.01%} in AI relative to EE2-3.2b. The second aspect proposes to apply the same CABAC bin budget used by EE2-3.2 to TSRC and shows minor loss of {0.00%, 0.03%, 0.02%} on CTC and gain of {-0.45%, -0.49%, -0.54%} on TGM class in AI relative to EE2-3.2b.

It was commented by the proponent that the second aspect is not relevant anymore given the adoption of EE2-3.2b\*. In terms of the first aspect, it was commented that there is no performance difference as this aspect is a very minor tweaking of the design.

The cross checker of the first aspect commented that the difference between this first aspect and what was adopted in EE2-3.2b\* is not significant at all.

No action at this time.

[JVET-AG0261](https://jvet-experts.org/doc_end_user/current_document.php?id=13834) Crosscheck of JVET-AG0068 on CABAC bin budget [X. Li (Google)] [late]

[JVET-AG0268](https://jvet-experts.org/doc_end_user/current_document.php?id=13841) Crosscheck of JVET-AG0068 (EE2-3.2 related: On CABAC bin budget) [H.-J. Jhu (Kwai)] [late]

[JVET-AG0080](https://jvet-experts.org/doc_end_user/current_document.php?id=13636) EE2-related: Extend block vector prediction for IntraTMP merge candidates [N. Qiu, J. Huo, Y. Ma, F. Yang (Xidian Univ.)]

EE2-1.7b, including EE2-1.2, EE2-1.3,EE2-1.4 and EE2-1.5.

Contribution JVET-AF0137(EE2-1.2) proposed extending the IntraTMP candidate list with all the merge candidates block vectors.

Contribution JVET-AF0226(EE2-1.3) proposed further including block-vector based prediction obtained from IntraTMP or IBC modes of the neighboring blocks in the candidate list of SGPM.

Contribution JVET-AF0275(EE2-1.4 and EE2-1.5) proposed IntraTMP extension to DIMD and LIC.

Contribution JVET-AF0129(EE2-1.8) proposed an auto relocated block vector prediction (AR-BVP) is introduced into IBC merge/AMVP candidate list construction.

This contribution merges both EE2-1.7b and EE2-1.8. In this contribution, an extension block vector prediction (EBVP) list is introduced into IntraTMP merge candidate list construction. EBVP can be derived as a guiding BV plus a BV of a reference block, located by the guiding BV.

On top of the ECM-11.0, simulation results of EE2-1.7b+EE2-1.8 are reported as below:

AI:

Class F: {-0.88%, -1.07%, -0.65% 105.4%, 106.6%, xx%};

Class TGM: {-1.63%, -1.51%, -1.54%, 103.2%, 106.7%, xx%};

On top of the ECM-11.0, simulation results of the proposed method are reported as below:

AI:

Class F: {-1.13%, -1.11%, -0.83% 106.7%, 108.3%, xx%};

Class TGM: {-2.22%, -2.08%, -2.08%, 104.4%, 107.9%, xx%};

The EBVP element of this proposal is similar to AR-BVP in EE2-1.8, but does not restrict the range of EBVP to be within the IBC search range (in comparison, the adopted AR-BVP has this restriction).

For the related contribution JVET-AG0063, the AR-BVP was also used without this restriction. The reason given by the proponents and confirmed by other experts was that, for IntraTMP merge candidates, the adjacent and non-adjacent candidates also do not have this restriction. It was suggested that restricting the BV range for all IntraTMP merge candidates (including adjacent, non-adjacent, and the proposed EBVP/AR-BVP) to be within IntraTMP search range should also be tested if we go for an EE test.

At the time of presentation, the proposal contains only results for screen content. It was reported by the proponents that although simulation for natural content is not complete, coding gain is expected.

The proponent of JVET-AG0063 confirmed that this proposal is indeed very similar, and coding gain for natural content is expected.

It was questioned why the encoder/decoder run time increased so much. The proponent said their run time was not accurate, the cross checker reported around 3% run time increase. This proposal also reported that EE2-1.7b+EE2-1.8 run time increase to be ~5%, which does not match the corresponding EE test results.

Presentation to be uploaded.

In the presentation of JVET-AG0080, the proponent showed the following luma gain on top of EE2-1.7b + EE2-1.8c in AI:

Class C: 0.07%, Class E: 0.04%, Class F: 0.25%, Class TGM: 0.61%

Other classes are not yet available at the time of presentation. During the discussion of JVET-AG0063, it was commented that the proposed method seemed to be a reasonable extension of the two EE tests that were adopted.

Investigate JVET-AG0063 and JVET-AG0080 in the next round of EE. The following elements should be tested separately:

* The addition of AR-BVP/EBVP candidates
* Template matching based sorting of IntraTMP merge candidates, including performing template matching based sorting on a smaller number of maximum candidates (i.e. less than 50)
* Restricting the BV range for all IntraTMP merge candidates (including adjacent, non-adjacent, and the proposed AR-BVP/EBVP) to be within IBC search range

[JVET-AG0298](https://jvet-experts.org/doc_end_user/current_document.php?id=13871) Crosscheck of JVET-AG0080 (EE2-related: Extend block vector prediction for IntraTMP merge candidates) [X. Li (Alibaba)] [late]

[JVET-AG0150](https://jvet-experts.org/doc_end_user/current_document.php?id=13706) EE2-related: Adaptive GPM blending [L. Zhao, K. Zhang, L. Zhang (Bytedance)]

This proposal presents the results of an enhanced GPM blending method on top of Test 2.11f, which combines Test 2.9, Test 2.10 and Test 2.11a. In Test 2.11f, a regression-based GPM blending method is proposed, where the splitting shape and the blending function are derived from the template. This proposal presents an adaptive blending width selection method based on the block size. It is reported that on top of ECM-11.0, simulation results of the proposed method are as below:

RA: -0.31 % (Y), -0.31 % (U), -0.42% (V), 102.3% (EncT), 99.8% (DecT)

On top of Test 2.11f, the results are reported as:

RA: -0.03 % (Y), -0.02 % (U), -0.14% (V), 99.9%% (EncT), 99.6% (DecT)

LB results not yet available at the time of presentation.

For larger blocks (min(width, height) >= 32), this proposal replaces the blending width candidate list in ECM-11 with a new candidate list, where each of the entry in the proposed new candidate list is twice the value of the corresponding entry in the existing candidate list. Effectively, this allows more samples to be blended for larger blocks.

It was commented that the gain seemed to be small, but the change also seemed to be small and logical.

It was commented that GPM tends to have larger gain in LB configuration, so more gain might be expected from this contribution in LB config.

Multiple experts expressed interest in investigating this in the next round of EE.

Investigate in next round of EE.

[JVET-AG0172](https://jvet-experts.org/doc_end_user/current_document.php?id=13728) EE2-related: Chroma LIC derivation with template costs [T. M. Bae, S. Deshpande (Sharp)]

This input document proposes deriving and utilizing chroma LIC flags. The derived chroma LIC flags are used to control LIC processing of chroma components separately. The derivation process of chroma LIC flags avoids the signalling of these chroma flags. When the LIC flag of the CU is enabled, chroma LIC flags are derived by utilizing the template costs that are used to derive the LIC flag of merge candidates in ECM. On top of ECM-11.0, the performance of the proposed method is reported as follows:

LB: {-0.07%, 0.26%, 0.11% }

No RA results were provided in the contribution.

In ECM-11, one LIC flag is sent to control whether LIC is enabled for all three color components. This proposal suggests to use the signaled flag to control LIC for luma component only, and use template cost to determine whether LIC should be enabled for the two chroma components.

It was asked why RA results couldn’t be made available, the proponent said that different thresholds to determine whether to enable LIC for chroma or not based on the proposed method are needed for RA config, however optimal values for those thresholds have not been found yet.

Overall, the current gain in LB is small (small luma gain and some chroma loss), and performance in RA is not available.

Further study.

[JVET-AG0208](https://jvet-experts.org/doc_end_user/current_document.php?id=13764) EE2-related: On LFNST/NSPT index signalling [M. Koo, J. Zhao, J. Lim, S. Kim (LGE)]

In JVET-AF0082, it was proposed to utilize LFNST/NSPT in inter coding, which is being investigated in the EE2 (Test-3.3). In EE Test-3.3, LFNST/NSPT is not enabled with SBT. However, it was found that an LFNST/NSPT index is signalled as 0 unnecessarily when SBT is applied to the corresponding CU. Therefore, in this contribution, it is proposed to remove the unnecessary LFNST/NSPT index signalling when SBT is applied.

The BD-rate and complexity changes relative to ECM-11.0:

* RA: -0.34%/0.12%/0.11% for Y/U/V with 103% EncT and 101% DecT
* LD: -0.XX%/-0.XX%/-0.XX% for Y/U/V with 1XX% EncT and 1XX% DecT.

The BD-rate and complexity changes relative to EE2 Test-3.3:

* RA: -0.03%/-0.11%/0.01% for Y/U/V with 100% EncT and 100% DecT
* LD: -0.09%/0.10%/0.01% for Y/U/V with 101% EncT and 100% DecT.

The above LD results (with anchor as ECM-11.0 and as EE2 Test-3.3 were obtained by enabling LFNST/NSPT for both test and anchor.

The proponent of EE2 Test-3.3 confirmed the problem reported in this contribution, and suggested that the proposed fix was a logical fix, and should be adopted along with EE2 Test-3.3. This was agreed by another expert.It was questioned why the proposed change incurs a 1% encoder run time increase in LDB. This proponent suggests this could be due to noise in measuring run time. However, given the decision made during EE2 review was only to turn on EE2 Test-3.3 for RA but not for LDB/LDP, this question about run time increase is LDB is less relevant in terms of decision making at this moment.

JVET-AG0230 is asserted to be a related contribution, which contains this SBT-related fix and some other changes related to IBC coded blocks. See notes under JVET-AG0230 as well.

Decision: adopt JVET-AG0208 (and the first aspect of JVET-AG0230, which is exactly the same). Same as the decision for EE2 Test-3.3, this should be turned on for RA config only. Further study on the performance vs. complexity impact due to inter LFNST/NSPT in the LDB/LDP config is encouraged.

[JVET-AG0227](https://jvet-experts.org/doc_end_user/current_document.php?id=13800) Crosscheck of JVET-AG0208 (EE2-related: On LFNST/NSPT index signalling) [F. Wang (OPPO)] [late]

[JVET-AG0230](https://jvet-experts.org/doc_end_user/current_document.php?id=13803) EE2-3.3 related: On Inter-LFNST [S. Puri, K. Naser, C. Bonnineau, F. Le Leannec (InterDigital)] [late]

The contribution JVET-AF0082 proposes to use LFNST/NSPT for inter blocks. In this contribution, a small fix is proposed on top of JVET-AF0082 to avoid signaling LFNST index when inter block is SBT coded. Next, the concept of LFNST/NSPT is further extended to the IBC coded blocks.

Compared to ECM-11, the average BDR gains and runtimes of the extensions reported in this contribution are as follows:

-0.02% 0.00% 0.01% 101.2% 97.5% for the AI configuration,

-0.34% 0.17% 0.12% 100.5% 96.1% for the RA configuration,

-1.66% -0.15% -0.56% 111.2% 97.5% 101.2%for the Low Delay B configuration.

Since LFNST is disabled in current CTC of ECM-11, the LDB results reported above are obtained by enabling LFNST for both the anchor and test.

Compared to JVET-AG0208, the SBT related fix is exactly the same as what was proposed in [JVET-AG0208](https://jvet-experts.org/doc_end_user/current_document.php?id=13764).In addition, enabling LFNST/NSPT for IBC coded blocks is proposed in this contribution. The proponent reported that the coding gain due to this aspect is {-0.02% 0.00% 0.01%} for {Y U V} and {101.2%, 97.5%} for {encT, decT} for the AI configuration for natural content, while the coding performance impact for RA and LDB is not readily available at the time of presentation.

As IBC was more frequently used for screen content, it was asked what is the performance impact due this IBC-related aspect on screen content. At the time of presentation, the only data available is {-0.05% -0.21% -0.08%} for {Y U V} and {100.9% 99.7%} for {encT, decT} for class F in AI config. Other data is not available. Given IBC was used more frequently in screen content, higher encoder run time increase should be expected compared to natural content. However, contradictory observation was made in the reported data.

The proponent of EE2 Test-3.3 commented that this IBC-related aspect seems interesting, and when they tried it before they were not able to obtain coding gain.

It was asked how LFNST/NSPT is handled for intraTMP coded blocks. DIMD is used to derive the intra prediction mode to determine LFNST/NSPT index. This is already done in ECM-11.

It was commented by multiple experts that the gain achievable by the IBC-related aspect is rather small, and not justified by the encoder run time increase of ~1%.

Further study of this second aspect is encouraged.

[JVET-AG0309](https://jvet-experts.org/doc_end_user/current_document.php?id=13882) Crosscheck of JVET-AG0230 (EE2-3.3 related: On Inter-LFNST) [M. Koo (LGE)] [late]

[JVET-AG0231](https://jvet-experts.org/doc_end_user/current_document.php?id=13804) EE2-1.7b related: unrestricted 1.7b performances [F. Le Léannec, K. Naser, T. Dumas, Y. Chen, M. Radosavljević, T. Poirier (InterDigital)] [late]

The proponent mentioned that this contribution is only meant to provide additional information on EE2-1.7b, and there is no need for presentation.

[JVET-AG0233](https://jvet-experts.org/doc_end_user/current_document.php?id=13806) EE2-related: Adaptive precision for CCALF coefficients [N. Hu, M. Karczewicz, V. Seregin, H. Wang (Qualcomm)] [late]

In adaptive loop filter of ECM-11.0, when filtering a luma sample with an on-line derived filter, 7 bits are used to represent the fractional part of a coefficient. Similarly, when filtering a chroma sample with a cross component adaptive loop filter (CCALF), 7 bits are used to represent the fractional part of a coefficient. Each CCALF coefficient can only be a value in the form of power of 2. In test EE2-4.3, the number of bits used to represent the fractional part of a luma coefficient is adaptive from 5 to 8, inclusively. A two-bit syntax element is signalled for each set of luma filter set to indicate the number of bits. In this contribution, the idea of EE test is applied to CCALF, the number of bits used to represent the fractional part of a CCALF coefficient is adaptive from 7 to 10, inclusively. In addition, the power of 2 constraint is removed. The proposed method was implemented on top of ECM-11.0, with common test condition, simulation results are reported as follows.

AI: {Y U V} {0.00% -0.25% -0.30%} {EncT DecT} {100.0% 99.9%}RA and LDB results not complete at the time of presentation.

It was asked if the constaint for CCALF coefficients to be power of 2 is maintained in this proposal. The proponent said that this constaint is removed for the results reported, although the constaint can be kept if desired. It was questioned what would be the performance difference by just removing the coeffient constraint of being power of 2. The proponent has not tested this.

JVET-AG0065 is reported to be a related contribution.

Investigate in the next round of EE. Study the effect of coefficient constraint removal and adaptive precision separately for CCALF. Also consider possible combination with JVET-AG0065.

In later meeting discussion, further notes were added to JVET-AG0065. See notes also under JVET-AG0065.

[JVET-AG0304](https://jvet-experts.org/doc_end_user/current_document.php?id=13877) Crosscheck of JVET-AG0233 (EE2-related: Adaptive precision for CCALF coefficients) [N. Song (OPPO)] [late] [miss]

[JVET-AG0237](https://jvet-experts.org/doc_end_user/current_document.php?id=13810) EE2-3.3 related: Fix on LFNST/NSPT index signalling [M. Koo, J. Zhao, J. Lim, S. Kim (LGE)] [late]

In EE2-3.3 (JVET-AG0061: Utilizing LFNST/NSPT for inter coding), it is expected that LFNST/NSPT index signalling for inter coding is aligned with intra case and performed only when last non-zero coefficient of Luma component is not at DC position. However, in current EE2-3.3 SW, an LFNST/NSPT index can be signalled even when last coefficient of Luma component is at DC position. In this contribution, a fix for this problem is proposed.

The BD-rate and complexity changes relative to ECM-11.0:

* RA: -0.XX%/-0.XX%/-0.XX% for Y/U/V with 10X% EncT and 10X% DecT
* LD: -0.XX%/-0.XX%/-0.XX% for Y/U/V with 1XX% EncT and 1XX% DecT.

The BD-rate and complexity changes relative to EE2-3.3:

* RA: -0.XX%/-0.XX%/-0.XX% for Y/U/V with 10X% EncT and 10X% DecT
* LD: -0.XX%/-0.XX%/-0.XX% for Y/U/V with 1XX% EncT and 1XX% DecT.

The above two LD results were obtained by enabling LFNST/NSPT for both test and anchor. It was asked if any complete results for RA are available from the proponent and two cross-checkers. Such is not the case at the time of presentation.

This proposal suggests to determine DC\_cond based only on luma DC condition, and not on Cb/Cr DC condition as done in EE2-3.3. The proponent of EE2-3.3 (and another expert) commented that this suggested change seems to be a logical fix.

In a later discussion, completed simulation results were provided as follows:

The BD-rate and complexity changes relative to ECM-11.0:

* RA\*: -0.30%/0.04%/0.11% for Y/U/V with 102% EncT and 100% DecT
* LD: -0.86%/0.06%/0.01% for Y/U/V with 104% EncT and 101% DecT.

The BD-rate and complexity changes relative to EE2-3.3:

* RA\*: 0.01%/-0.19%/0.02% for Y/U/V with 100% EncT and 100% DecT
* LD: 0.01%/-0.46%/-0.35% for Y/U/V with 100% EncT and 100% DecT.

The simulation results (\* = QP 27 of ParkRunning not complete yet and copied from anchor) show some coding loss of 0.02% in class A1 and class B, and coding gain of 0.01% in class C.

It was commented by one expert that the proposed change is a simplification compared to the adopted method in EE2-3.3. However, the proponent asserted that this is not a simplification, and that it rather targets redundant signaling.

As supplemental results, the proponent reported BD-rate and complexity changes relative to EE2-3.3 + JVET-AG0208 (both of which had been adopted) to be {Y, U, V: -0.01%, -0.08%, -0.10%} and {EncT, DecT: 100%, 100%}. However, given this result was not cross checked and could not be cross checked before this meeting ends, it was suggested to investigate this in EE.

Investigate in EE.

[JVET-AG0305](https://jvet-experts.org/doc_end_user/current_document.php?id=13878) Crosscheck for JVET-AG0237 EE2-3.3 related: Fix on LFNST/NSPT index signalling [[S. Puri](mailto:saurabh.puri@interdigital.com), K. Naser (InterDigital)] [late]

[JVET-AG0306](https://jvet-experts.org/doc_end_user/current_document.php?id=13879) Crosscheck of JVET-AG0237 (EE2-3.3 related: Fix on LFNST/NSPT index signalling) [F. Wang (OPPO)] [late]

[JVET-AG0310](https://jvet-experts.org/doc_end_user/current_document.php?id=13883) EE2-1.14-related: EIP with bias and clipping [J. Lainema, P. Astola (Nokia)] [late]

This contribution proposes to include a bias term to the EIP (extrapolation-based intra prediction of EE2-1.14) filters. The bias term is proposed to substitute one of the sample inputs of the EIP filters to keep the number of coefficients and thus the complexity of the filters unchanged with respect to EE2-1.14. Furthermore, output of the EIP filter is proposed to be clipped to the value range of the reference samples instead of the full sample bitdepth range. It is asserted that inclusion of these aspects will stabilize and improve EIP performance.

It is reported the proposed filter modifications achieve the following coding gains for AI configuration when implemented on top of EE2-1.14b (adopted to ECM this meeting) {for Y, U, V}:

AI - {-0.20%, -0.14%, -0.18%} for original EE2-1.14b  
AI - {-0.23%, -0.12%, -0.16%} EE2-1.14b with a bias term  
AI - {-0.21%, -0.17%, -0.11%} EE2-1.14b with clipping  
AI - {-0.25%, -0.15%, -0.19%} EE2-1.14b with both bias term and clipping

It is further asserted the ECM runtimes and peak memory consumption should be unchanged over the EE test as the filter size remains unchanged and there is already clipping to the full sample range in the original EIP.

The proposal replaces the extrapolation filter used in EE2-1.14b with a form that includes a bias term. Filter tap is not changed, the sample used in EE2-1.14b that is farthest away from the current sample is replaced with the bias term.

Aside from the bias term, a clipping operation is introduced to the EIP in EE2-1.14b.

Results showing the performance impact due to each of these two elements are provided in the contribution, and shown to be additive.

It was asked if adding a bias term (rather than replacing the farthest sample) could bring further benefit. The proponent tried this but reported that practically there was no further gain.

Investigate in EE. Study the impact of each of the two elements separately on top of the next version of ECM.

### ECM modifications and software improvements beyond EE2 (33)

Contributions in this area were discussed at 2105–2305 on Thursday 18 Jan. 2024 (chaired by Y. Ye), at 0735–0925 and 2100–2310 on Friday 19 Jan. 2024 (chaired by Y. Ye and M. Wien), at 2100-2310 on Monday 22 Jan. 2024 (chaired by Y. Ye and M. Wien), at 2330-0045+1 on Wednesday 24 Jan. 2024 (chaired by Y. Ye).

#### General (2)

[JVET-AG0096](https://jvet-experts.org/doc_end_user/current_document.php?id=13652) Non-EE2: On temporal buffer handling [Z. Deng, K. Zhang, L. Zhao, L. Zhang (Bytedance)]

In ECM-11.0, some coding tools can access larger than one CTU of temporal buffer, but some coding tools follow the one CTU temporal buffer restriction. This contribution proposes to remove the one CTU temporal buffer restriction for all coding tools in ECM. The impact on coding efficiency and runtimes of the proposed test over ECM-11.0 are reportedly for {Y, U, V, EncT, DecT}:

RA { -0.02%, -0.04%, -0.01%, 100.3%, 100.4% }; LDB { -0.21%, 0.10%, 0.10%, 100.3%, 100.2% }

Presentation deck not available at time of presentation. Proponent will upload.

It is asserted that in ECM-11.0, different tempopral motion tools use different temporal buffer restrictions in the collocated pictures, some tools use a one-CTU-row buffer, and other tools (i.e. the temporal CCP candidate derivation process in intraCCP merge and interCCP merge modes) use the entire collocated picture.

The proponent’s main motivation is to solve the inconsistency issue, but considers the gain to be “free”.

However, it was commented that this inconsistency in ECM-11.0 might not be the most important at this stage of exploration, and there are two ways to solve this inconsistency (if we do decide to solve it): 1. As proposed, remove restriction for all tools, and 2. Impose the one-CTU-row restriction on intraCCP merge and interCCP merge.

It was commented by one expert that this inconsistency could cause different temporal prediction tools to behave differently (e.g. those not subject to the constraint might appear to provide more gain), which could affect our ability to accurately evaluate coding tool performance.

Multiple experts support investigating this in EE to better understand the impact on coding performance due to the buffer constaints.

Investigate in EE, both methods (removing temporal buffer constaint for all relevant tools, and imposing the one-CTU-row restriction on all relevant tools) are to be studied.

[JVET-AG0291](https://jvet-experts.org/doc_end_user/current_document.php?id=13864) Crosscheck of JVET-AG0096 (Non-EE2: On temporal buffer handling) [[Z. Lv (vivo)](mailto:zhuoyi.lv@vivo.com)] [late]

[JVET-AG0120](https://jvet-experts.org/doc_end_user/current_document.php?id=13676) Non-EE2: On line buffer restriction [Z. Deng, K. Zhang, L. Zhang (Bytedance)]

In ECM-11.0, many coding tools can access more than one line buffer, but a few coding tools still follow the limitation of one line buffer restriction as in VVC. This contribution proposes to remove the line buffer restriction for all coding tools in ECM. The impact on coding efficiency and runtimes of the proposed test over ECM-11.0 are reportedly for {Y, U, V, EncT, DecT}:

AI { -0.05%, 0.06%, 0.02%, 100.4% , 100.3%}; RA { -0.01%, -0.02%, -0.05%, 100.4%, 100.0%}

Presentation deck not available at time of presentation. Proponent will upload.

This proposal reports that the reconstruction sample and motion used by various ECM coding tools are not constrained in the same way, some tools use a line buffer that consists of one line outside of current CTU (which is the same as in VVC) and other tools use much bigger line buffer size. For example, for non-adjacent spatial candidates for intra MPM and inter merge mode, up to four CTU rows are used.

For this proposal, if we apply the same restriction used in VVC to all coding tools, then it would break the “spirit” of many ECM tools.

The proposal lists three intra coding tools (extended MRL, template based MRL, and intra luma fusion) that are currently using the VVC-time line buffer restriction. When line buffer restriction is loosened for these three tools, small coding gain (0.05% for AI and 0.01% for RA) is achieved. Up to 12 reference lines outside of the current CTU can be used.

It was commented that if we would put this in EE, there does not seem much to study. However, it was commented by one expert that we had many intra coding related adoptions this time, and it would be interesting to see the performance impact of this proposal on the next ECM version.

Investigate in EE.

[JVET-AG0263](https://jvet-experts.org/doc_end_user/current_document.php?id=13836) Crosscheck of JVET-AG0120 (Non-EE2: On line buffer restriction) [H. Huang (OPPO)] [late]

#### Intra and CIIP (13)

[JVET-AG0075](https://jvet-experts.org/doc_end_user/current_document.php?id=13631) AHG12: Adaptive MRL Fusion [S. Blasi, J. Lainema (Nokia)]

This contribution proposes to modify MRL Fusion in ECM. The MRL indexes used to obtain the additional MRL Fusion predictors are selected depending on the current MRL index. Also, the fusion weights are determined adaptively for each block using a template-based search. The impact on coding efficiency and runtimes of the proposed adaptive MRL Fusion over ECM-11.0 are reportedly {for Y, U, V, EncT, DecT, EncVmPeak, DecVmPeak}:

AI {-0.07%, -0.05%, -0.02%, 100.6%, 101.4%, 100.5%, 100.9%}

It was asked what kind of line buffer restriction is applied to the proposed adaptive MRL fusion method. The proponent reported that the same restriction as in the current ECM (i.e. one line outside of current CTU) is used. It was further commented that the interaction between this proposed method and JVET-AG0120 should be investigated, if we decide to put this in EE.

This proposal replaces the MRL fusion method in ECM, and no additional signaling is introduced. No other MRL-related tools (e.g. template based MRL) are modified.

There are the two elements in this proposal, the second element which uses template-based search to determine the fusion weights adaptively, should be tested separately, vs. a fixed weight strategy.

Investigate in EE, results are to be provided for the following:

* The first element (3-line predictor) in this proposal
* Both elements (3-line predictor + adaptive weights) in this proposal
* Combination with JVET-AG0120 which removes the line buffer restriction

[JVET-AG0078](https://jvet-experts.org/doc_end_user/current_document.php?id=13634) AHG12: Intra-prediction using Merged Histogram of Gradients [S. Blasi, I. Zupancic, P. Astola, J. Lainema (Nokia)]

This contribution proposes a new intra-prediction mode based on the computation of a Merged Histogram of Gradients using information from neighbouring blocks. The Merged Histogram of Gradients is then used to derive a set of intra modes and weights that are blended together to compute the intra prediction for a block. The impact on coding efficiency and runtimes of the proposed method over ECM-11.0 are reportedly {for Y, U, V, EncT, DecT, EncVmPeak, DecVmPeak}:

AI {-0.10%, 0.00%, 0.00%, 100.6%, 99.6%, 100.4%, 100.0%}

RA {-0.04%, 0.02%, -0.03%, 99.6%, 100.7%, 99.8%, 100.4%}proposal shows interesting gain especially for 4K sequences, with limited impact on run time. The proponent mentioned that aside from normative restriction on the use of the new mode (block size, availability of MIMD and DIMD neighbor), there is also some fast encoding algorithm in the proposal that achieves better performance vs. complexity tradeoff than bruteforce RDO.

Compared to DIMD merge, this proposal includes a new algorithm that composes a synthetic histogram of gradients based on neighboring blocks (which do not have to be DIMD blocks). Both adjacent and non-adjacent neighboring blocks are used.

Investigate in EE. Results are requested for using the normative restriction as proposed, and for removing the normative restriction.

[JVET-AG0084](https://jvet-experts.org/doc_end_user/current_document.php?id=13640) AHG12: DIMD Merge List [M. Blestel, P. Andrivon (Ofinno)]

Contribution JVET-AG0084 proposes to extend the DIMD merge process that was introduced in JVET-AE0071, refined in JVET-AF0120 and JVET-AF0106, by introducing the generation of a DIMD merge candidates list. The DIMD merge list is built reusing the Decoder-side Intra Mode Derivation (DIMD) information and merged DIMD information (as described in JVET-AF0120) from spatial adjacent and non-adjacent neighboring blocks. On top of ECM-11.0, the reported PSNR-Y, Cb, Cr, BD-rate and {EncT, DecT} results are as follows:

* AI: -0.12%, 0.00%, 0.05% {101.9%, 100%}

This proposal creates a DIMD merge candidate list consisting of a set of DIMD parameters from adjacent, non-adjacent and merged DIMD neighbors, whereas JVET-AG0078 introduces a new way to calculate the DIMD parameters for a current DIMD block. Therefore, some combination would be interesting to be tested.

Compared to DIMD merge mode proposed two meetings ago, this proposal includes more neighbors (e.g. non-adjacent) and achieves higher coding gain.

It was asked how large the candidate list is, this current proposal only uses 2 candidates. The proponent commented that higher coding gain can be achieved with more candidates, though it will incur higher complexity. One expert commented that the number of candidates may be made to depend on the block size to achieve a better performance vs complexity tradeoff.

Investigate in EE. Combination test with JVET-AG0078 should be tested.

[JVET-AG0254](https://jvet-experts.org/doc_end_user/current_document.php?id=13827) Crosscheck of JVET-AG0084 (AHG12: DIMD Merge List) [R. G. Youvalari, M. Abdoli (Xiaomi)] [late]

[JVET-AG0106](https://jvet-experts.org/doc_end_user/current_document.php?id=13662) AHG12: TIMD merge mode [R. G. Youvalari, M. Abdoli (Xiaomi)] [late]

This contribution proposes a new intra prediction method called as TIMD merge mode for improving the performance of ECM-11.0. The proposed method inherits the TIMD modes and its associated information (fusion flag, fusion weights and wide-angle conditions of TIMD modes) from previously TIMD or TIMD merge coded blocks. The difference of the proposed method compared to the TIMD mode of ECM-11.0 is related to the selection prcess of the intra modes for TIMD. The block prediction aspect of the proposed method is the same as in TIMD mode. The impact on coding efficiency and runtimes of proposed method over ECM-11.0 are reportedly {for Y, U, V, EncT, DecT}:

CTC classes: AI { -0.09%, 0.05%, 0.04%, 103.6%, 99.7%}

Class F: AI { -0.02%, 0.00%, 0.04%, 103.9%, 100.3%}, RA { -0.06%, -0.13%, -0.15%, 100%, 10x%}

Class TGM: AI { 0.01%, -0.01%, -0.02%, 104.1%, 99.9%}, RA { -0.01%, -0.04%, -0.03%, 101%, 10x%}

Currently in ECM, there is no TIMD merge mode. This proposal adds the TIMD merge, which allows to use TIMD related information from previously coded block to improve coding efficiency.

It was commented that currently the encoder run time is a bit high, and it was suggested to develop non-normaive fast encoding algorithm to reduce encoder run time and improve tradeoff.

Currently, two TIMD merge candidates are used and one of them is selected based on SATD. RDO is performed on the TIMD merge candidate thus selected to compare with other modes (regular TIMD and other intra coding modes). It was suggested to investigate the performance vs complexity tradeoff using both TIMD merge candidates in RDO process in combination with other fast encoding algorithms.

In terms of signaling, one flag is signaled to indicate whether the TIMD merge mode is used, and another flag is signaled to indicate which one of the two merge candidates is used.

Investigate in EE. Fast encoder algorithm should be investigated to reduce encoder run time without too much penalty on coding performane.

[JVET-AG0113](https://jvet-experts.org/doc_end_user/current_document.php?id=13669) Non-EE2: FIBC Extension [J. Kim, J. Kang, H. Han, H. Choi (HNU), W. Lim, S.-C. Lim (ETRI)]

The filtering methods for IBC prediction, specifically FIBC and IBC-LIC, are included in ECM-11.0. While IBC-LIC supports multiple template shapes and multiple models, FIBC utilizes only L-shaped templates and single model. This contribution introduces an extension method for FIBC to enable support for multiple template shapes and multiple models. Experimental results of the proposed method on top of ECM-11.0 are reported as follows:

AI: Class F {-0.03%, -0.11%, 0.26%; 103.9%, 99.9%}, Class TGM {-0.07%, -0.03%, -0.03%; 104.3%, 99.9%};

RA: Class F {-0.09% , -0.21%, -0.10%, 101.7%, 99.6%}, Class TGM {-0.07%, -0.08%, -0.02%; 100.9%, 100.3%}

It was commented that the proposed method seemed to be a simple extension (i.e. applying something that is already being done in IBC-LIC to FIBC), but it does not seem to be effective. The coding gain is only for screen content and is very small.

Further study.

[JVET-AG0293](https://jvet-experts.org/doc_end_user/current_document.php?id=13866) Crosscheck of JVET-AG0113 (Non-EE2: FIBC Extension) [H.-J. Jhu (Kwai)] [late]

[JVET-AG0121](https://jvet-experts.org/doc_end_user/current_document.php?id=13677) Non-EE2: Block vector guided LUT for chroma prediction [J. Huo, X. Hao, M. Chen, N. Qiu, Z. Zhang, Y. Ma, F. Yang (Xidian Univ.), M. Li, F. Wang, J. Ren (OPPO)]

In this contribution, a new chroma intra prediction mode, block vector guided look-up table (BVG-LUT) mode, is proposed for screen content. This method is used for chroma components when dual tree is activated in intra slice. BVG-LUT uses block vectors of the co-located IBC or intraTMP coded luma blocks to identify the reference area for deriving the cross-component LUT, and uses the co-located luma samples for chroma prediction. For non 4:4:4 sequences, the luma samples are derived by multiple down-sampling filters. It is reported that the coding performance of ECM 11.0 is as below:

For AI configuration:

F: -0.02%, -0.49 %, and -0.44 %, with 103.4 % EncT, 101.5 % DecT.

TGM: -0.59%, -1.01 %, and -1.35 %, with 102.6 % EncT, 98.8 % DecT.

For RA configuration:

F: -0.04%, -0.22 %, and -0.20 %, with 101.0 % EncT, 99.8 % DecT.

TGM: -0.20%, -0.40 %, and -0.53 %, with 100.9 % EncT, 100.0 % DecT.

This proposal utilizes block vectors from IBC and intraTMP to look up chroma values as prediction for chroma components, and this is signaled as a new chroma prediction mode. It was commented by the proponent that although IBC and intraTMP are enabled for natural content in CTC, this proposed change is not expected to bring any performance gain for natural content, and is therefore only proposed for screen content. It was asked if this LUT is built for each chroma component separately. This was confirmed by the proponent.

It was questioned how the LUT behaves for higher bit-depth content such as 10-bit, this has not been experimented because the screen content CTC contains only 8-bit sequences.

It was commented that the coding gain vs. runtime tradeoff as currently reported does not appear to be attractive. The proponent suggested that fast encoding algorithm might be used to reduce encoder runtime.

Currently, one of multiple downsampling filters is selected, which could cause encoding time increase. According to the proponent, using a fixed downsampling filter could achieve similar gain and incur less encoding time increase.

Further study.

[JVET-AG0299](https://jvet-experts.org/doc_end_user/current_document.php?id=13872) Crosscheck of JVET-AG0121 (Non-EE2: Block vector guided LUT for chroma prediction) [X. Li (Alibaba)] [late]

[JVET-AG0138](https://jvet-experts.org/doc_end_user/current_document.php?id=13694) Non-EE2: Chroma intra prediction mode reordering [X. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba)]

Discussion chaired by M. Wien

This contribution proposes to reorder the non-CCP chroma intra prediction modes. First, a chroma intra prediction mode list is constructed. Then the modes in the list are applied to predict the collocated luma block and reordered based on SAD cost. The first N modes in the list are allowed to predict the current chroma block. It is reported that on top of ECM-11.0, the overall coding performance impact for {Y, U, V, EncT, DecT} is {-0.05%, -0.01%, -0.05%, 100.3%, 101.5%} for AI and {-x.xx%, -x.xx%, -x.xx%, xxx.x%, xxx.x%} for RA.

It was reported that the simulation results for Classes A1 and A2 for RA were not finished before the presentation. The proponent asserted that no changes with respect to syntax elements has been applied. It was asserted that the ordering of the mode list remains stable also in the case of e.g. transmission losses. It was remarked that it seems that the gains mainly come from the luma. It was asserted that for reording, the collocated luma block is used, i.e. the pixels at the chroma block boundary are not relevant.

Two experts expressed support for having an EE on this proposal.

The maximum size of the list was requested. It was commented that restrictions on the selection of modes of reference blocks would apply. It was commented that the reordering of modes in the list and the addition of modes to the list should be investigated as separate aspects in a corresponding EE.

Investigate in EE as described in the previous paragraph.

[JVET-AG0273](https://jvet-experts.org/doc_end_user/current_document.php?id=13846) Crosscheck of JVET-AG0138 (Non-EE2: Chroma intra prediction mode reordering) [Z. Deng (Bytedance)] [late]

[JVET-AG0141](https://jvet-experts.org/doc_end_user/current_document.php?id=13697) AHG12: Occurrence-Based Intra Coding (OBIC) [R. G. Youvalari, M. Abdoli (Xiaomi)] [late]

This contribution proposes a new intra prediction method called as Occurrence-based Intra Coding (OBIC) for improving the performance of ECM-11.0. The proposed method derives the intra prediction modes based on the occurrence of the intra modes in the spatial neighborhood of the block. The proposed method uses blending of up to five intra modes, with highest occurrence, along with the planar mode to obtain the prediction of the block. The blending weights are decided based on each mode’s occurrence in the neighborhood.

The impact on coding efficiency and runtimes of proposed method over ECM-11.0 are reportedly {for Y, U, V, EncT, DecT}:

CTC classes: AI { -0.12%, 0.03%, -0.04%, 103.2%, 100.1%}

Class F: AI { -0.02%, 0.03%, 0.10%, 102.3%, 101.4%}, RA { -0.08%, -0.22%, -0.30%, 101.2%, 10x%}

Class TGM: AI { -0.02%, -0.05%, -0.07%, 101.6%, 102.2%}, RA { 0.02%, 0.02%, 0.06%, 100.1%, 10x%}

This proposal adds OBIC mode as a submode of DIMD, where when DIMD is used, an additional flag is signaled to indicate whether histogram of gradients (as in current ECM) or histogram of occurrence (as in the proposal) is used. The current implementation uses a bruteforce RDO to decide if the new mode is used, therefore causing encoder run time increase of approximately 3%. The proponent commented that fast encoding algorithm can be used to achieve a better performance vs. runtime tradeoff. JVET-AG0084 and JVET-AG0078 are reported to be related contributions as they both propose submodes of DIMD. However, these two proposals continue to use histogram of gradients as in current ECM whereas this proposal uses histogram of occurrence. It was commented that combination tests should be performed to understand if there is overlap in terms of performance gain among these proposals.

Investigate in EE. Proponent is requested to investigate encoder run time reduction. Also, combination tests with JVET-AG0084 and JVET-AG0078 should be investigated.

[JVET-AG0186](https://jvet-experts.org/doc_end_user/current_document.php?id=13742) Non-EE2: FIBC Extension [H.-J. Jhu, X. Xiu, W. Chen, C.-W. Kuo, N. Yan, C. Ma, X. Wang (Kwai)]

This contribution proposes to support multiple template shapes and multiple models for the FIBC. Compared to ECM-11.0 anchors, simulation results are summarized as follows:

For class F sequences:

AI { -0.13%, -0.12%, -0.07%, 101.2%, 99.2% }, RA { -0.13%, -0.16%, -0.31%, 100.3%, 99.9% }

For TGM sequences:

AI { -0.06%, -0.02%, -0.09%, 101.0%, 100.1% }, RA { -0.08%, -0.13%, -0.05%, 100.4%, 99.9% }

Conceptually, this proposal is very similar to JVET-AG0113. It reports somewhat better performance vs. complexity tradeoff compared to JVET-AG0113, and the proponent claims that it could be due to better encoder implementation. It was commented that the performance gain in this contribution, even though better than JVET-AG0113, is still rather small, esp. considering it is only for screen content.

Further study.

[JVET-AG0303](https://jvet-experts.org/doc_end_user/current_document.php?id=13876) Crosscheck of JVET-AG0186 (Non-EE2: FIBC Extension) [W. Lim, S.-C. Lim (ETRI), J. Kim, H. Choi (HNU)] [late]

[JVET-AG0193](https://jvet-experts.org/doc_end_user/current_document.php?id=13749) Non-EE2: Enhancements on IntraTMP [W. Chen, X. Xiu, C. Ma, H.-J. Jhu, C.-W. Kuo, N. Yan, X. Wang (Kwai)]

This contribution proposes to improve the IntraTMP from two aspects. In the first aspect, it is proposed to use the SATD as the cost function for the refinement process of integer BV search. In the second aspect, it is proposed to use the reordering method to reduce the signaling overhead of intraTMP BV. The simulation results on top of ECM-11.0 are summarized as follows for AI config:

natural content (Y/U/V): -0.08 %/ -0.04 %/ -0.04 %

class F (Y/U/V): 0.02 %/ 0.03 %/ 0.07 %

class TGM (Y/U/V): -0.09 %/ -0.08 %/ -0.05 %

Approximately, EncT 102% and DecT 101% are reported. The proponent reported that the encoder/decoder run time increase is likely more due to replacing SAD with SATD in cost function in the refinement stage, rather than the reordering of BV.

It was asked what is the performamce gain from each of these two elements separately. The proponent reported that roughly half of the gain came from each element. JVET-AG0243 is reported to be a related contribution, which uses both SAD and SATD and calculates two sets of block vectors.

It was asked if there is anyway to reduce encoder run time. The proponent commented that early termination methods could be employed to achieve this goal (of complexity reduction).

It was commented that although this proposal is for intraTMP, the gain for screen content is not interesting at all.

Investigate in EE. Separate tests on each of the two proposed elements are requested. Though the gain is not targeting screen content, results for class F and class TGM are requested (to examine whether the proposed elements have negative impact on screen content).

[JVET-AG0197](https://jvet-experts.org/doc_end_user/current_document.php?id=13753) Non-EE2: Matrix based intra prediction replacing conventional intra modes [B. Ray, H. Wang, V. Seregin, M. Karczewicz, P. Garus, G. Verba(Qualcomm)]

In this proposal, a matrix based intra prediction replacing some directional intra modes is proposed where the weights are intra mode and block shape dependent which are applied to L shaped causal template to generate final prediction.

On top of the ECM-11.0, simulation result of the proposed method is:

AI: -0.46%, -0.42%, -0.42%, EncT: 102.2%, DecT: 100.8%.

The proposed matrix based intra prediction replaces the traditional intra prediction process (not DIMD or TIMD). It does not introduce new modes or additional signaling. It is applied to luma only. The matrices are trained using the NNVC training set. It was asked if this is at risk of overfitting. The proponent suggests that different training strategies may be used during EE round to check for overfitting. It was commented that the training process (training script etc) should be part of EE.

In total, how many matrics? Though the proponent does not have an accurate number, an estimate of 5M of total memory to store matrices was given (each coefficient is 16-bit in current implementation). How many multiplications? This can be provided during EE round.

It was asked if downsampling of the reference template is performed in a similar manner as MIP. Currently no. But this could be beneficial.

Currently, the proposed method is only applied to blocks up to 16x16. It was asked if applying this to larger blocks might bring benefits as well. This could be investigated during EE.

Investigate in EE. Training script to be part of the EE test. Detailed information on computation complexity (# matrices, # multiplications for different block sizes) is requested. Downsampling of the reference template, applying this to larger blocks, etc. should be explored as well (to the extent practical).

[JVET-AG0202](https://jvet-experts.org/doc_end_user/current_document.php?id=13758) Non-EE2: Geometry partitioning mode with inter prediction and intra block copy Y. Wang, K. Zhang, L. Zhang (Bytedance)

This contribution presents a method of geometry partitioning mode (GPM) with inter prediction and intra block copy (GPM-inter/IBC). In GPM-inter/IBC, the two sub-partitions divided geometrically are generated using inter prediction and IBC. On top of ECM-11.0, simulation results of the proposed method are reported as below:

RA: Class F {-0.03%, 0.02%, -0.07%; 100.7%, 99.9%}, Class TGM {-0.39%, -0.39%, -0.41%; 99.9%, 100.5%};

LDB: Class F{-0.21%, -0.37%, -0.12%; 101.1%, 101.5%}, Class TGM {-0.77%, -0.79%, -0.79%; 100.2%, 98.5%}.It was asked if this proposal has performance gain for natural content. The proponent said only very minor gain in RA (slightly larger gain in LDB) could be achieved, and therefore it is proposed only for screen content.

This proposal uses the same IBC candidate list as in current ECM (which could contain block vector from intraTMP) to signal the block vector used to generate the IBC prediction.

It was commented that the gains are rather small, esp. considering they are only for screen content.

Further study.

[JVET-AG0243](https://jvet-experts.org/doc_end_user/current_document.php?id=13816) AHG12: Additional Metric for IntraTMP [K. Naser, F. Le Léannec, T. Poirier, H. Guermoud, T. Dumas (InterDigital)] [late]

In IntraTMP mode, the best N block vectors are found using SAD metric. Specifically, the candidate block vectors are evaluated in template matching process using SAD cost. This contribution proposes further including additional metric, to find a new set of block vectors. A flag is signaled to the decoder to indicate the metric choice. On top of ECM-11.0, the following results are obtained:

SATD as additional metric  
AI: -0.10% -0.04% -0.06% 102.7% 102.7%

MR-SAD as additional metric  
AI: -0.09% -0.04% -0.11% 101.7% 100.8%

The proposal adds a new metric (either SATD or MR-SAD) in cost function for intraTMP. The encoder selects which one is used (existing SAD metric or the additional SATD/MR-SAD metric) and signals the choice to decoder. It was asked whether profiling has been done to check which metric is used how often. The proponent has not studied this aspect.

JVET-AG0193 is a related contribution, but does not signal which metric is used. It just replaces SAD with SATD in the intraTMP refinement stage.

Investigate in EE. Reduction of encoder/decoder run time increase should be investigated. Interaction with JVET-AG0193, JVET-AG0063 and JVET-AG0080 should be studied, e.g. in the form of combination tests.

#### Inter (7)

[JVET-AG0073](https://jvet-experts.org/doc_end_user/current_document.php?id=13629) Non-EE2: Chained motion vector prediction [Y. Kidani, H. Kato, K. Kawamura (KDDI)]

This contribution proposes a chained motion vector (MV) prediction as a new inter merge candidate list construction. Inspired by auto-relocated block vector (BV) prediction proposed in JVET-AF0129, the chained MV prediction candidates can be derived as the accumulation of the recursively traced MVs and BVs based on the pre-derived MVs. The chained MV prediction candidates are inserted after the HMVP prediction candidates.

On top of the ECM-11.0, simulation results {Y, U, V, EncT, DecT} of the proposed method for overall test sequences in CTC are reported below:

The length of the chained MV prediction trace depth is 1, as like EE2-1.8a:

RA: {-0.10%, -0.20%, -0.27%, 99.9%, 100.5%};

LDB: {-0.42%, -0.20%, -0.32%, 100.1%, 102.1%};

LDP: {-0.28%, -0.46%, -0.54%, 100.1%, 101.7%};

No constraint for the length of the chained MV prediction trace depth, as like EE2-1.8c:

RA: {-0.12%, -0.19%, -0.20%, 99.7%, 100.7%};

LDB: {-0.38%, -0.04%, -0.16%, 100.1%, 102.5%};

LDP: {-0.28%, -0.43%, -0.25%, 100.1%, 102.1%}.

With the proposed method, it could become necessary to fetch motion information (motion vectors, reference indices, etc) from reference pictures that are not collocated pictures. This means that all reference pictures will need to store motion data and effectively become collocated pictures. It was commented that an EE test should be designed to investigate the tradeoff between coding performance and additional storage required for motion information.

It was asked if the chained motion vector could go outside of the collocated CTU row in the collocated picture. The proponent said that clipping is performed to ensure that only motion from collocated CTU row is utilized. JVET-AG0096 proposes to remove the constraint on one-CTU-row and has been agreed to be investigated in EE. Therefore, combination test with the aspect of removing the one-CTU-row constraint from JVET-AG0096 should be conducted.

It was asked up to how many CMVP candidates are added. This depends on the value of trace path. Up to 32 are added.

It was asked how much gain comes from reordering. The proponent has not experimented with this aspect separately. However, given that reordering is already performed in ECM, this is not necessarily a new element introduced by this proposal.

It was asked if temporal motion vector scaling is performed. The proponent said no scaling is performed.

It was asked if multihypothesis prediction (i.e. more than 2 predictions) could result from the CMVP candidate. The proponent said that all CMVP candidates are either uni-prediction or bi-preidction.

It was commented that the performance provided by the proposed method when chained MV prediction trace depth is set to 1 seems to provide more reasonable tradeoff, and this trace depth is also adopted as part of EE2-1.8a. Therefore, the EE test only needs to test this trace depth = 1 option.

It was commented by multiple experts that the gain vs. runtime tradeoff looks interesting.

Investigate in EE. The following aspects should be studied separately:

* Investigate the tradeoff between coding gain and additional MV storage, e.g. provide results when only motion from collocated pictures are used
* Investigate the combination with JVET-AG0096 which removes the one-CTU-row constraint

[JVET-AG0240](https://jvet-experts.org/doc_end_user/current_document.php?id=13813) Crosscheck of JVET-AG0073 (Non-EE2: Chained motion vector prediction) [N. Zhang (Bytedance)] [late]

[JVET-AG0257](https://jvet-experts.org/doc_end_user/current_document.php?id=13830) Crosscheck report of JVET-AG0073: Non-EE2: Chained motion vector prediction [K. Naser (InterDigital)] [late] [miss]

[JVET-AG0074](https://jvet-experts.org/doc_end_user/current_document.php?id=13630) Non-EE2: IntraTMP with HMVP Candidates [C. Zhu, G. Li, T. Tang, L. Luo, H. Guo (UESTC), Y. Huo, Y. Liu (Transsion)]

JVET-AF0079 proposed that IntraTMP BV is stored for HMVP at the last meeting. However, block vectors of History-based Motion Vector Prediction (HMVP) list are only used in IBC mode, but are not used in IntraTMP mode. This contribution proposes extending the IntraTMP candidate list with the HMVP candidate block vectors. The following results are obtained:

AI: Class F : {-0.06%, 0.03%, 0.16%, 99.7%, 99.4%},

AI: Class TGM : {-0.10%, -0.07 %, -0.05%, 97.9%, 99.5%},

RA: Class F : {0.00%, 0.02%, 0.05%, 99.8% , 100.0%},

RA: Class TGM : {-0.13%, -0.35%, -0.30%, 93.8%, 99.7%};

It was commented that the gains for screen content are rather small. It was asked if the proposed method is expected to have gains for natural content. The proponents showed gains of 0.01% for class C and 0.10% for class E in AI. However, such data on natural content are not available in the uploaded package and not cross checked.

It was commented that the contribution seems to belong to the intra category.

Further study.

[JVET-AG0281](https://jvet-experts.org/doc_end_user/current_document.php?id=13854) Crosscheck of JVET-AG0074 (Non-EE2: IntraTMP with HMVP Candidates) [M. Radosavljević (InterDigital)] [late]

[JVET-AG0125](https://jvet-experts.org/doc_end_user/current_document.php?id=13681) AHG12: Parallel friendly use of boundary distortion for DMVR [K. Andersson, R. Yu (Ericsson)]

In JVET-AF0057 a normative subjective enhancement of DMVR was proposed, wherein a subblock boundary distortion metric was introduced in the subblock bilateral matching search. Concerns were raised regarding parallel computation of subblock motions due to introduction of dependency to neighbouring subblock boundary samples.

In the present contribution, an improved method over the JVET-AF0057 method is proposed. The subblock boundary distortion calculations are moved to a refinement step after the subblock bilateral matching search. In the refinement step, the initial motion vector pair and the best bilateral matching candidate motion vector pair are compared with respect to boundary distortion. The motion vector pair that has least sum of boundary distortion, motion vector cost and bilateral matching distortion is selected as the refined motion vector pair.

Objective performance under ECM-11.0 CTC is reported to be:

Random access (Y/U/V): -0.01%/0.02%/-0.01% Dec time: 100.3 %

It is asserted that the proposed method is more parallel friendly compared to the JVET-AF0057 method while the subjective improvement is maintained. It is proposed to include the method to ECM.

The proponent ran the decoder several times and asserted that the decoding time should be reliable. The slight increase in decoder run time was reportedly attributed to DMVR being used slightly more frequently with the proposed change.

It was commented that JVET-AF0057 from the same proponent was a similar contribution brought to the last meeting in October 2023, and a non-normative fix from JVET-AF0057 was included in ECM (turned off in CTC).

It was commented that to include a normative method that claims to be beneficial in terms of subjective quality, then visual testing should be performed. This has not been done for this proposal at this moment.

It was commented that for us to adopt normative method, a cross check report should be provided and that the claimed benefit in terms of parallel processing friendliness should also be cross checked. This has not been done at this moment.

Further study, especially in terms of setting up a subjective viewing session at the April meeting, as well as having the parallelization aspect cross checked by other experts.

[JVET-AG0126](https://jvet-experts.org/doc_end_user/current_document.php?id=13682) AhG12: On ECM temporal partitioning prediction [G. Laroche, P. Onno (Canon)]

This contribution presents a temporal prediction method of the partitioning parameters and syntax elements. In this temporal prediction method, initially presented in the contributions JVET-AD0130, JVET-AE0132, the split modes allowances, the partitioning depths and the split syntax elements ordering are derived for each block according to the partitioning parameters of the current frame and parameters obtained from a temporal area. Compared to the previous contributions, the proposed method has been modified to increase the coding efficiency.

Compared to ECM-11.0, the average BDR gains and runtimes reported in this contribution are as follows:

-0.21% -0.14% -0.12% 98.3% 100.4% for the RA configuration,

-0.02% 0.04% -0.04% 96.0% 100.2% for the Low Delay B configuration.

It was asked if the reported runtime decrease was reliable. The proponent reported that the timing results were obtained from computing cloud with reliable timing information.

It was commented that the proposal requires some additional memory to store partition information of the collocated pictures, but the additional memory is relatively small compared to the motion information that already needs to be stored.

It was commented that the performance vs. complexity tradeoff is quite attractive, esp. given the encoding runtime reduction.

There are three main elements in the proposal: adaptation of maximum MTT depth, temporal prediction of split modes, and temporal prediction of split syntax. The gain of the maximum MTT depth element is estimated to be approximately 2/3 and complexity reduction, with the remaining gain coming from the other two elements.

It was commented that class F shows some loss in RA, and it was requested to provide results for screen content in the EE test.

It was asked how the QP of the collocated picture affects the temporal prediction. The proponent commented that prediction is more effective when the QP of the collocated picture is lower (i.e. better quality).

Investigate in EE. Three elements of the proposal to be studied as follows:

* Temporal prediction of split modes
* Temporal prediction of split modes + Maximum MTT depth adaptation
* Temporal prediction of split modes + Maximum MTT depth adaptation + Prediction of split syntax element order

Class F and class TGM results are requested.

[JVET-AG0149](https://jvet-experts.org/doc_end_user/current_document.php?id=13705) Non-EE2: Improvements to subblock merge mode [J. Chen, R.-L. Liao, Y. Zheng, X. Li, Y. Ye (Alibaba)]

Discussion chaired by M. Wien

In this contribution, it is proposed to improve the subblock merge mode by adding more subblock temporal motion vector prediction (SbTMVP) candidates to the subblock merge list and extending the template matching (TM) based refinement for SbTMVP candidates and affine candidates in subblock merge mode. It is reported that on top of ECM-11.0, the proposed improvements overall achieve {-0.22%, -0.24%, -0.26%} BD-reduction in RA and {-0.19%, 0.00%, 0.04% } BD-reduction in LD.

In the second version, more experimental results are provided.

In the third version, full results are provided.

The results of this contribution were considered to provide interesting coding gains by multiple experts. It was suggested to investigate the multiple included aspects separately in order to identify the benefit of each. The proponent estimates the contributions of the separate aspects to be roughly balanced.

The number of SbTMVP candidates to be added to the list was asked. It was reported that the number is not fixed. At most four candidates would be added dependent on the availability of candidates.

The process of refining subblocks was asked for. It was reported that each candidate would be refined separately. No template would be applied for subblocks inside the blocks, but only for the boundary subblocks.

Investigate in EE with separate tests for the SbTMVP aspect and the affine candidates aspect.

[JVET-AG0282](https://jvet-experts.org/doc_end_user/current_document.php?id=13855) Crosscheck of JVET-AG0149 (Non-EE2: Improvements to subblock merge mode) [L. Zhao (Bytedance)] [late]

[JVET-AG0194](https://jvet-experts.org/doc_end_user/current_document.php?id=13750) Non-EE2: Reference filtering for inter-prediction [A. Filippov, V. Rufitskiy, K. Suverov (Ofinno)]

The contribution proposes an additional method of filtering (interpolated) reference samples for inter predicted blocks. The proposed technique involves filtering with spatial filters that are derived from templates of reconstructed and reference blocks. On top of ECM-11.0, the following BD-rate gain results are reportedly obtained for the proposed method {Y, U, V}: {-0.11%, -0.11%, -0.08%} and {-0.12%, 0.02%, 0.05%} for RA and LDB configurations, respectively.

Encoder and decoder runtimes are not provided in the contribution. The proponent estimated enc/dec runtime increase to be 3%/1% respectively. The proponent asserted that the proposed filter is only 7-tap and therefore derivation at the decoder is not overly complex.

Two classes of filters are used in the current implementation of the proposal.

The proposed method is applied to uni-prediction only in the current implementation. The proponent found the design for bi-prediction to be more complex and therefore is not proposing it.

Presentation slides are not available at the time of presentation. Proponent will upload.

Similar method was proposed before, and the gains shown before were larger. The proponent asserted that this could be due to other inter prediction adoption in the meantime.

Investigate in EE.

[JVET-AG0195](https://jvet-experts.org/doc_end_user/current_document.php?id=13751) Non-EE2: LIC model parameter inheritance for merge modes [C.-C. Chen, H. Huang, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes to enable LIC (Local Illumination Compensation) model inheritance for merge modes to infer LIC model parameters from merged candidate motion with LIC enabled. When the new merge candidate is selected, the derivation process of LIC model parameters could be bypassed completely. In the proposal, two methods are tested for mode selection. First determines LIC inheritance at LIC flag swapping stage, while the other leaves the new merge candidates in ARMC for reordering. The experimental results on top of ECM-11.0 are reported, as follows:

[Test 1] Determine LIC model inheritance at LIC flag swapping stage

\* RA(Y/U/V): {-0.05%, 0.04%, -0.07%, EncT 99.1%, DecT 99.7%};

\* LDB(Y/U/V): {-0.13%, -0.15%, 0.01%, EncT 99.2%, DecT 99.6%};

[Test 2] Determine LIC model inheritance through ARMC

\* RA(Y/U/V): {-0.07%, 0.04%, 0.01%, EncT 100.2%, DecT 100.3%};

\* LDB(Y/U/V): {-0.11%, 0.20%, -0.18%, EncT 100.4%, DecT 99.9%};

It was asked if the test 1 in this proposal interacts with EE2-2.1 (of which the combination test 2.6f is to be revisited at the time of this presentation). The proponent stated that the even if the LIC flag signaling for merge mode, as proposed in EE2-2.1, would be adopted, test 1 which proposes to inherit the LIC model parameters would still be applicable.

It was commented by the proponent that the encoding time reduction in test 1 was due to increased use of LIC, which means BDMVR would not be used, which contributed to encoding time reduction.

One expert commented that this proposal allows the LIC model parameters to be inherited rather than being derived, which should be beneficial.

The proposed method is also applied to GPM mode, i.e. allowing the GPM candidate to inherit LIC model parameters. It was asked how this GPM part performed. The proponent stated this had not been tested separately.

It was commented by multiple experts that the proposed methods bring desirable performance vs. complexity tradeoff, and support investigation in EE.

The two tests in this proposal are mutually exclusive. Currently, test 2 seems to bring better tradeoff (more gain but also no runtime increase). However, given the LIC related adoptions at this meeting, it was claimed by the proponent that it could be interesting to perform both tests on top of the next version of ECM to better understand tool interaction.

Investigate in EE. Both methods (test 1 and test 2) will be separately tested. Also test the performance from applying this to the GPM mode separately.

[JVET-AG0312](https://jvet-experts.org/doc_end_user/current_document.php?id=13885) Crosscheck of JVET-AG0195 (Non-EE2: LIC model parameter inheritance for merge modes) [Y. Kidani, K. Kawamura (KDDI)] [late]

#### Cross Component Prediction (1)

[JVET-AG0200](https://jvet-experts.org/doc_end_user/current_document.php?id=13756) Non-EE2: Inter CCP merge mode with zero luma CBF [Z. Deng, K. Zhang, L. Zhang (Bytedance)]

In ECM-11.0, inter CCP merge mode is allowed only if luma CBF is non-zero. In this contribution, it is proposed to allow inter CCP merge mode when the luma CBF is equal to 0. The proposed method is applied on SCC sequences. The impact on coding efficiency and runtimes of the proposed method over ECM-11.0 are reportedly for {Y, U, V, EncT, DecT}:

RA :

class F { 0.02%, -0.28%, -0.49%, 100.9%, 101.1% }

class TGM { -1.26%, -3.61%, -3.61%, 101.1%, 100.7% }

LDB :

class F { -0.17%, -1.43%, -1.69%, 101.5%, 100.8% }

class TGM { -0.90%, -5.74%, -5.84%, 101.2%, 100.9% }

In current ECM, inter CCP merge mode is applied to both natural content and screen content. However, this proposal only applies the proposed change to screen content.

The proposed method requires an additional CU-level flag to be signaled when merge skip mode is applied, and when luma CBF or root CBF is equal to 0, which does not give good performance vs. complexity tradeoff for natural content in preliminary implementation. It was commented that having different signaling conditions for inter CCP merge mode for natural content vs. screen content is not desirable. The proponent is asked to bring results for natural content as well, using the same normative condition for natural and screen content.

Investigate in EE. Natural content results requested along with screen content results.

#### In-Loop Filters (2)

[JVET-AG0065](https://jvet-experts.org/doc_end_user/current_document.php?id=13621) Non-EE2: Adaptive coefficient precision for CCALF [N. Song, Y. Yu, H. Yu, D. Wang (OPPO)]

This contribution proposes a method of using adaptive coefficient precision for CCALF to replace the fixed 7-bit precision currently used in ECM. The proposed method is implemented on top of ECM-11.0 software, and its performance is reported as follows:

AI: { 0.00% Y, -0.20% U, -0.25% V, 100.2% EncT, 100.5% DecT }

RA: { 0.00% Y, -0.26% U, -0.39% V, 99.9% EncT, 100.0% DecT }

EE2-4.3 which proposes to use adaptive coefficient precision for luma ALF was adopted in this meeting. The proponent of EE2-4.3 reported that adaptive coefficient precision for chroma ALF was attempted but brought no further gain.

This contribution is closely related to JVET-AG0233, which also proposes to apply adaptive coefficient precision to CCALF, and was already agreed to be investigated in the next round of EE.

The set of precisions in this contribution {7, 8, 9, 10} is the same as proposed in JVET-AG0233. Both propopents agreed to test the two contributions (JVET-AG0233 and JVET-AG0065) as one EE test set.

It was commented that the set of precisions for luma ALF is {5,6,7,8}, and it would be interesting to test this set on CCALF as well. This was agreed.

Investigate in EE together with JVET-AG0233.

Tests to be conducted should include:

* Study the effect of coefficient constraint removal and adaptive precision separately for CCALF
* Study using the set of precisions {5,6,7,8}, which is used for luma ALF, on CCALF adaptive precicion.

[JVET-AG0264](https://jvet-experts.org/doc_end_user/current_document.php?id=13837) Crosscheck of JVET-AG0065 (Non-EE2: Adaptive coefficient precision for CCALF) [W. Yin (Bytedance)] [late]

[JVET-AG0289](https://jvet-experts.org/doc_end_user/current_document.php?id=13862) Crosscheck of JVET-AG0065 (Non-EE2: Adaptive Coefficient Precision for CCALF) [N. Hu (Qualcomm)] [late]

[JVET-AG0198](https://jvet-experts.org/doc_end_user/current_document.php?id=13754) Non-EE2: Coding Information based ALF Classification [W. Yin, K. Zhang, Y. Wang, Z. Deng, L. Zhao, N. Zhang, M. Salehifar, L. Zhang (Bytedance)]

In this contribution, a coding information-based classification method is proposed for Luma Adaptive Loop Filter (ALF). In the proposed method, the prediction mode and the partitioning information are introduced into the Luma-ALF classification process to build up an additional rule.

On top of ECM-11.0, simulation results of the proposed method are reported as below:

AI: 0.01%, -0.01%, -0.01%, 100.1%%, 100.3%.

RA: -0.03%%, -0.01%%, -0.01%, 100.2%, 100.4%.

LB: -0.10%, -0.07%, 0.12%, 100.5%, 100.1%.

The number of ALF classes as proposed in this contribution is reduced to 24 (from 25 in current ECM).

It was commented by multiple experts that the description in the document lacks some detail. Further detail to be added, and a revised document to be uploaded.

It was asked if separate tests had been performed on each type of the coding information (prediction mode and partition info) used in the proposal. The proponent reported that prediction mode information brings a bit more gain.

It was commented that the gains were not significant, and most of the gain in RA came from class C.

It was commented by ALF expert that the LB gain is interesting, and changes seem to be small, and support was expressed to study in the next round of EE.

In a later discussion, the proponent reported that an updated document had been provided in v2 package. This was checked by non-proponent and commented to be sufficient.

Investigate in EE. Separate tests on adding each type of the coding information (prediction mode and partition info) are requested.

[JVET-AG0321](https://jvet-experts.org/doc_end_user/current_document.php?id=13894) Crosscheck of JVET-AG0198 (Non-EE2: Coding Information based ALF Classification) [N. Hu (Qualcomm)] [late] [miss]

#### Entropy coding, transforms, and transform coefficient coding (7)

[JVET-AG0062](https://jvet-experts.org/doc_end_user/current_document.php?id=13618) Non-EE2: Multiple Transform Sets Selection for LFNST/NSPT [F. Wang, Y. Yu, H. Yu, D. Wang (OPPO)]

This contribution proposes a multiple transform set selection method for intra blocks coded with LFNST/NSPT. In the proposed method, CUs coded with DIMD, TIMD, MIP, SGPM and ITMP modes can choose an alternative LFNST/NSPT transform set. The transform set selection is signaled in the bit-stream. On top of ECM-11.0, the performance of the proposed method is reported as follows:

AI {-0.08% Y, 0.06% U, 0.05% V, 107.9% EncT, 100.7% DecT}

Complete RA results were not available at the time of presentation. Class B results were reported to be {-0.06% Y, 0.10% U, 0.05% V, 106.3% EncT, 100.0% DecT}.

The proponent asserted that the encoding runtime increase was mainly due to unoptimized encoder implementation.

The proposal introduces an additional CU-level flag for CUs coded with DIMD, TIMD, MIP, SGPM and ITMP modes. It was asked if there are other conditions that determine whether this flag is signaled that might introduce parsing dependency, the proponent reported that no parsing dependency would be introduced.

When the two intra prediction modes used to select the transform set index are very similar (i.e. less than or equal to 4), what will happen? The proponent asserted that the usage of DIMD derived mode could guarantee that the second mode to be sufficiently different from the first.

Transform kernels are not modified by this proposal.

It was commented that selecting from different transforms usually requires full RDO, and therefore the encoder runtime increase could be difficult to bring down.

It was commented that the idea in this contribution was interesting, however the current tradeoff did not seem to be promising.

However, multiple experts expressed interest in studying this in the next round of EE.

Investigate in EE. Reduction in encoding time without compromising the gain would be needed for this method to be considered valuable.

[JVET-AG0307](https://jvet-experts.org/doc_end_user/current_document.php?id=13880) Crosscheck of JVET-AG0062 (Non-EE2: Multiple Transform Set Selection for LFNST/NSPT) [X. Li (Alibaba)] [late]

[JVET-AG0064](https://jvet-experts.org/doc_end_user/current_document.php?id=13620) AHG12: On Context modeling in Chroma Coefficient Coding [L. Xu, H. Yu, Y. Yu, D. Wang (OPPO)]

This contribution proposes to use the same and consistent spatial neighbors to derive context model index and Rice parameter for the chroma blocks not coded by LFNST or NSPT. The proposed method is implemented on the ECM-11.0 software.

The simulation results of the proposed method under CTC are shown as follows,

RA: -0.00% (Y), -0.02% (U), -0.04% (V)

The simulation results of the proposed method when DualITree=0 are shown as follows,

AI: 0.01%(Y), -0.22%(U), -0.18%(V)

The proposed modification only affects single tree, and is not applicable to dual tree.

Cross checker (also original proponent of JVET-AE0102) confirmed the results.

It was commented that the performance difference due to the proposed modification is negilible in the context of the current stage of exploration.

It was commented that making the proposed modification could also make certain aspects “more complex”.

No action at this stage.

[JVET-AG0284](https://jvet-experts.org/doc_end_user/current_document.php?id=13857) Crosscheck of JVET-AG0064 (AHG12: On Context modeling in Chroma Coefficient Coding) [P. Nikitin (Qualcomm)] [late]

[JVET-AG0102](https://jvet-experts.org/doc_end_user/current_document.php?id=13658) AHG12: New context model parameters for low delay B condition [R.-L. Liao, Y. Ye, J. Chen, X. Li (Alibaba)]

This contribution proposes to separate context model parameters of low delay B slices from those of non-low delay B slices. In addition to three sets of context model parameters for B-/P-/I-slice types, a fourth set of context model parameters is proposed and is trained using the bitstream generated by ECM-11.0 under low delay B condition. A SPS flag is signaled to indicate whether the original set or the newly proposed set of context model parameters are used to initialize the context model of B slices. It is reported that on top of ECM-11.0, the overall coding performance for {Y, U, V} is { -0.04%, 0.10%, 0.08%} for AI, { -0.15%, 0.38%, 0.16%} for RA, { -0.18%, 0.63%, 0.18%} for LB and { -0.11%, 0.33%, 0.20%} for LP. Compared with EE2-5.1 which retrains CABAC initialization parameters for all slices, the overall coding performance for {Y, U, V} is {-0.01%, 0.05%, 0.07%} for AI, {-0.01%, 0.10%, 0.01%} for RA, {-0.12%, -0.13%, -0.37%} for LB and {-0.14%, -0.22%, -0.54%} for LP.stated that the software for training from EE 5.1 has been used. It was commented that the training for LDB is currently somewhat unreliable with AI and RA typically showing gains while LDB may be not doing so. It was commented by experts that separating the context for the LDB setting may be beneficial. The proposal adds a context table and only uses it for LDB with the use being signaled in the SPS. It was commented that having a configuration dependent setting might not be favorable. It was asked if setting the flag based on the properties of the B prediction (i.e., both MVs coming from the same direction and not from both sides). It was commented by the proponent that the training set included the CTC and other test sequences from the JVET ftp site.

It would be interesting to understand what aspect of the proposal induces the observed gains. Further work on this is encouraged.

Investigate in EE. The test shall include switching of the context initialization based on the slice level LDB flag and the sequence level solution.

[JVET-AG0311](https://jvet-experts.org/doc_end_user/current_document.php?id=13884) Crosscheck of JVET-AG0102 (AHG12: New context model parameters for low delay B condition) [Y. Kidani, K. Kawamura (KDDI)] [late]

[JVET-AG0108](https://jvet-experts.org/doc_end_user/current_document.php?id=13664) AhG12 Entropy coding extension [F. Galpin, F. Lo Bianco, C. Salmon-Legagneur, M. Balcilar (InterDigital)]

This contribution proposes to add more flexibility to the bins coding in current ECM. First, some bins coded in EP are transformed to be coded with an associated context. Second, an additional coding type is introduced in between the CABAC and the EP coding where the encoding uses a fixed probability without adaptation. The method is tested on some bins which were coded in EP. It is reported that on top of ECM-11.0, the overall coding performance impact for {Y, U, V} is {0.00% -0.11% -0.15%} {-0.08% -0.22% -0.23%} in AI and RA configurations respectively.

The systematic context numbering aspect would be considered normative in a formal standardization phase, but only requires shifting around of some contexts in ECM software at this exploration stage. This aspect does not contribute to coding performance gain.

The proposal also includes an aspect that shifts context coding to non-context coding but settles upon non-equal probability rather than equal probability, which was considered interesting by multiple experts.

It was commented that the “overfitting to current CTC sequences” problem might be a more pronounced problem for this proposal, as it would settle on fixed probability and not be able to adapt from that point on.

Investigate in EE. Proponents are requested to bring results that are trained on different training sets, and especially training data should include some non-CTC content (e.g. BVI content).

[JVET-AG0110](https://jvet-experts.org/doc_end_user/current_document.php?id=13666) AHG12: 16 States TCQ with State Exchange [M. Balcilar, K. Naser, Y. Chen, F. Galpin, F. Le Léannec (InterDigital)]

This contribution proposes to increase the number of trellis code quantization (TCQ) states in dependent quantization (DQ). The number of states is increased from 8 to 16, where the 8 new states use similar transition rule as the original ones.

The following results {Y, U, V; EncT, DecT, VmPeak} are obtained on top of ECM-11.0:

* RA: { -0.08%, -0.07%, -0.04% ; 100.9%, 99.3%, 101.2%}.
* LDB:{ -0.15%, 0.61%, -0.22% ; 101.0%, 100.3%, 101.6%}.

It was commented by the proponent that the proposed change does not provide good tradeoff for AI, as it introduces around 7-8% encoding run time increase, and 0.1% gain for AI. In the proposal the proposed 16-state TCQ is only enabled for RA and LDB configs. The proponent suggests that this could be due to naïve encoder implementation, and a more sophisticated encoder could potentially achieve better tradeoff for AI.

It was asked by multiple experts if the proposed new states and transitions can be achieved by encoder-only modifications, i.e. if the current 8-state TCQ is capable of achieving these transitions. The proponent claimed that the proposed normative changes are more straightforward.

It was asked for the proponent to show an example of a transition that is not achievable by the current 8-state TCQ. It was commented that the description in the proposal is not sufficient to understand.

It was commented that offline discussion was initiated by non-proponent but not answered by proponent. As a general principle, offline discussion is encouraged as it often facilitates a more efficient meeting discussion.

Multiple experts commented that the performance vs. runtime tradeoff appears interesting.

It was requested for the proponents to provide a description in the EE document (JVET-AG2024) as detailed as possible within the timeline.

Investigate in EE. Besides providing software and results, proponents are also requested to provide at least one clear example of a transition that is not achievable by the current 8-state TCQ, and get this cross checked.

[JVET-AG0185](https://jvet-experts.org/doc_end_user/current_document.php?id=13741) Non-EE2: Slice based Rice parameter selection for transform skip residual coding [H.-J. Jhu, X. Xiu, W. Chen, C.-W. Kuo, N. Yan, C. Ma, X. Wang (Kwai)]

In this proposal, the same slice-based Rice parameter syntax in the VVC is used to improve the coding efficiency of screen content sequences (SCC) in the ECM. Compared to ECM-11.0 anchors, simulation results are summarized as follows:

For class F sequences:

AI { -0.07%, -0.22%, -0.09%, 100.2%, 99.0% }, RA { -0.08%, 0.03%, 0.03%, 99.9%, 99.1% }, LD { -0.16%, -0.08%, -0.52%, 99.3%, 98.5%\* }

For TGM sequences:

AI { -0.22%, -0.20%, -0.25%, 100.2%, 99.2% }, RA { -0.08%, -0.09%, -0.08%, 100.3%, 100.0% } LD { -0.10%, -0.07%, -0.14%, 100.2%, 97.2%\* }

\*missing rate point is copied from anchor

The slice-based Rice parameter syntax used in this proposal only exists in version 2 of VVC spec, and not in version 1, and only applied to transform skip residual coding. This Rice parameter syntax was shown to benefit natural content and screen content, both of higher bit depth (12-bit and beyond) at the time. This proposal uses a similar method (with different parameter than VVC v2) on screen content of normal bitdepth and shows some small gain. It was questioned why the same parameter as in VVC v2 could not be used.

The proponent reported that the same encoder strategy (to select the value of the Rice parameter) as in VVC v2 is used.

It was asked if this was tried on natural content, the proponent reported internal test showed no coding benefit.

The reported gain is rather small esp. considering it only works on screen content.

At this stage, it is more important to explore coding tools or algorithms that provide more substantial performance.

No action at this time.

[JVET-AG0277](https://jvet-experts.org/doc_end_user/current_document.php?id=13850) Crosscheck of JVET-AG0185 (Non-EE2: Slice based Rice parameter selection for transform skip residual coding) [L. Xu (OPPO)] [late]

[JVET-AG0287](https://jvet-experts.org/doc_end_user/current_document.php?id=13860) Non-EE2: Rice parameter derivation for transform skip residual coding [M. Coban, M. Karczewicz (Qualcomm)] [late]

Adaptive template-based Rice parameter derivation for transform skip residual coding in VTM-6.0 was later replaced by usage of a constant Rice parameter value in VVC. This contribution proposes to use the Rice parameter derivation method in VTM-6.0 for transform skip residual coding. The proposed method provided following results on top of ECM-11.0:

For class F sequences:

AI { -0.05%, -0.16%, -0.01%}, RA { -0.06%, -0.13%, 0.28%}

For TGM sequences:

AI { -0.23%, -0.27%, -0.31%}, RA { -0.21%, -0.26%, -0.18%}

This proposal suggests to bring back a Rice parameter derivation method that was available in an early-stage VTM (VTM-6.0) but replaced by a different method in VVC standard.

The reported gain is rather small esp. considering it only works on screen content.

No action at this time.

#### RPR (1)

[JVET-AG0116](https://jvet-experts.org/doc_end_user/current_document.php?id=13672) AHG12: GOP-based RPR encoder control for ECM [K. Andersson, J. Ström, P. Wennersten, R. Yu, W. Ahmad (Ericsson)]

GOP-based RPR encoder control has been available for VTM since VTM-19.0. It selects coding in a reduced resolution when the PSNR impact of re-scaling from source resolution to reduced resolution is expected to be limited. ECM does not have such functionality. In JVET-AF0058 the GOP-based RPR encoder control was implemented in ECM and was demonstrated to show benefit for high QPs, but that contribution lacked results for CTC. The present contribution includes results on CTC.

Objective performance (BDR) is reported as:

ECM-11.0 with GOP-based RPR versus ECM-11.0:

* RA Y: -0,06%, U: 0,75%, V: 0,71%
* LDB Y: 0,01%, U: 0,09%, V: 0,03%

It is also asserted that subjective quality can be improved by the proposal.

It is proposed to add the GOP-based RPR encoder control from VTM to ECM to enable further experimentation of RPR functionality in ECM for improved coding efficiency.

In v2 results based on MS-SSIM were added.

* RA -0,29% 0,22% 0,05%
* LDB: -0,05% -0,03% -0,06%

The comparison performed in this contribution uses the normal CTC (rather than RPR CTC) as the anchor. The goal is to show that for 4K sequences, by switching the resolution on a GOP basis, a small luma performance gain can be achieved (though chroma performance loss is observed).

The proposed method is encoder only.

The proponent reported subjective improvement in offline viewing. No examples were provided in the contribution.

The cross checker suggests that the proposed change aligns the RPR encoder of ECM (no GOP-based RPR) and VTM (GOP-based RPR).

The proponent reported that this is the only difference between RPR encoder of ECM and VTM.

It was commented that such modifications are not the core activities for the ECM exploration.

Decision (SW): adopt JVET-AG0116.

[JVET-AG0319](https://jvet-experts.org/doc_end_user/current_document.php?id=13892) Crosscheck of JVET-AG0116 (AHG12: GOP-based RPR encoder control for ECM) [J. Lainema (Nokia)] [late]

### CTC for EE2/ECM and general ECM improvements (1)

Contributions in this area were discussed at 2345–XXXX on Thursday 25 Jan. 2024 (chaired by JRO).

[JVET-AG0334](https://jvet-experts.org/doc_end_user/current_document.php?id=13907) AHG12: Fix for coding at high QPs [K. Andersson, C. Hollmann (Ericsson)] [late]

It was discovered that ECM-11.0 cannot encode in low-delay configuration for QP equal to 55 or higher. The reason is that for high QP the lambda as part of ARMC can exceed 10 bits but the lambda is encoded with at most 10 bits.

One solution is to set the lambda for the highest QP to 1023 instead of 1125.369910 which then would fit the current encoding of the lambda with at most 10 bits.

One alternative solution is to increase the number of bits for encoding lambda to 11 bits.

It is proposed to include one of the proposed fixes to enable encoding at high QP with ECM.

It was confirmed by the original proponents of the lambda syntax element that there is a problem. They suggest that the second solution (increasing from 10 to 11 bits) would be preferable.

In CTC, this does not have any impact, and only low delay B at high QP would be affected.

The problem was detected when investigations were made in the context of AHG8 with high QP encoding.

It was commented that the second solution might increase the bit rate, whereas the first solution might just be an encoder bug fix, clipping the value to the highest allowed.

Decision(SW/BF): Adopt JVET-AG0334 first solution.

.

# High-level syntax (HLS) and related proposals (46)

## AHG9: SEI messages on NNPF aspects other than grouping (2)

Contributions in this area were discussed at 0730–0920 on Thursday 18 Jan. 2024 (chaired by JRO).

[JVET-AG0089](https://jvet-experts.org/doc_end_user/current_document.php?id=13645) AHG9: Temporal extrapolation purpose for the neural-network post-filter characteristics SEI message [M. M. Hannuksela, F. Cricri, H. Zhang (Nokia)]

The neural-network post-filter characteristics (NNPFC) SEI message includes the nnpfc\_purpose syntax element, which enables to signal any number of purposes to which the defined neural-network post-filter (NNPF) is intended. For example, the purposes include picture rate upsampling through interpolating pictures between one or more pairs of input pictures.

It is asserted that there are use cases for temporal extrapolation, i.e., generating one or more pictures following all the input pictures in output order. Hence, this contribution proposes the addition of temporal extrapolation as one of the purposes of nnpfc\_purpose. This addition includes the following aspects:

1. A new bit in nnpfc\_purpose is dedicated for temporal extrapolation.
2. When nnpfc\_purpose indicates temporal extrapolation, nnpfc\_extrapolated\_pics\_minus1 plus 1 indicates the number of extrapolated pictures generated by the NNPF.
3. The constraints to avoid conflicting or ambiguous output order are updated to concern extrapolated pictures in addition to interpolated pictures.

The main difference relative to frame rate upsampling would be that the extrapolated picture would be after the input picture (not allowed by the current syntax).

It was asked if a network designed for interpolation could also be used for extrapolation? Potentially, but it is more likely that another dedicated network would be designed.

It was pointed out that in picture extrapolation applications (navigation, cloud gaming) it is also common that multiple hypotheses are generated for the same temporal position. This would not be supported by the proposed syntax, but could be added and would make another differentiation from the current picture rate upsampling.

From the comments made during the review, the idea of defining a separate purpose was supported and assessed to be important for various application.

Decision: Include JVET-AG0089 in TuC. The option of a network generating several instances at same temporal position shall be further studied, as there may be some general conflict with constraints in the existing NNPF architecture (disallowing several outputs at same time instance).

[JVET-AG0192](https://jvet-experts.org/doc_end_user/current_document.php?id=13748) AHG9: Miscellaneous NNPF items related to VUI [J. Xu, Y.-K. Wang (Bytedance)]

This contribution proposes the following changes to the VSEI text:

Item 1: Fix a typo when vui\_chroma\_sample\_loc\_type\_frame is mentioned in the VSEI text.

Item 2: Fix two bugs for the constraints on nnpfc\_matrix\_coeffs and add two NOTEs in the semantics of nnpfc\_matrix\_coeffs to clarify complicated cases.

Item 3: Fix an asserted bug on the inference of nnpfc\_full\_range\_flag.

Item 4: Add a missing inference and a NOTE on nnpfc\_chroma\_sample\_loc\_type\_frame.

Item 5: Enable sending of nnpfc\_matrix\_coeffs and nnpfc\_full\_range\_flag when nnpfc\_out\_format\_idc is not equal to 1.

Item 6: Add a NOTE about the aspect ratio of NNPF output pictures when ResolutionResamplingFlag is equal to 1.

Item 7: Enable sending the overscan information for NNPF output pictures.

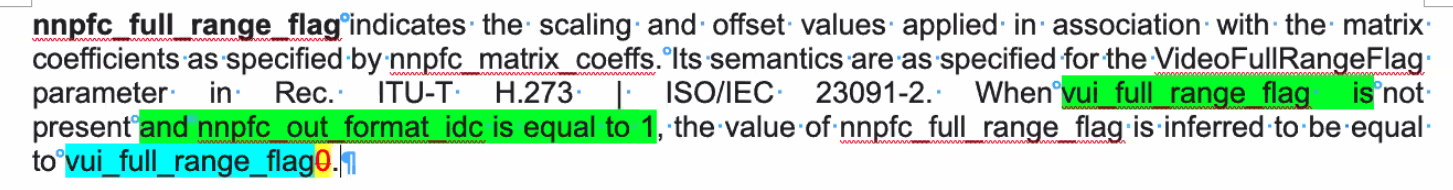
Item 1 – agreed

Item 2 – it was suggested that the first suggested change (formulating a third condition about presence of the syntax elements) should better be formulated similar to the second change (formulating an “and” in the first condition)

Generally agreed that this additional constraint is necessary, but better be formulated as described in the last paragraph.

It was further asked in which cases these cases might occur? It was commented that other options might exist to resolve the problem. It was pointed out that additional problems might exist if only one of the two bitdepth elements is present. This should be further studied.

Item 3 – generally agreed that there is an inconsistency, but it was pointed out that some additional condition might need to be added under which further circumstances the inference would be made at all (e.g., if output would be floating point or not). M. Hannuksela was asked to report on this latter aspect. The following text was suggested as a result of the offline communication and agreed (to be included in update of JVET-AG0192):



Item 4 – it was commented that inferring the chroma location being unknown is something that a decoding device naturally would do if the location type is not present. This additional semantics of inference may not be necessary.

Item 5 – it was asked if releasing the constraint that matrix coefficients are only existing if the output is integer would be backward compatible? Proponents argue that this could be seen as an oversight. Another expert mentions that the constraint was imposed because CICP is defining matrix coefficients in integer. Further, a comment is made that also a semantics change would be necessary currently not included in the proposal. It was pointed out that colour primaries and matrix coefficients are not necessarily related, such that they may not be inferred for the floating point output. Further study on this suggested change appears necessary, also in context of studying the relevance of colour conversion applied to floating point.

Item 6 – It was commented that this additional note is not needed, as a decoding device would naturally assume that the aspct ratio at output is the same as at input if not explicitly signalled differently.

Item 7 – This appears rather as a new proposal than solving a problem. It requires additional syntax to indicate that a spatial extension of a picture is made by synthesis rather than upsampling. It was suggested that another option to achieve this would be to define a new purpose (similar as for temporal extrapolation), which also might give the decoding device more flexible to use it or not. Interesting proposal, further study reommended.

Agreed to include items 1-3 (with slight modifications as indicated above) in JVET-AG1004.

## AHG9: SEI processing order and processing order nesting SEI message aspects (9)

Contributions in this area were discussed at 2100–2310 on Thursday 18 Jan. 2024 (chaired by JRO), and at 2330-0130+1 on Thursday 18 Jan. 2024 (BoG chaired by Jill Boyce).

[JVET-AG0052](https://jvet-experts.org/doc_end_user/current_document.php?id=13608) AHG9: On using SEI processing order for NNPF grouping [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

This contribution proposed some changes to sei processing order SEI message as the result of studying the SEI message for enabling it to also support grouping of NNPF SEI messages. The changes are asserted to enable the sei message to also support parallel grouping and to handle the case where SEI messages with prefix information and no prefix information present are included in sei processing order SEI message.

In the revision 1 version of this contribution, snippet text from JVET-AF0061 that describes how cascading and alternative grouping can be supported using sei processing order SEI message. In addition, example of cases how this proposal adds the capability to support parallel grouping is also added.

With the proposed flag from this contribution, below is the example how the three different grouping and their combination can be supported:

Case 1: Alternative grouping (filter A or filter B): none of filter A and filter B are included in an SPO SEI message. Decoder would choose to invoke filter A or filter B only.

Case 2: Cascading grouping (filter A then filter B): filter A and filter B are included in an SPO SEI message with processing order value of A is lower than processing order value of B.

Case 3: Parallel grouping (filter A and filter B in parallel): filter A and filter B are included in an SPO SEI message with processing order value of A and B equal and po\_parallel\_processing\_enabled\_flag equals to 1.

Case 4: Alternative and cascading grouping (filter A then B or filter C then D): filter A and filter B are included in an SPO SEI message with processing order value of A is lower than processing order value of B; filter C and filter C are included in another SPO SEI message with processing order value of C is lower than processing order value of D.

Case 5: Alternative and parallel grouping (filter A and filter B or filter C and filter D, in parallel): filter A and filter B are included in an SPO SEI message with processing order value of A and B equal and po\_parallel\_processing\_enabled\_flag equals to 1; filter C and filter D are included in another SPO SEI message with processing order value of C and D equal and po\_parallel\_processing\_enabled\_flag equals to 1.

Case 6: Cascading and parallel grouping (filter A then filter B and C in parallel): filter A, filter B, and filter C are included in an SPO SEI message with processing order value of A is lower than processing order values of B and C, processing order value of B is equal to that of C, and po\_parallel\_processing\_enabled\_flag equals to 1.

First item of contribution: parallel topology:

As an example for the need of parallel configurations, an example is given where different outputs are to be generated for different purpose, e.g. human and machine consumption.

It was asked if it would be mandatory to generate both outputs? If it is left to the application to generate only one of them, it would be the same as alternatives.

It was pointed out that a parallel configuration could be achieved by using two different SPO SEI. This was the view of the last meeting. It would anyway need to be up to the application to understand the purpose, and run both of them if needed and sufficient resources are available.

For the application to understand the purpose, it might be beneficial to extend the list of purposes specified.

Several experts expressed concerns that including the parallel option in the SEI message might introduce some complication.

It was asked if the two outputs of the parallel outputs would be combined? Not according to the proponent’s opinion. Another expert suggests that according to his opinion cases could exist where both outputs are required. This might in particular be the case when the parallel processing is not done at the last stage, but a subsequent stage requires both.

One expert commented that in the last meeting, the following statement was produced:

po\_sei\_processing\_order[ m ] equal to po\_sei\_processing\_order[ n ] indicates that there is no preferred order of processing between the types of SEI messages associated with indexes m and n (e.g., they can indicate different properties that are both applicable at that stage, or alternative processes that can be applied, or one can indicate a property and the other can indicate a process).

It was commented that “e.g.” means it is an example, but not the only possibility.

If the parallel usage is at the last stage of the processing pipeline, it can be solved by using two different SPOs, and let the decoder decide whether to select one of them or run both (depending on needs).

If a parallel configuration is needed in the middle of the processing pipeline (i.e. a subsequent stage requires both as input), defining this in the SPO SEI would be necessary. However, reasonable use cases from current SEIs are not obvious at this moment, such that an appropriate definition might not be possible at this time, and further study is needed.

Second item: Coexistence of SEI message with and without prefix in SPO SEI:

***Option 1:*** Define a constraint to disallow situation of mix of SEI messages with the same payload type to be included in sei processing order SEI message in which one of the SEI message has no prefix information while the rest have prefix information.

***Option 2:*** Specify that when two or more SEI messages with the same payload types are included in one or more sei processing order SEI messages and one of the SEI messages has no prefix information while the rest have, processing order of those SEI messages are treated independently (i.e., the ones with no prefix information represent all the SEI messages of that type but excluding those included in the sei processing order SEI message with prefix information).

***Option 3:*** Specify that when two or more SEI messages with the same payload types are included in one or more sei processing order SEI messages and one of the SEI messages has no prefix information while the rest have, those SEI messages with prefix information are considered part of the one without prefix information and both of the following apply:

* The sei processing order SEI message that contains the SEI messages with prefix information shall not contain SEI message with other payload type.
* The SEI message with no prefix information and the SEI messages with prefix information shall not be included in the same sei processing order SEI message.

It was commented that the intent of prefix is differentiation between SEI messages of same payload type. From that point of view, something like option 1 appears appropriate. One expert commented that he believes that such a constraint already exists. It might however have been removed. Check offline.

It was further commented that in case of wrapping prefix might not be needed. Therefore, option 1 might need to be conditioned on whether wrapping is used or not.

Agreed that a constraint is needed. Exact text was developed offline (included in v4 of JVET-AG0052, yellow highlighted), which was presented in JVET on Thursday 25 Jan. 2150.

Agreed to include item 2 option 1 (from JVET-AG0052v4) into JVET-AG2027.

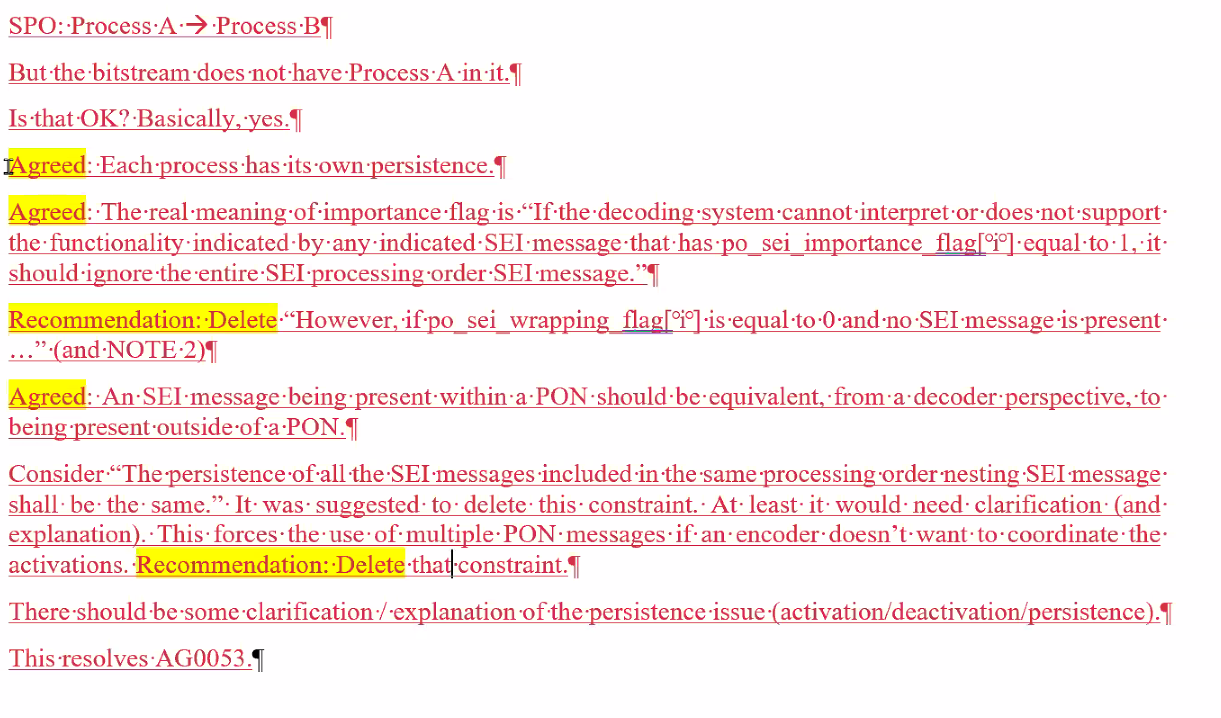
[JVET-AG0053](https://jvet-experts.org/doc_end_user/current_document.php?id=13609) AHG9: On activation and cancelling persistence of SEI message included in a processing order nesting SEI message [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

This contribution asserted that there are problems with the activation and persistency cancellation of processing-order-nested SEI messages. To address the asserted problems, the following changes are proposed:

1. Regarding activation of processing-order-nested SEI messages, do one of the following options:
   1. Option 1: Specify that when PON SEI message is present, the SEI message(s) included in the PON SEI message is active, unless they need to be activated by corresponding activation SEI message(s).
   2. Option 2: Specify a constraint that disallows SEI messages that need to be activated by other SEI messages and SEI messages that do not need to be activated by other SEI messages be contained in the same processing order nesting SEI message.
2. Add an id (i.e., pon\_id) in PON SEI message. This id is used to cancel the persistence of all SEI messages included in the PON SEI message.
3. Add cancel flag (i.e., pon\_cancel\_flag) in PON SEI message. When pon\_cancel\_flag is equal to 1, it cancels the persistence of all SEI message included in the previous PON SEI message with the same pon\_id.
4. Add the following constraints:
   * The value of persistent flag of all SEI messages included in the same PON SEI message shall be the same.
   * The value of cancel flag of all SEI messages included in a PON SEI message shall be equal to 0.

It was commented that the current concept in PON and SPO is that the individual SEI messages control their activation. The constraint was however introduced that all SEI messages in PON shall have same persistence, and if cancelled this is done altogether. This was introduced to simplify implementation, but would in principle not be necessary.

Discuss offline with proponents of PON to get an understanding of the current concept, identify if a problem exists or additional mechanisms may be useful for simplifying implementation. As a result of a side activity led by G. Sullivan, the following solution was presented in JVET at 2200 on Thu 25 Jan, and agreed to be included in JVET-AG2027 :



[JVET-AG0054](https://jvet-experts.org/doc_end_user/current_document.php?id=13610) AHG9: On the case an SEI message included in multiple sei processing order SEI messages [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

This contribution discussed the case when an SEI message is included in two or more sei processing order SEI message and whether indirect preferred processing order can be established / inferred. The asserted case is in the situation when an SEI sei\_A is included in two different SEI processing order SEI messages spo\_1 and spo\_2. If spo\_1 includes another SEI message sei\_B in which sei\_A has lower processing order than sei\_B (i.e., the preferred order is sei\_A then sei\_B) and spo\_2 includes another SEI message sei\_C in which sei\_A has higher processing order than sei\_C (i.e., the preferred order is sei\_C then sei\_A). In such case, can the preferred processing order between sei\_B and sei\_C be inferred?

In this contribution, it is proposed to clarify the handling of such case. Two options are provided for the handling of the case. They are:

***Option 1:*** The scope of processing order value of a particular SEI message is within a CVS. For this option, the following should be applied:

1. Specify that when an SEI message is included in two or more SEI processing order SEI message in a CVS, the processing order values of that SEI message shall be the same in all those SEI processing order SEI message.
2. Add a note to explain that preferred processing order may be inferred / indirectly established from two or more SEI processing order SEI messages.

***Option 2:*** The scope of processing order value of a particular SEI message is within the SEI processing order SEI message that contains it. For this option, the following changes should be applied:

1. Specify a constraint such that two SEI messages are included in two different SEI processing order SEI messages, the preferred processing order between the two SEI messages in all those SEI processing order SEI message shall be consistent.
2. Add a note to provide example of such consistency.

It was commented that this would only occur when the po\_id is different. It was further commented that imposing always the same processing order might be too restrictive. For example, the two POs could have different amount of complexity, and a decoder may select one of them depending on its processing ressources..

Not obvious that there is a problem, and a restriction is not necessary.

No action.

Subsequent contributions initially reviewed in BoG (no separate BoG report)

[JVET-AG0105](https://jvet-experts.org/doc_end_user/current_document.php?id=13661) AHG9: On the SEI processing order SEI message [Y. Gao, Y.-X. Bai, S.-W. Xie, M.-H. Jia, C. Huang, P. Wu (ZTE)]

Chaired by J. Boyce.

This contribution proposes the following changes related to the SEI processing order SEI message:

Adding a flag (po\_sei\_prcessing\_degree\_flag[i]) for processing method to specify whether or not the decoder should ignore the entire SPO SEI message if it cannot interpret or does not support the functionality indicated by any indicated SEI message that has po\_sei\_importance\_flag[ i ] equal to 1 and po\_sei\_wrapping\_flag[ i ] equal to 0. This flag is set to 1 when the decoder should ignore the entire SPO SEI message and set to 0 when the decoder should still process the SEI messages that po\_sei\_processing\_order[ i ] is less than that of the previously mentioned important SEI messages in processing order.

An expert asserted that it is possible to accomplish the same goal using the existing syntax, although may be more complex when there are many SEI messages included in the SPO SEI message. No action.

Further study on improving the SPO SEI message is encouraged.

[JVET-AG0165](https://jvet-experts.org/doc_end_user/current_document.php?id=13721) AHG9: On the processing order nesting SEI message [L. Chen, O. Chubach, Y. Huang, S. Lei (MediaTek)]

Chaired by J. Boyce.

This contribution proposes to modify the syntax structure for signlling the number of the processing-order-nested SEI messages and the processing orders for each of the target processing order id pon\_target\_po\_id[i], adding flexibilities for multiple SEI processing order SEI messages with different values of po\_id in CVS.

An expert indicated that this proposes a different approach than had been intentionally selected in the original design and that the desired functionality could be accomplished by sending multiple PON SEI messages. In the existing design it is possible to use the same SEI message payload multiple times without requiring the payload being sent multiple times.

The proponent suggested adding a clarifying note in the semantics to describe the flexibility supported when multiple PON SEI messages are sent. This is delegated to the editors.

[JVET-AG0166](https://jvet-experts.org/doc_end_user/current_document.php?id=13722) AHG9: On grouping the SEI prefix data as a new SEI message [L. Chen, O. Chubach, Y. Huang, S. Lei (MediaTek)]

Chaired by J. Boyce.

Considering both the wrapping flag and the prefix flag are used as the special cases for the SEI processing order SEI message, to further simplify the syntax structure of the SEI processing order SEI message, this contribution proposes to group the SEI prefix data related syntax structure as a new SEI message – the processing order prefix matching SEI message.

The modified text for the SEI processing order SEI message in this document (JVET-AG0166) and a separate contribution JVET-AG00165 are combined in another contribution JVET-AG0168 for overall clarity and completeness.

An expert asserted that the proposed change would require that a decoder associate the SPO SEI message and the proposed new SEI message. Another expert suggested that a specified ordering of the SEI messages could make the association easier. No action.

[JVET-AG0168](https://jvet-experts.org/doc_end_user/current_document.php?id=13724) AHG9: Proposed update for the SEI processing order SEI message [L. Chen, O. Chubach, Y. Huang, S. Lei (MediaTek)]

Proponent did not request presentation of this contribution.

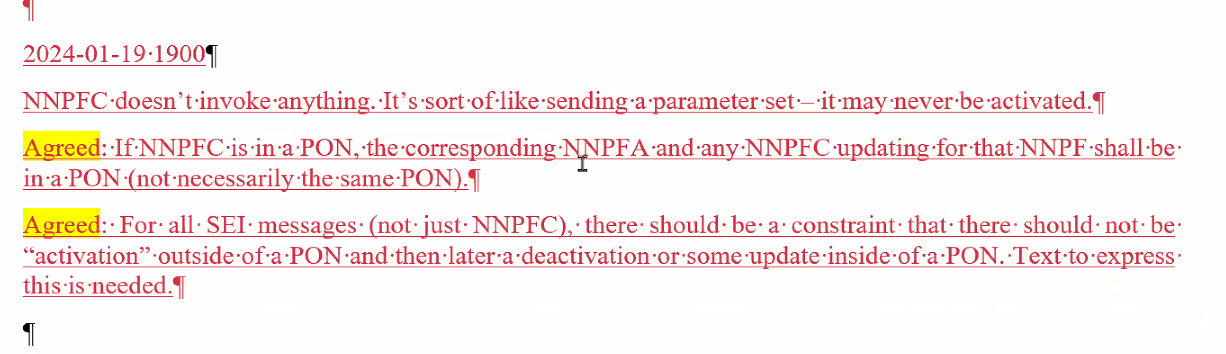
[JVET-AG0169](https://jvet-experts.org/doc_end_user/current_document.php?id=13725) AHG9: Comments on use of NNPF SEI messages in the SEI processing order SEI message [L. Chen, O. Chubach, Y. Huang, S. Lei (MediaTek)]

Chaired by J. Boyce.

This contribution proposes the semantic text for indicating NNPF SEI messages in the SPO SEI message.

Some clarification may be needed, such as NNPF update and activation when inside an SPO SEI message.

Initial results of side activity organized by G. Sullivan were reported Monday 22 January 2250.



It was asked if the opposite of the last bullet (activation inside, deactivation outside) should also be prohibited?

It was reported that had been discussed but was deemed not to be a problem, because a legacy device would not activate it, so deactivation of something that was not active is harmless.

To be left to discretion of editors to develop text.

[JVET-AG0180](https://jvet-experts.org/doc_end_user/current_document.php?id=13736) AHG9: On the SEI processing order SEI message [Y. Sanchez, R. Skupin, C. Hellge, T. Schierl (HHI)]

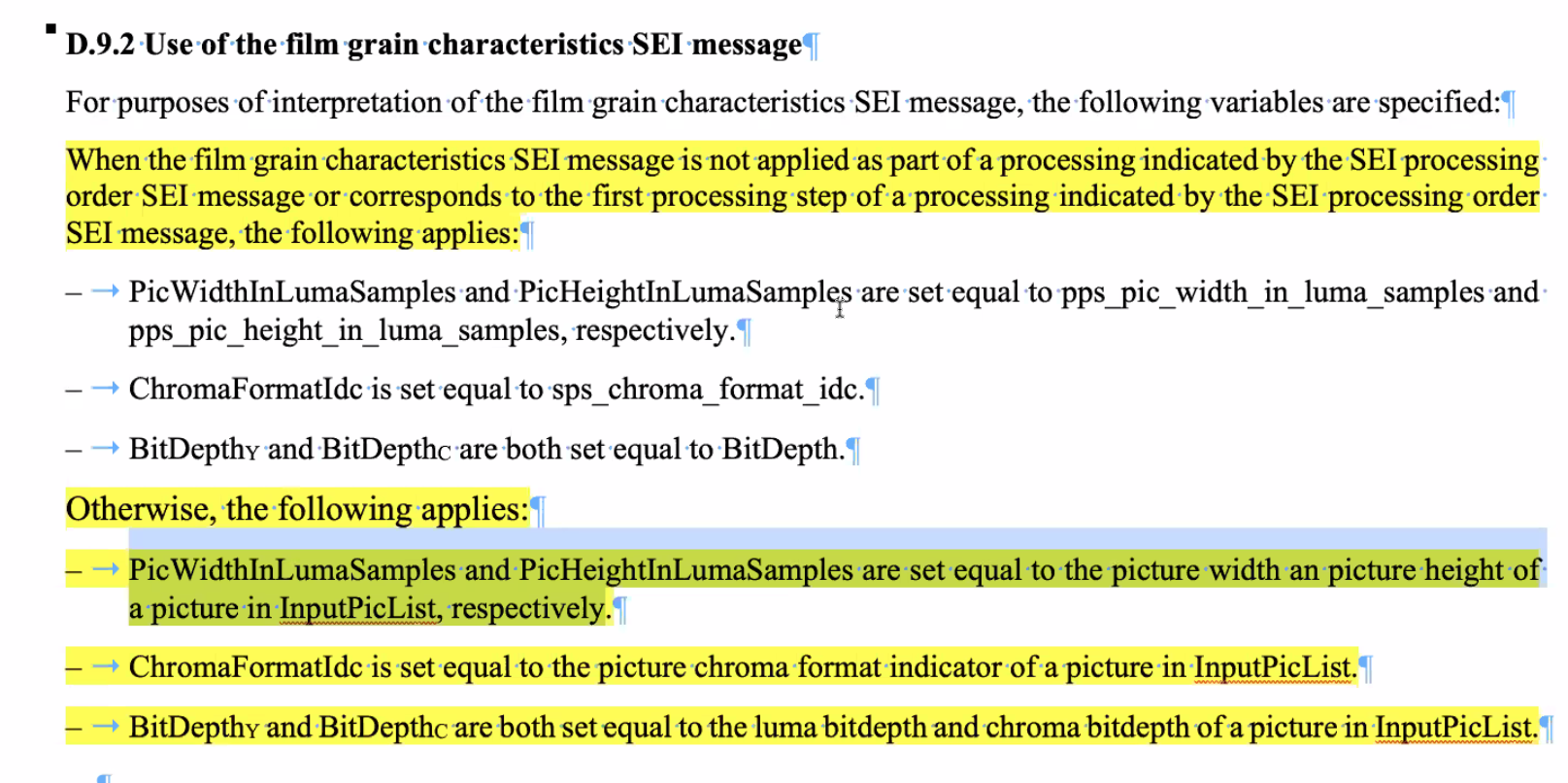
Chaired by J. Boyce.

In this document a proposal is described that updates the SEI messages allowed to be used with the SEI processing order SEI message and that suggest modifications to clearly specify the input to the different processing steps of the process indicated by the SEI processing order SEI message.

It was asserted by an expert that filler data SEI was intentionally included. Film grain SEI references decoded pictures while most other SEI messages reference cropped output pictures. The reason might be because the processing of film grain should be aligned with the transform blocks, while the cropped region may be smaller. This requires some clarification.

After offline discussion between experts, the aspect of cropped output pictures was identified to be relevant. A follow-up presentation was given at 2210 on Thu 25 Jan.

As a conclusion, the following solution was suggested:



It was commented that, as there is no expectation of a normative output, the misalignment with transform block boundaries would not be too severe in case of the FGC SEI.

It was agreed that the change as shown above (from JVET-AG0180) shall be included in JVET-AG2032 TuC.

## AHG9: SEI messages related to generative face video (3)

Contributions in this area were discussed at XXXX–XXXX on XXday XX Jan. 2024 (chaired by XXX).

See notes under 4.16

[JVET-AG0087](https://jvet-experts.org/doc_end_user/current_document.php?id=13643) AHG9: On the generative face video SEI message [M. M. Hannuksela, F. Cricri, H. Zhang (Nokia)]

[JVET-AG0088](https://jvet-experts.org/doc_end_user/current_document.php?id=13644) AHG9: Usage of the neural-network post-filter characteristics SEI message to define the generator NN of the generative face video SEI message [M. M. Hannuksela, F. Cricri, H. Zhang (Nokia)]

[JVET-AG0203](https://jvet-experts.org/doc_end_user/current_document.php?id=13759) AHG9/AHG16: Common text for proposed generative face video SEI message [J. Chen, B. Chen, Y. Ye (Alibaba), S. Yin, S. Wang (CityU), S. McCarthy, P. Yin, G.-M. Su, A. K. Choudhury, W. Husak, G. J. Sullivan (Dolby)]

## AHG9: Source picture timing information SEI message aspects (4)

Contributions in this area were discussed at 0515–0740 on Friday 19 Jan. 2024 (chaired by JRO).

[JVET-AG0070](https://jvet-experts.org/doc_end_user/current_document.php?id=13626) AHG9: Comments on Source Picture Timing Information Message [S. Deshpande, J. Samuelsson-Allendes (Sharp)]

Some modifications are proposed to the Source Picture Timing Information SEI Message. Following is proposed:

* Proposal 1: It is asserted that the current signalling in source picture timing SEI message includes only an integer scale factor for temporal sublayer interval. It is proposed to signal the scale factor for temporal sublayer interval as a numerator and a denominator.
* Proposal 2: It is proposed to signal spti\_num\_units\_in\_elemental\_interval with a minus 1 coding to avoid signalling the value 0.

Alternatively, it is proposed to add a constraint to prevent signalling the value 0 for spti\_num\_units\_in\_elemental\_interval.

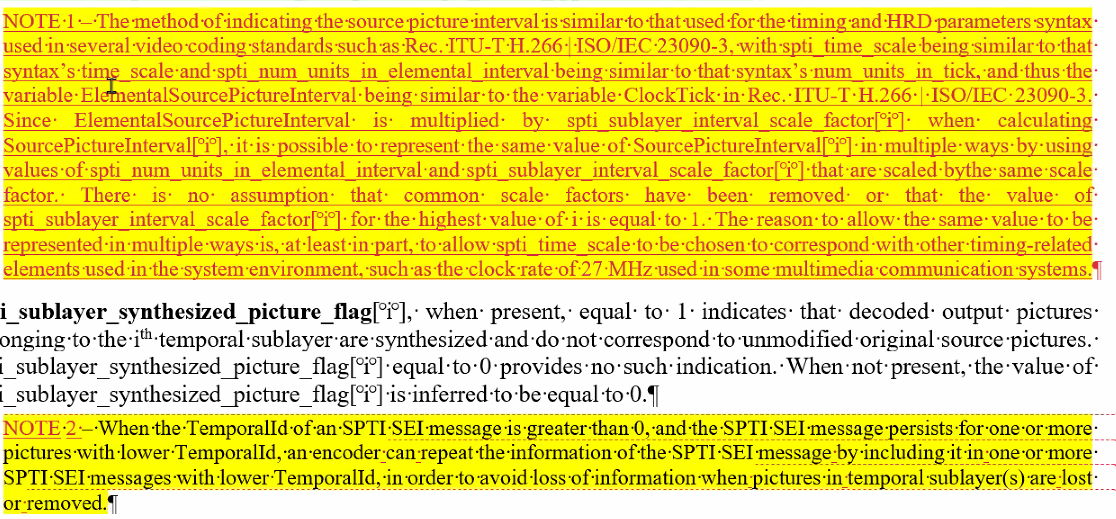
* Proposal 3: An asserted bug-fix is proposed for the derivation of the variable temporalReversalFlag.

Proposal 1:

It was commented that elemental source picture interval can be a very high number (e.g. 27 MHz), such that the problem may not exist.

The contributor assumes that elemental source picture interval refers to the frame rate of the camera, and gives the example where a sequence that was captured with a high speed camera is upsampled and comes in a higher temporal layer. There is however not such a constraint in the SEI message semantics. If post processing is done, the elemental source picture interval should refer to the post processed sequence’s highest temporal layer.

After similar discussion in the previous meeting, a note had been put into the specification text clarifying this:



(it was found there is an error in the note above which will be fixed in v4 of JVET-AF2032)

By choosing elemental source picture interval as being a very large number (seeing it as a clock rather than the camera’s frame rate), the case that is mentioned would not occur.

It was suggested that changing the names of syntax elements and variables might resolve a possible misunderstanding. The SEI message does not really need to refer to physical properties of a camera, but rather expresses how fast the content shown in the sequence should refer to the time axis of its interpretation. This could also be content that was generated synthetically.

No need for action, but editorial clarification would be desirable.

Proposal 2: It is proposed to either express the syntax element spti\_num\_units\_in\_elemental\_interval as “minus1”, or forbid the value 0.

The original proponents mention that by purpose the definition was done consistent with definition in HRD, which disallows the value 0. This was obviously forgotten.

Agreed to go for option 2 (forbidding value 0).

Proposal 3: This is an obvious bug (handling the case spti\_source\_type\_present\_flag equal 0).

Agreed to implement this, editors to select one of the two options proposed (or equivalent)

[JVET-AG0082](https://jvet-experts.org/doc_end_user/current_document.php?id=13638) AHG9: On Source Picture Timing SEI message [J. Samuelsson-Allendes, S. Deshpande (Sharp)]

This input document relates to the Source Picture Timing Information (SPTI) SEI message included in the Technology under Consideration for future extensions of VSEI (draft 2). It is proposed to modify how TemporalId is used to derive SourcePictureInterval. It is asserted that the proposed modification makes the SPTI signalling more robust and able to handle more usage scenarios.

The source picture timing information SEI message offers content providers the means for providing receivers with information about the source picture timing, which may be different from the output picture timing. In the current version of the SPTI SEI message draft the source picture interval is defined for “consecutive pictures” and mentions “the previous decoded output picture”. We assert that it might not always be clear to the receiver which picture is meant by “the previous decoded output picture”. In particular in cases that involves temporal scalability, step-wise temporal access and/or drop of individual pictures due to congestion.

It is proposed to specify the previous picture as the one in temporal layer zero, as this is always present. It is pointed out in the discussion tha the proposed approach would deviate from the original intent. If it was always referring to TId zero, it might be necessary to also introduce a scale factor.

It is agreed that there is an issue in case that temporal layers are removed, but another solution would be necessary. After offline consideration, un update was made available in JVET-AG0082v2 (green highlighted text)

Agreed to include the modification from JVET-AG0082v2 in the SPTI SEI (which is to be included in prelim. WD of VSEIv4)

[JVET-AG0188](https://jvet-experts.org/doc_end_user/current_document.php?id=13744) AHG9: On source picture timing information SEI message specification text [J. R. Arumugam, L. Jawale (Ittiam), P. Yin, G. J. Sullivan, S. McCarthy (Dolby)]

This contribution proposes two options to resolve an asserted error introduced in the TuC for VSEI when an aspect of JVET-AF0069 was adopted at the previous JVET meeting. Specifically, it was agreed to change the size of the syntax element spti\_num\_units\_in\_elemental\_interval from u(32) to u(18). However, it was later noticed that the corresponding semantics provide an example of using a value of spti\_num\_units\_in\_elemental\_interval of 1,080,000, which is greater than the range supported by u(18). The syntax design had also been intended to be aligned with the timing and HRD parameters syntax, with spti\_num\_units\_in\_elemental\_interval corresponding to num\_units\_in\_tick, and num\_units\_in\_tick is coded as u(32). Part of the reason a large range of values is needed is for the the spti\_time\_scale (which corresponds to the time\_scale in the timing and HRD parameters syntax) to support the 27 MHz clock rate used in MPEG-2 Systems. The contribution proposes to resolve the mismatch by changing spti\_num\_units\_in\_elemental\_interval back to u(32). Another option is to change the example values used in the semantic for spti\_num\_units\_in\_elemental\_interval, but it is asserted that this approach would not provide the desired consistency with the timing and HRD parameters syntax.

At the previous JVET meeting, it was agreed for the source picture timing information (SPTI) SEI message to change the syntax element spti\_num\_units\_in\_elemental\_interval from u(32) to u(18). However, the corresponding semantics provide an example of using a value of spti\_num\_units\_in\_elemental\_interval of 1,080,000, which is greater than supported by u(18).

Two options are proposed. Option 1 is preferred by the authors.

* Option 1 – Revert to u(32). No change in existing semantics. These changes would help avoid confusion by providing consistency between the HRD specification, the SII SEI message, and the SPTI SEI message.
* Option 2 – Change value of spti\_num\_units\_in\_elemental\_interval of 1,080,000 to a value supported by u(18), for example 108,000. The example value of spti\_time\_scale could be changed similarly to 2,700,000 so that the elemental source picture interval remains 0.04 seconds.

Obvious bug. The intent of the change at the last meeting was saving some bits. This is not of highest relevance in an SEI message.

Agreed to implement option 1.

[JVET-AG0191](https://jvet-experts.org/doc_end_user/current_document.php?id=13747) AHG9: Reference software for source picture timing information SEI message [J. R. Arumugam, L. Jawale (Ittiam), P. Yin, S. McCarthy (Dolby)]

This contribution provides reference software for the source picture timing information (SPTI) SEI message to enable study of the SPTI SEI message by JVET experts. The software provided is based on the TuC branch of VTM 22.2 and the specification text in Technologies under consideration for future extensions of VSEI (version 2) (JVET-AF2032).

Decision(SW): Adopt JVET-AG0191

It was clarified that any software related to a technology specified in the TuC JVET-nn2032 should be qualified to be merged in the TuC branch, such that other experts can study this tehnology.

## AHG9: Encoder optimization information SEI message aspects (3)

Contributions in this area were discussed at 2320–0100+1 on Monday 22 Jan. 2024 (chaired by JRO).

[JVET-AG0083](https://jvet-experts.org/doc_end_user/current_document.php?id=13639) AHG9: On feature-based optimization type [C. Kim, Hendry, J. Lim, S. Kim (LGE)]

This contribution proposes to add feature-based optimization type in Encoder Optimization Information (EOI) SEI message. The feature-based optimization type indicates that the pictures have been pre-processed or encoded based on the quality of features. This helps clarify the criteria and properties of the encoder optimization.

It was asked if that would ever be used for the case of human consumption? Likely not.

Would it be better to signal that a special distortion criterion has been used, from which then, along with the information that something has been optimized for machine or human usage or both, the encode could draw conclusions.

It was asked what a decoder would conclude (or which action it would take) when knowing that feature based optimization was done.

Not clear that inclusion of that type would give benefit. If a decoder knows that something has been optimized for machine consumption, it would be good enough to use the decoded video.

[JVET-AG0086](https://jvet-experts.org/doc_end_user/current_document.php?id=13642) AHG9: On the encoder optimization information SEI message [M. M. Hannuksela, F. Cricri, H. Zhang (Nokia)]

The encoder optimization information (EOI) SEI message is one of the SEI messages included in the Technologies under Consideration (JVET-AF2032). This contribution proposes the following changes to EOI SEI message:

1. 2-bit indicators the eoi\_for\_human\_viewing\_idc and eoi\_for\_machine\_analysis\_idc are proposed to replace the respective flags present in JVET-AF2032. The values for the eoi\_for\_human\_viewing\_idc are specified as follows (and respective values are specified for eoi\_for\_machine\_analysis\_idc):

3: optimized for human viewing

2: suitable but not specifically optimized for human viewing

1: unsuitable for human viewing

0: unknown if the video is suitable for human viewing

1. It is proposed to enable indicating temporal sublayers where optimization has been applied as follows:
2. When an EOI SEI message potentially persists for multiple pictures, the temporal sublayers to which the EOI SEI message applies are indicated as follows:
3. A syntax element (eoi\_num\_sublayers) indicates the number of sublayers.
4. If eoi\_num\_sublayers is equal to 7, it indicates that the EOI SEI message applies to all sublayers.
5. Otherwise, eoi\_num\_sublayers indicates the number of sublayers to which the EOI SEI message applies and eoi\_min\_tid is present to indicate the smallest sublayer identifier to which the EOI SEI message applies.
6. The semantics of eoi\_persistence\_flag equal to 1 is updated to concern only the temporal sublayers indicated by eoi\_num\_sublayers and eoi\_min\_tid.
7. Constraints are added to avoid EOI SEI messages with conflicting syntax element values present in the same picture unit. It is allowed to have two EOI SEI messages with different SEI payload content in the same picture unit, when the EOI SEI messages apply different ranges of temporal sublayers.

The first part of the proposal allows better expressing to which extent something has been optimized for human or machine consumption, by replacing the current flags by 2-bit syntax elements.

It was asked which benefit the value 0 (unknown) would have? In combination of both syntax elements, a decoder might still draw useful conclusions. Also the combination 0/0 (both unknown) might be useful, drawing conclusions from optimization type. Combination 1/1 would not be useful.

Generally, the finer granularity is agreed to be beneficial. The proposed modification 1. was agreed for inclusion in TuC. It should be investigated if both syntax elements could be combined into one, avoiding unreasonable combinations (not with the goal of saving bits). Further study on this aspect.

Modification 2. allows more specific EOI for individual temporal sublayers (or ranges of sublayers).

It was generally agreed that defining EOI on sublayer basis could have advantages (and practical use cases, such as doing object based optimization only at certain sublayers). Several experts were requesting to think about alternative ways to accommodate this functionality.

Proponents were asked to update the contribution with more practical use cases. This was presented at 2230 on Thursday 25 Jan (v2 of JVET-AG0086).

Two options are presented (replace

1. Option a Using separate EOI messages, one per sublayer range
2. Option b Including all sublayer ranges in the loop

Two experts commented that the option a appears simpler, but both options are assessed to be appropriate.

As in general it appeared that the question of sublayer support for EOI might require more consideration, it was suggested to put this into TuC, from which it could later be transferred to the WD of VSEIv4 (where the EOI will be included).

Agreed to include option a in JVET-AG2032 TuC.[JVET-AG0213](https://jvet-experts.org/doc_end_user/current_document.php?id=13769) AHG1/AHG2/AHG8: On project management related to the encoder optimization information SEI message [M. M. Hannuksela, A. Aminlou, F. Cricri, H. Zhang (Nokia)]

This document proposes the following items:

1. Addition of the encoder optimization information (EOI) SEI message into a working draft of a VSEI amendment.
2. Addition of a description of the EOI SEI message in the draft technical report on optimization of encoders and receiving systems for machine analysis of coded video content (hereafter, referred to as the technical report).
3. Aligning the completion timelines of the technical report and VSEI v4 so that the technical report is not completed before VSEI v4.

Further discuss in context of JVET-AG0204 and timeline for technical report (4.10)

Also include consideration of JVET-AG0081

## AHG9: Object mask information SEI message aspects (1)

Contributions in this area were discussed at 0105–0120 on Tuesday 23 Jan. 2024 (chaired by JRO).

[JVET-AG0148](https://jvet-experts.org/doc_end_user/current_document.php?id=13704) AHG9: On object mask information SEI message [J. Chen, Y. Ye, S. Wang (Alibaba)]

This contribution studies the object mask information (OMI) SEI message in the technologies under consideration (TuC) for future extensions of VSEI, and it is proposed to separate the object mask identifiers from the auxiliary picture sample values to make the design cleaner and more logical.

Further, it is suggested to establish at least a preliminary working draft of additional SEI messages for VSEI and include the OMI SEI message in it.

In the 2nd version, it is further proposed to include OMI SEI message in the technical report (TR) on optimization of encoders and receiving systems for machine analysis of coded video content.

Further discuss in context of JVET-AG0204 and timeline for technical report (4.10)

The suggested change is to have mask IDs as increasing numbers starting from 0, and the numbers put into the aux picture is independent from that. It was pointed out that the bit depth of aux\_sample\_values also needs to be signalled. Also the value num\_Aux\_Layer may need to be signalled. Typo for cancel flag needs correction. Agreed to implement this change in TuC.

Post-discussion note: Clarification about possible parsing dependency. It was later confirmed that there was a parsing issue, but this was resolved in JVET-AG0148v3.

Agreed to include in the preliminary WD vor VSEIv4 JVET-AG2034.

Also agreed to include a paragraph about the SEI message in the TR JVET-AG2030 (wording in JVET-AG0148 should be editorially improved).

As a general remark, moving software from TuC branch to VTM branch might be considered when a CD is issued.

## AHG9: SEI message aspects related to film grain (5)

Contributions in this area were discussed at XXXX–XXXX on XXday XX Jan. 2024 (chaired by XXX).

See notes under 4.12

[JVET-AG0101](https://jvet-experts.org/doc_end_user/current_document.php?id=13657) AHG9: Film grain adaptive SEI message [Y. Gao, S.-W. Xie, Y.-X. Bai, M.-H. Jia, C. Huang, P. Wu (ZTE)]

[JVET-AG0140](https://jvet-experts.org/doc_end_user/current_document.php?id=13696) AHG9/AHG13: FGS Extension SEI message for spatial adaptation [G. Teniou, S. Wenger, A. Hinds (Tencent)]

[JVET-AG0160](https://jvet-experts.org/doc_end_user/current_document.php?id=13716) AHG9/AHG13: FGS Extension SEI message useful descriptors [G. Teniou, S. Wenger, A. Hinds (Tencent)]

[JVET-AG0215](https://jvet-experts.org/doc_end_user/current_document.php?id=13771) AHG9/AHG13: Region-dependent film grain characteristics [P. de Lagrange, E. François, M. Le Pendu, C. Salmon-Legagneur (InterDigital)] [late]

[JVET-AG0328](https://jvet-experts.org/doc_end_user/current_document.php?id=13901) AHG9/AHG13: unified proposal on film grain regions characteristics SEI message [G. Teniou, S. Wenger, A. Hinds (Tencent), P. de Lagrange, E. François, D. Doyen, M. Le Pendu, C. Salmon-Legagneur (InterDigital), P. Wu, Y. Gao, S. Xie, Y. Bai, M. Jia, C. Huang (ZTE)] [late]

## AHG9: Other SEI topics (11)

Contributions in this area were discussed at 0640–0710, at 0730-0935, and at 2100–XXXX on Friday 19 Jan. 2024 (chaired by JRO).

[JVET-AG0044](https://jvet-experts.org/doc_end_user/current_document.php?id=13598) AHG9: Copyright SEI message [S. Wenger, A. Hinds, G. Teniou (Tencent)]

Proposed for VSEI is an SEI message to convey copyright information in a video stream, to enable marking the video stream with a copyright related statement such as a copyright notice. Two alternative proposals are offered: free-form st(v) text, and syntax-structured copyright information. No preference between the two options were indicated by the authors.

It is emphasized by proponents that associating copyright with elementary streams becomes more important nowadays, observing increased occurrence of media manipulation.

It was reported that H.263 has such an SEI message, but none of the newer standards, nor VSEI. Otherwise, similar mechanisms exist in the systems layer. “Proposal 1” is similar as the definition of H.263 (just plain text).

It is commented that such an SEI message would not prevent from violating copyright. It might help copyright owners to enforce their rights by other means, or warn prospective violators not to manipulate streams.

It was commented that people could also use a text SEI message for that purpose.

Decision: Include JVET-AG0044 proposal 1 in TuC

[JVET-AG0045](https://jvet-experts.org/doc_end_user/current_document.php?id=13599) AHG9: AI marking SEI [S. Wenger, A. Hinds, G. Teniou (Tencent]

The contribution proposes syntax and semantics for an SEI message addressing the purpose of marking one or more coded pictures as created or modified by generative AI. It follows up on the concerns raised in Hannover related to contribution JVET-AF0145. As requested in the previous meeting, the subject contribution contains evidence and references to legislative initiatives mandating the marking of content created or modified by generative AI, as well as a short summary why a copyright SEI message would not be an appropriate mechanism to convey AI marking. Two solutions are proposed: a free-form text SEI marker (comparable what was proposed in JVET-AF0145) and a structured syntax containing fields that are currently envisioned in the legislative processes reviewed by the authors.

(from powerpoint – issues raised in Hannover as pertaining to the labelling aspect of JVET-AF0145)

Q1. Why one SEI message for two different purposes? Use SEI payloadType as demux point.

A1. Proponents agree. Implemented here and in companion contribution related to AI instruction

Q2. Why put marking in video elementary stream (ES) and not in the system layer

A2. Conceivably, only one ES of a multimedia stream can get modified by AI, hence only that stream should be marked for truthful representation of AI involvement

Q3. Why SEI when proposed regulation requires “conspicuous marking”

A3. Overly elaborate marking (including textual information required by proposed regulation) would interfere with artistic intent. Proposed regulation generally appreciates that and requires only marking of AI involvement (like: a logo) in video content, but also requires more details in metadata. This contribution concerns metadata.

Q4. Why not use the “Copyright” SEI message.

A4: Please see contribution document on why the copyright SEI is not appropriate. In short. Copyright statements and AI marking address distinct legal concerns and are, to some extent, incompatible with each other.

It was suggested that some editorial improvements would be necessary (e.g., notes are always informative)

Decision: Include JVET-AG0045 proposal 1 in TuC

[JVET-AG0049](https://jvet-experts.org/doc_end_user/current_document.php?id=13603) AHG9: SEI message extension of VVC for computer-generated hologram use [H. Kojima, K. Nonaka, R. Koiso, K. Kawamura, H. Kato (KDDI)]

In JVET-AE0182, an extension of SEI message on VVC for realizing computer-generated holography use was proposed. This proposal is an updated version of JVET-AE0182 with additional preliminary study experiments and modifications to the SEI message, such as persistence scope variables. Transmission of object wave, whitch is the intermediate data representation of computer-generatad hologram, is suitable for the practical use in terms of high computational complexity on object wave generation and the data flexibility of object wave. Therefore, this contribution describes a method of transmitting object waves using existing video codecs, and show the flow of the object wave transmission for computer-generated hologram use and the results of preliminary study of the flow by using VTM. The proposed extension provides the ability to compress the object wave and convert object wave into computer-generated holograms tailored to the playback device, enabling a wide range of use of computer-generated holograms.

Results are shown comparing RA vs. AI encoding, showing approx. 50% bit rate reduction. It was commented that this is significantly less benefit of using temporal redundancy than with normal video. Also, the phase has a spatially irregular “noisy” structure, and requires significantly more rate than video. Compressed rates are in the range of Gbit/s.

It was asked how commonly this format is used. One expert commented that it may be very specific for one type of holographic displays, more in prototype status. Would be too early defining a specific SEI message for that.

Several experts expressed that the object wave method looks interesting. However, it was pointed out that it would be more desirable defining a format that could be used over a wider range of holographic displays.

Further study recommended – from market perspectives, it is not foreseen to be appropriate for defining an SEI at the current point.

[JVET-AG0051](https://jvet-experts.org/doc_end_user/current_document.php?id=13607) AHG9: On design for new SEI RBSP and SEI message [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

This contribution proposed new SEI RBSP and SEI message to be included in the TuC for future extensions of VSEI to overcome both issues of handling large SEI payload size and important / application-required SEI message.

The proposed points are:

1. Define a new RBSP for SEI which may be called versatile SEI RBSP. A versatile SEI RBSP contains only one versatile SEI message.
2. In addition to containing one versatile SEI message, the bits of the first byte of a versatile RBSP are allocated for signalling information that may be important to be accessible for entities other than the decoder such as MANEs or any other system interfaces. As for now, we may assign one flag to indicate whether the SEI message is important / application-required so that it should not be randomly removed while the rest 7 bits are reserved.
3. A versatile SEI message data structure is designed as follow:
   1. Payload type
   2. Payload data
4. A versatile SEI RBSP may be carried in NAL unit that is defined by codec specification. Existing codecs may add two new NAL unit types to carry prefix versatile SEI RBSP and suffix versatile SEI RBSP.

From the discussion, defining a new RBSP is seen as an alternative for the lsei message (which needs to signal the size). Both could not be interpreted by legacy devices, but the RBSP approach could be more flexible.

Such a new RBSP could also include several SEI messages.

An advantage compared to lsei is the lower latency. It was commented that both would require defining a new NAL unit type in VVC and other standards. Otherwise, like the lsei approach would require some more changes in VVC than just adding the codepoint.

The second aspect of this proposal (vsei\_importance for MANE) is believed to have some flaw according to proponents of lsei.

It is unclear at this moment which of the both approaches is better. At the time being, both might coexist in TuC, but put them in an annex with a clear statement that this is currently not seen to be relevant for existing standards due to the fact of touching basic concepts of SEI transport, and kept for further study in the context of potential future standards.

Agreed to include JVET-AG0051 in TuC, and moved to a separate section together with lsei. Proponents of both approaches are asked to work out possible combination.

[JVET-AG0081](https://jvet-experts.org/doc_end_user/current_document.php?id=13637) AHG9: On signalling privacy protection information in SEI message [C. Kim, Hendry, J. Lim, S. Kim (LGE)]

This contribution proposes to have a mechanism to convey privacy protection information (PPI) in an SEI message. The privacy protection can be in the form of removal and / or alteration of personal information. This can affect both machine analysis and human perception. Therefore, providing information about the removal or alteration of personal information can be valuable. Privacy protection information (PPI) indicates the methods applied to protect privacy and the type of personal information intended to be protected.

Two options are proposed to convey PPI privacy protection information in an SEI message as follows:

* Option 1: Define privacy protection optimization as a new optimization type in Encoder Optimization Information (EOI) SEI message.
* Option 2: Define a new SEI message to carry the privacy protection information.

It was commented that removal should be an another replacement type

It was commented that definition of info type (what kind of object specifically is being modified) has suc a large amount of variations that its precise definition is almost impossible (and possibly not necessary, as a device that knows which replacement type was used may be able to infer that.

It was commented to investigate whether JPEG has done something like this (in JPsec?)

Generally, including the general principle (with modifications as described) in EOI (option 1) is agreed to be the best option for now, would be beneficial for both huma and machine consumption, like other elements of that message. Agreed to be included as optimization type in EOI SEI message (in next version TuC).

[JVET-AG0107](https://jvet-experts.org/doc_end_user/current_document.php?id=13663) AHG9: On phase indication SEI message [T. Chujoh, T. Ikai (Sharp), K. Kawamura (KDDI)]

VVC can change the picture resolution in a sequence by using RPR; however, most video applications display decoded pictures with a constant resolution, even with decoded resolution changes dynamically. When the same phase information for the resampling process of RPR is not identified between the transmitting and receiving sides, pixels of the picture may shift in time when the resolution changes dynamically in the sequence. To avoid this phenomenon, the phase indication SEI message has been adopted in VSEI to share the phase information between the transmitting and receiving sides. The remaining problem is that the bitstreams, defined as the phase information implicitly but without the SEI, are not interoperable. This contribution proposes the following:

1. The phase indication SEI message is moved to the VVC text, which is made as a normative SEI message.
2. Syntax elements of the phase indication SEI message have the inferred values. Their values mean the co-sited positions. Then, even if there is no phase indication SEI message in the bitstream, the phase information is defined.

These proposals do not change the conformance bitstreams and the reference software.

It was commented that it does not make a difference whether something is in VSEI or VVC. It was further commented that VUI may be a better place. It also seems to be more a display issue which is out of scope of our standards.

It was asked why the co-sited position should be defined as default when the SEI is not present? It would still have the undesired effect when the downsampler generates a different position. However, it might be assumed that an implement would use a downsampler that matches the inferred value when no SEI is present. This could be augmented with an informative note for the downsampler in case of RPR.

It was asked how serious the problem is? For some sequences, it may not even be visible. Further, other effects such as loss of structure is happening at the switching point.

Further study on the possibility of defining a default phase position when the SEI message is not present, and add an informative remark what an encoder should do. Revisit if time allows.

[JVET-AG0167](https://jvet-experts.org/doc_end_user/current_document.php?id=13723) AHG9: Text prompt for generative AI SEI [A. Hinds, G. Teniou, S. Wenger (Tencent)]

Text prompts instruct an AI model to produce a desired output in generative AI image and video generator applications. In such applications the prompts can be accompanied with a sequence of images or video frames upon which the AI modifies the imagery according to the instructions provided in the prompt. The prompt may direct the model to alter the imagery in any number of ways depending on how the model has been trained. For example, a simple prompt might direct an AI to alter the ambient daylight for one or more frames of a cityscape scene to give the appearance of the scene occurring during a sunset (as opposed to early morning or mid-day). This contribution proposes that a new SEI be created to store the prompt data directly into the video stream for generative AI applications. Such an SEI serves to facilitate a rigorous association of the prompt data with the precise video frames that should be targeted by the AI consistent with industry best practices. For example, dynamic HDR SEI messages are likewise carried in the video stream as opposed to carriage in a sidecar file.

It was asked what the use case is? Examples are given for realtime applications, such as replacing or hiding certain parts of scenes. Typically the SEI would be provided by the content producer. Unlike the case where the generative AI would be run before encoding (such that the consumer would always see the modified scene), this would be optional or could be selected based on user preference.

It was further reported that such processing would not necessarily be performed at the end user’s decoder, it might also be executed at a middle point, where a kind of transcoding would be executed inserting the modified content.

It was commented that generative AI is developing and improving quickly, and over time there will likely be evolution how it is controlled (instructed) to make the result potentially less random.

It was asked, considering that lots of generative AI applications are existing, how much the generated result is matching the result. It is unlikely that it would really be operated at the end user site. More likely, it is generated somewhere in the distribution chain, where it is more controllable and can guarantee a certain degree of quality of the result.

It was commented that a problem might be that certain restrictions might be imposed to prevent inappropriate text.

The whole application scenario needs to better understood. In a commercial service between a service provider and the operator of a middle box, a private data SEI message could be used. Not clear that a dedicated SEI message would be necessary for this with the intent of interoperability in a wider sense between various devices.

It was commented that also NNPF could be used to define generative AI.

Further study.

[JVET-AG0182](https://jvet-experts.org/doc_end_user/current_document.php?id=13738) AHG9: JPEG segments SEI message [P. de Lagrange, D. Doyen, E. François, F. Urban, C. Salmon-Legagneur (InterDigital)]

During the 32nd JVET meeting, the contribution JVET-AF0141 proposed new SEI messages to embed externally defined camera-related metadata structures (JFIF, EXIF or XMP). This is said to provide somewhat redundant functionality, with EXIF additionally raising concern because of the need to refer to specifications from non-SDO bodies like CIPA and JEITA. This contribution proposes to replace or merge the JFIF SEI message with a new SEI message embedding generic JPEG segments, that is said to go around these issues while providing more functionality. Indeed, some JPEG segments named “APP segments” are free to use for applications with any kind of payload and are commonly used to embed such camera-related metadata. This solution is also reported to offer additional possibilities such as embedding full JPEG pictures in a video bitstream, for instance for indexing purpose or hi-resolution snapshots.

A comment was made that the statement about “redundant functionality” about JFIF, EXIF, XMP is not really true.

Another comment was made that this significantly deviates from the original intent of SEI messages, which should not be seen as yet another transport mechanism. JPEG segments may contain something that has no relation at all with the video (other than the thumbnails in JFIF).

If the intent was to find a way to resolve the EXIF referencing problem, this should not be resolved by allowing lots of other elements to come into SEI messages.

No action on JVET-AG0182

[JVET-AG0183](https://jvet-experts.org/doc_end_user/current_document.php?id=13739) AHG9: TIFF data SEI message [P. de Lagrange, D. Doyen, E. François, F. Urban, C. Salmon-Legagneur (InterDigital)]

During the 32nd JVET meeting, the contribution JVET-AF0141 proposed new SEI messages to embed externally-defined camera-related metadata structures, for example EXIF. One concern raised about the EXIF SEI message proposal was that it includes a reference to specifications from non-SDO bodies (CIPA and JEITA). Based on the observations that EXIF data is actually a TIFF structure, and that TIFF is both an ITU-T SG16 publication and ISO standard 12639, this contribution proposes an alternate solution to go around this issue, by designing a new SEI message embedding generic TIFF data instead. This solution is also said to reduce functional redundancy with the XMP SEI message, while allowing broader use cases, including additional possibilities such as embedding full TIFF pictures in a video bitstream, for instance to attach information in the form of pictures (like scanned documents).

It was commented that EXIF is still further developing, and its current version contains lots of metadata that TIFF files do not carry.

It was suggested that we should rather target finding a solution referencing EXIF appropriately by seeking communication with ISO and ITU authorities.

Scepticism was expressed whether the statement above about TIFF being an ITU standard is correct. If it exists, it may be an outdated version.

A TIFF file could be very large. No urgent need is seen to put TIFF images into a video stream.

[JVET-AG0184](https://jvet-experts.org/doc_end_user/current_document.php?id=13740) AHG9: Text comment SEI message [P. de Lagrange, D. Doyen, E. François, F. Urban, C. Salmon-Legagneur (InterDigital)]

Chaired by J. Boyce.

This contribution proposes a new SEI message allowing to embed text comments in a video bitstream. Examples of use cases are annotations for parts of a video sequence, descriptive annotations for indexing purpose, user generated comments. In the proposed "text comment SEI message”, text comments are associated with ids to allow overlapping persistence (and potentially defining different types of comments).

Possible use cases include annotations, carrying technical data, user generated comments.

Has some similarity to AG0184 AI marking, but has different syntax for carrying text. AG0184 uses st(v), this contribution used b(8).

This contribution also allows overlapping text labelling, so persistence could differ for individual text comments.

H.263 Annex W had a feature for carrying text with different purposes.

Annotated regions SEI message also supports text labels, but limits to 255 characters. AR SEI also has an indication of what language is used. It was suggested that having a language tag could be useful.

Systems specs already have timed text feature, which includes presentation information for displaying the text. Would be useful to understand the systems timed text features better to understand any differences in capability with the contribution.

It would be possible to have a purpose id that allows text to be used for multiple purposes, potentially including copyright and AI marking. External means could potentially be used for signalling the purpose, for example based on the id. Another alternative is to use separate SEI messages for each purpose but have a consistent text signalling mechanism, perhaps using a common syntax function. There is not a shortage of SEI message payload type values.

Further study encouraged.

[JVET-AG0232](https://www.jvet-experts.org/doc_end_user/current_document.php?id=13805) AHG9: Support of non-parallel MPI layers in the MPII SEI message [Y. Li (SJTU), Y.-K. Wang (Bytedance), Y. Xu, K. Yang (SJTU)] [late]

Chaired by J. Boyce.

This contribution proposes an extension to the multiplane image information (MPII) SEI message in the VSEI TuC (JVET-AF2032), such that the multiplane image layers can be either parallel or non-parallel to each other.

In v2 of this document, some changes were made for optimizing the number of bits used for signalling of the normal vectors of multiplane image (MPI) layers for the case when the MPI layers might be non-parallel. And some text errors were fixed.

JVET-AF0167 provided SW for MPII SEI and provided instructions for end-to-end example to render using an open source implementation for Adaptive MPI.

MPI rendering is simpler for parallel planes than for the proposed non-parallel planes.

Further study requested to study difference in rendering complexity vs. parallel planes and to justify the need for standardization for this variant of MPI, including providing information about open source software for rendering MPI non-parallel planes.

[JVET-AG0332](https://jvet-experts.org/doc_end_user/current_document.php?id=13905) AHG9: Test material for the multi-plane image information SEI message [W. Husak, S. McCarthy, T. Lu (Dolby)] [late]

This contribution provides information for test material for testing of the multi-plane image information SEI message described in the TuC for future extensions of VSEI.

Only one sequence. Using it would require signing a license agreement with Dolby. Would not be possible to provide it via JVET ftp.

It was commented that the license agreement would allow Dolby to modify the content

The functionality could also be tested with MIV content, but that has higher number of texture layers (64, whereas this has 16).

It was commented that for MIV testing, usually also a “pose trace” is provided.

Modified MIV content (with only 16 layers, such that it can be used with MPI) is available for download, and the link is included in the software package. This is also only one sequence.

It was commented that using only one or two sequences to investigate the benefit of a technology may be not sufficient.

Add a mandate to AHG9 about discussing the testing of MPI, and prepare a viewing at Rennes, and coordinate with the MIV AHG of WG 4 how it could be compared to corresponding MIV coded content.

## Non-SEI HLS aspects (5)

Contributions in this area were discussed at 2100–XXXX on Monday 22 Jan. 2024 (chaired by JRO).

[JVET-AG0046](https://jvet-experts.org/doc_end_user/current_document.php?id=13600) Application-required NAL Units [G. Teniou, S. Wenger (Tencent)]

The present contribution follows up on JVET-AF0149 from the Hannover meeting by Tencent. JVET-AF0149v4 was briefly discussed om Hannover but referred to further study, as substantial changes were introduced during the Hannover meeting. This contribution is substantially similar to option 2, sub-option with req\_nu\_count\_minus1. It is argued that certain NAL unit types or SEI message types can be ignored by an encoder but are in practice required for an underlying system to work. Such NAL units or SEI messages are called “required NAL units”. Homogenous system standards frequently make such NAL units and/or SEI messages mandatory to implement and use in both the sending and receiving system. If a bitstream designed for such a homogenous system were transported in a heterogenous environment, then middleboxes in the network may not be aware of the required nature of the NAL units and dispose of them. The contribution argues this problem can be addressed by introducing a prefix NAL unit type that marks n following NAL units as required.

From the presentation: An example where this would be needed is the “hybrid” concept in VCM, where VVC intra picture coding would be replaced by another method for which the coded picture might be transported in an SEI message. Inter coding would be done by VVC.

Another example given are mandatory SEI messages in DVB specs. It is however commented that DVB streams are not typically transported via MANEs.

Some potential problems were pointed out with the concept of inserting a previously undefined NUT at a leading position in an AU.

It was further pointed out that with the current design a MANE would not be able to interpret.

Proposal withdrawn.

It was requested to bring the proposal back only when a use case exists where an SEI message needs to survive when transported via a MANE. This does not seem to be the case with existing SEI messages.

[JVET-AG0077](https://jvet-experts.org/doc_end_user/current_document.php?id=13633) AHG9: On Picture Modality Information [J. Gao, H.-B. Teo, C.-S. Lim, K. Abe, V. Drugeon (Panasonic)]

This contribution proposes the following items on the picture modality type included in the technologies under consideration for future extensions of VSEI (version 2), JVET-AF2032.

* Common Aspects:

1. Signal minimum and maximum wavelength which indicates the spectral band of optical radiation represented by the pictures in the CLVS, by adding the following syntax elements:

* Add **vui\_min\_wavelength\_mantissa** and **vui\_max\_wavelength\_mantissa** to specify the mantissa part of the min and max wavelength indicating the spectral band of optical radiation represented by the pictures in the CLVS.
* Add **vui\_min\_wavelength\_exponent\_plus15** and **vui\_max\_wavelength\_exponent\_plus15** to specify the exponent part of the max and min wavelength indicating the spectral band of optical radiation represented by the pictures in the CLVS.
* The unit of wavelength is specified in meters.

1. Add **vui\_modality\_type\_extension\_bits** and **vui\_reserved\_modality\_type\_extension** for future extension of picture modality information.
2. Change the type of picture modalities from thermal infrared picture to ultraviolet picture, when vui\_modality\_type is equal to 3, as in Table 4.
3. Add a constraint on vui\_colour\_description\_present\_flag:

- When vui\_modality\_type is equal to 2 or 3, vui\_colour\_description\_present\_flag shall be equal to 0.

* Option 1 Aspect:

1. Add **vui\_spectrum\_range** to specify the spectrum band of the optical radiation wavelength represented by the pictures in the CLVS in a predefined spectrum range table. When vui\_spectrum\_range is equal to 1, the user is allowed to signal the min and max wavelength as specified in the proposed item 1.
2. Add constraints on vui\_spectrum\_range:

* When vui\_modality\_type is equal to 1, vui\_spectrum\_range should be 0, 1, or 7.
* When vui\_modality\_type is equal to 2, vui\_spectrum\_range should be 0, 1, or in the range of 8 to 10.
* When vui\_modality\_type is equal to 3, vui\_spectrum\_range should be 0, 1, or in the range of 2 to 6.
* Option 2 Aspect:

1. Add **vui\_spectrum\_range\_present\_flag** to enable signaling of the min and max wavelength directly, without a predefined spectrum range table. When mantissa is equal to 0 or is not present, it denotes that the wavelength is unspecified.

In this contribution, two syntax designs are proposed. It’s proposed to adopt one of the following options:

* Option 1 Aspect (items 5 and 6) + Common Aspects (items 1 to 4).
* Option 2 Aspect (item 7) + Common Aspects (items 1 to 4).

It was commented that it would be better to have a backward compatible approach which could more straightforward be interpreted by legacy devices, e.g. via colour primaries or SEI, rather than modifying VUI syntax substantially.

The proponents argue that they could extend the VUI in a backward compatible way via an SPS extension, but it was commented that this is more difficult to implement in legacy decoders rather than a new colour codepoint of SEI, which are out of the scope of the core decoder.

[JVET-AG0079](https://jvet-experts.org/doc_end_user/current_document.php?id=13635) AHG2/AHG9: VUI extension mechanism and picture modality information for AVC and HEVC [J. Gao, H.-B. Teo, C.-S. Lim, K. Abe, V. Drugeon (Panasonic)]

In the 32nd JVET meeting, signaling the picture modality type in VUI parameters was adopted into the technologies under consideration for future extensions of VSEI (version 2), JVET-AF2032. In AVC and HEVC, the VUI parameters are signaled in the SPS without a specific VUI extension mechanism. This contribution proposes a VUI extension mechanism and picture modality information for AVC and HEVC.

The following items are proposed for the VUI extension mechanism in HEVC:

* Aspect 1:
* Modify **sps\_extension\_4bits** to **sps\_vui\_extension\_flag** and **sps\_extension\_3bits** in theSPS RBSP syntax structure.
* Add sps\_vui\_extension( ) in the SPS RBSP syntax structure, as specified in Annex E.
* Aspect 2:
* Add **vui\_extension\_bits** and **vui\_reserved\_extension** in the sps\_vui\_extension( ) syntax structure to enable future extensions to VUI parameters.

The following items are proposed for the VUI extension mechanism in AVC:

* Aspect 3:
* Add sps\_vui\_extension( ) in the SPS extension RBSP syntax structure, as specified in Annex E.
* Aspect 4:
* Add **additional\_extension2\_flag** and **additional\_extension2\_data\_flag** in the SPS extension RBSP syntax structure to enable future extensions to the SPS extension RBSP syntax structure.
* Aspect 5:
* Add **vui\_extension\_bits** and **vui\_reserved\_extension** in sps\_vui\_extension( ) syntax structure to enable future extensions to VUI parameters.

The picture modality information is added in the sps\_vui\_extension( ) syntax structure for both AVC and HEVC. The following items are proposed for adding picture modality information in AVC and HEVC:

* Aspect 6:
* Signal picture modality information such as picture modality type and spectrum range in the sps\_vui\_extension( ) syntax structure.
* Aspect 7:
* Add **modality\_type\_extension\_bits** and **reserved\_modality\_type\_extension** in the sps\_vui\_extension( ) syntax structure, for future extension of picture modality information.

It is proposed to adopt the proposed VUI extension mechanism and the proposed picture modality information into AVC and HEVC.

Considering the fact that support in all three standards (AVC, HEVC, VVC) is desirable, but the VUI approach requires specific variation for the three of them, it appears that the approach defining it via an SEI message would be more consistent, and also easier to interpret by legacy devices (without change to the core video stream).

With regard of using colour primaries for that purpose, it would still be necessary to define the wavelength.

It was agreed that the proposed syntax of JVET-AG0077 and JVET-AG0079 is appropriate, with a suggested modification of the precision of the exponent for wavelength (6 bit, such that 16 bits in total with mantissa). It should be implemented as a new SEI message rather than VUI.

Proponents were requested to provide a new document with the SEI concept. (see JVET-AG0322)

There was also discussion whether it might be useful to have an additional colour codepoint, and whether the SEI message should remove the type of “visible picture”. This is for future consideration.

[JVET-AG0322](https://jvet-experts.org/doc_end_user/current_document.php?id=13895) AHG9: On Picture Modality Information for SEI [J. Gao, H.-B. Teo, C.-S. Lim, K. Abe, V. Drugeon (Panasonic)] [late]

This contribution is a follow-up to address the revisit after the discussion on JVET-AG0077 and JVET-AG0079. It proposes the following items on the picture modality type which is currently in the technologies under consideration for future extensions of VSEI (version 2). It’s proposed to add picture modality information to VSEI, HEVC, and AVC.

1. Move picture modality information from VUI parameters to a new SEI message.
2. Signal minimum and maximum wavelength which indicates the spectral band of optical radiation represented by the pictures in the CLVS, by adding the following syntax elements:

* Add **mi\_min\_wavelength\_mantissa** and **mi\_max\_wavelength\_mantissa** to specify the mantissa part of the min and max wavelength indicating the spectral band of optical radiation represented by the pictures in the CLVS.
* Add **mi\_min\_wavelength\_exponent\_plus15** and **mi\_max\_wavelength\_exponent\_plus15** to specify the exponent part of the max and min wavelength indicating the spectral band of optical radiation represented by the pictures in the CLVS.
* The unit of wavelength is specified in meters.

1. Add **mi\_modality\_type\_extension\_bits** and **mi\_reserved\_modality\_type\_extension** for future extension of picture modality information.
2. Change the type of picture modalities from thermal infrared picture to ultraviolet picture, when mi\_modality\_type is equal to 3.
3. Add **mi\_spectrum\_range\_present\_flag** to enable signaling of the min and max wavelength directly. When mantissa is equal to 0 or is not present, it denotes that the wavelength is unspecified.
4. Add a constraint on VUI colour description syntax elements:

-When mi\_modality\_type is equal to 2 or 3, decoders shall ignore vui\_colour\_primaries, vui\_transfer\_characteristics, vui\_matrix\_coeffs, and vui\_full\_range\_flag indicated in the VUI parameters

It was suggested to make the syntax fully byte aligned, which could be achieved by using 5 bits for modality type syntax element. This appears sufficient, as only 5 different values are currently used.

It was further commented that an alternative option could be to signal the modality type via VUI (as per JVET-0308, which is for further study). Such a modification could be done later, as the remaining part of the SEI message would be needed anyway.

It was agreed to include the modality information SEI JVET-AG0322 in the preliminary WD for VSEIv4, JVET-AG2034.

It is noted that the contribution also includes corresponding text elements for AVC and HEVC. As the ongoing extensions of those are under DIS/DAM ballot, no decision can be done at the current meeting. It would require ballot comments to include those at the late stage of FDIS.

[JVET-AG0144](https://jvet-experts.org/doc_end_user/current_document.php?id=13700) AHG9: Carriage of depth and alpha maps as HEVC single-layer bitstreams [E. Thomas, E. Potetsianakis, E. Alexiou, R. Ghaznavi-Youvalari, M. Abdoli, M.-L. Champel (Xiaomi)]

For delivering XR experiences on XR head-mounted displays (HMD), it is argued that split rendering services recently emerged in which the XR scenes are pre-rendered by a remote entity and then transmitted as video to the HMD. Specific to AR applications, it is further claimed that those services transmit the pre-rendered views as well as the associated depth maps and alpha maps as multiple video bitstreams in different RTP streams, as defined in the 3GPP SA4 TS 26.565 under development.

However, it is also argued that the current HEVC specification merely enables the carriage of alpha planes and depth pictures as auxiliary pictures in a multi-layer bitstream. In addition, the presence of metadata describing the alpha planes and depth pictures is constrained to layers containing auxiliary pictures and thus cannot be present in a single-layer bitstream.

Therefore, it is proposed to enable split-rendering services by allowing the carriage of depth maps and alpha maps as HEVC single-layer bitstreams.

In order to enable the above use case, this contribution presents possible alternative solutions for each design goal:

1. Information indicating that the bitstream carries a depth map sequence or an alpha map sequence.
   * Option 1.1: Defining a content type parameter to identify depth map and alpha map sequences in a new SEI message.
   * Option 1.2: Defining code points in colour primaries for depth maps and alpha maps as colour primaries in the VUI.
2. Information indicating the mapping used between the depth/alpha map (mono-channel) and the luma and chroma components.
   * Option 2.1: Defining a component mapping parameter in the new SEI message.
   * Option 2.2: Defining a code points in matrix coefficients for mono-channel (luma only) in the VUI.
3. Information specifying how to interpret the sample values from the decoded pictures to reconstruct the depth/alpha maps.
   * Option 3.1: New SEI messages based on the multi-layer extension but defined for the single layer context (no reuse of annexes F and G)
   * Option 3.2: Modifications of existing SEI messages to be used in the single layer context

It was commented that JVET-AG0308 has the same target, but uses a solution via CICP instead.

It was commented by another expert that in principle this would already be possible by sending a depth layer without a base layer using main 10 profile with independent layer decoding flag (and rewriting layer id in NAL header).

It is agreed that this is a desirable functionality (to be supported with AVC, HEVC and VVC), but further study appears necessary to understand what is the best way (e.g., JVET-AG0308 has the same target) of achieving it, and which changes would be necessary for the different specs.

# Plenary meetings, joint meetings, BoG reports, and liaison communications

## JVET plenaries

An intermediate plenary was held on XXday XX Jan. XXXX. The following items were discussed:

* …

Further detail on scheduling is recorded in section 2.15.

Joint meetings involving JVET were held as follows:

* JVET, XXX on XXX, on XXday XX Jan. at XXXX–XXXX
* …

Further detail about these sessions with other groups is provided in the other subsections of this section.

General plenary wrap-up discussions are recorded under sections 8, 9, and 10.

## Information sharing meetings

Information sharing sessions with other WGs and AGs of the MPEG community were held on Monday 22 Jan. 0600–XXXX, Wednesday 24 Jan. 0600–XXXX, and Friday 26 Jan. 2100–XXXX.

The status and plans for the work in the MPEG WGs and AGs was reviewed at these information sharing sessions.

## Joint session 0600-0643 Wednesday 24 Jan on SEI rules and MPI/MIV: WG 2, WG 4, WG 5 / JVET, WG 7, VCEG

(The notes for this session were recorded by JRO.)

The following ideas were presented by J. Ostermann:

Requirements on SEI messages in case they duplicate existing MPEG functionality (ISO/IEC JTC1/SC29/WG11 N12546)

* + - 1. SEI message typically carries data or references to data defined inside or outside of MPEG. It defines interpretation of data carried in the video stream, sometimes a process description but never conformance
      2. Typically, lightweight documentation, limited amount of syntax
      3. Convincing quality demonstrated in showcase
      4. Substantially smaller functionality than existing standard
      5. No extension of functionality of SEI message
      6. SEI messages harmonized across WGs
      7. Specification should be aligned (OMAF functionality is replicated in VVC with aligned specification)

Several experts expressed (at first sight) that he might not generally agree on this. It was pointed out that some existing SEI messages (such as NNPF) would probably not comply with some of those rules. It was suggested to rather discuss based on concrete contributions on a case by case basis.

Regarding MPI, it was proposed that JVET studies the possibility of an SEI message containing a subset of MIV in cooperation with WG 4, or if necessary coordinate on extensions, or alternative solutions. This includes a quality assessment done in a proper way. WG 4 may take it in future amendments.

It was commented that the exact meaning of “cooperation with WG 4” needs clarification.

No agreement was reached on the formulation about “extensions”.

It was commented that some description of principles of SEI messages might also be valuable for external organizations who consider using them or requesting for new ones.

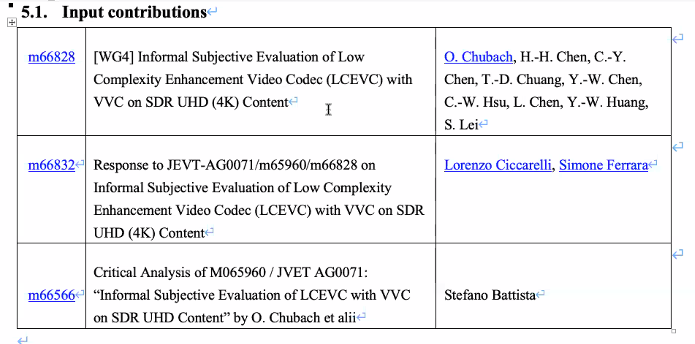
Input on these aspects is solicited.

Session ended 06:43.

## Joint session 0645-07000759 Wednesday 24 Jan on multi-layer coding tests: AG 5, WG 4, WG 5 / JVET, VCEG

(The notes for this session were recorded by JRO.)

Documents submitted to WG 4:



The first two documents are identical to those submitted to JVET (see section 4.5)

It was mentioned that the methodology of the informal test reported in JVET-AG0071 was not identical to the formal test conducted in the verification test. 51 non-expert subjects were participating in the formal test.

The informal test did not replicate the formal test condition, but rather extended conditions for a larger range of rates. For bit rate allocation between base and enhancement, the same formula was used as in the verification test. It was found that over a larger range, losses were found of LCEVC compared to VVC.

It was noted that in a majority of test cases, it was found that the largest amount of rate (99%) was allocated to the base layer (both in the formal and informal test).

It was commented by one expert that for customers interested in using LCEVC it is somewhat confusing that it is necessary to very carefully select the bit rates to get benefit. It was suggested that it would be beneficial if a clarification on that could be conveyed from the side of the standardization bodies.

It was commented that the sequences and rates used in the verification tests had been very carefully selected by the time they had been conducted. For some sequences, the amount of rate in enhancement layer was more significant.

It was commented that some results in JVET-AG0071 show similar problems that were found in the context of the preparation of layered VVC tests, that it is very difficult to demonstrate the benefit of scalable coding over a wide range. It would be important to provide software that somewhat does proper bit rate allocation between layers (probably necessary on a sequence by sequence basis). LTM has not been updated for a long time.

One expert commented that this might affect the credibility of verification tests of MPEG in general. The investigation in JVET-AG0071 might have tried to better match with market needs in terms of rate ranges.

Strong concern is raised about the test setup and analysis and methodology of JVET-AG0071 by the former coordinator of the verification test.

It was commented by LCEVC specialists that there is no confusion in the market. Other investigations exist that confirm the results of the verification tests. An update of LTM would be highly beneficial to help implementers in operating LCEVC in the intended way.

Session ended XX:XX.

(M. Wien taking notes from here to end of session)

Proposal on how to proceed:

- LTM software should be updated as soon as possible such that investigations on performance of LCEVC can be done by experts (inside and outside of MPEG) to assess the performance of LCEVC.

- Set up further investigations (probably in WG 4 together with AG 5) starting by the April meeting. Suggestion to set up an AhG on topic to prepare such activities.

- The method of subjective assessment for this purpose is to be clarified as well.

It was asked if HDR content could be included in the investigation. This should settled in the AhG.

One expert objected to setting up an AhG on this topic.

Closing the session at 07:59h.

## BoGs (4)

The following break-out groups were established at this meeting to conduct discussion and develop recommendations on particular subjects.

[JVET-AG0325](https://jvet-experts.org/doc_end_user/current_document.php?id=13898) BoG on NNVC activity [E. Alshina, F. Galpin]

The report was presented Wednesday Jan. 24, 0740

This report summarizes the discussion during BoG on NNVC activity established during JVET-AG meeting.

First session of BoG was conducted on Jan. 23, 05:00-07:00 (UTC).

Second session of BoG was conducted on Jan. 24, 06:00-07:00 (UTC).

**Introduction**

BoG was established with following mandates:

* Design of next EE1
* Call for training material
* Review section 5.1.5 (SADL implementation)

In total 162 experts attended the 1st and 124 experts attended the 2nd BoG session on-line session.

**Discussion summary**

***Design of next EE1***

Discussion was conducted under assumption NNVC-8 will have the following design:

* NNIntra (unchanged) enabled by default,
* NNLF LOP2 enabled by default,
* NNLF HOP3 (adopted [JVET-AG0174](https://jvet-experts.org/doc_end_user/current_document.php?id=13730)) disabled by default,
* NNLF Post filter disabled by default,
* NNSR disabled by default.
* Minor inference changes for LOP/HOP.

Proposals which have been recommended by JVET to be studied in EE1 are listed below, with clarified aspects to be studied. Tests are grouped into several categories with the hope to have at least some tests to be performed jointly and to have training cross-checked.

Recommendation: For test competing with technologies in NNVC-8 it is recommended to configure the proposed solution to have a complexity close to related NNVC(LOP/HOP), but not exceeding:

kMAC/pxl\_EE1\_test ≤ kMAC/pxl\_NNVC\_ (*must*),

Num Param\_EE1\_test ≤ Num Param\_NNVC\_ (*if possible*).

Exact parameters setting (number of channels, number of repetitions of ‘back bone blocks’) needs to be carefully adjusted to meet the requirements above. EE1 test description should specify network architecture and provides tentative parameters setting, but final setting (number of channels, number of repetitions of ‘back bone blocks’) will be announced in AhG11/14 teleconference. For description of NN architecture, proponent are encouraged to use figures provided in the attached ppt and compute kMAC/pxl using ‘calculator’ in[JVET-AD0380](https://jvet-experts.org/doc_end_user/current_document.php?id=12944)**.**

**Recommendations**

1. For test competing with technologies in NNVC-8 it is recommended to configure the proposed solution to have a complexity close to related NNVC(LOP/HOP), but not exceeding:

kMAC/pxl\_EE1\_test ≤ kMAC/pxl\_NNVC\_ (*must*),

Num Param\_EE1\_test ≤ Num Param\_NNVC\_ (*if possible*).

1. NN architecture to be tested shall be provided by end for JVET AG meeting (to appear in EE1 description draft). Exact parameters settings can be specified by 2nd AhG11/14 teleconference Feb. 28, (Wed).
2. Conduct one joint test ( combination of [JVET-AG0155](https://jvet-experts.org/doc_end_user/current_document.php?id=13711), [JVET-AF0206](https://jvet-experts.org/doc_end_user/current_document.php?id=13467), [JVET-AG0057](https://jvet-experts.org/doc_end_user/current_document.php?id=13613)) to study very low operation point.
3. Test EE1-1.0 to be completed in 2-3 weeks after JVET AG meeting, discussed and announced in 1st AhG11/14 teleconference Feb 13 (Tue).
4. Design four tests EE1-1.0…3 combining [JVET-AG0156](https://jvet-experts.org/doc_end_user/current_document.php?id=13712), [JVET-AG0163](https://jvet-experts.org/doc_end_user/current_document.php?id=13719), [JVET-AG0155](https://jvet-experts.org/doc_end_user/current_document.php?id=13711), [JVET-AG0069](https://jvet-experts.org/doc_end_user/current_document.php?id=13625), changing one aspect at the time.
5. Design three separate tests for 1) separate model [JVET-AG0175](https://jvet-experts.org/doc_end_user/current_document.php?id=13731), 2) wide activation [JVET-AG0179](https://jvet-experts.org/doc_end_user/current_document.php?id=13735), 3) transformers ([JVET-AG0163](https://jvet-experts.org/doc_end_user/current_document.php?id=13719)). If combination test proposed then must be limited to extension separate model to test 2) or 3). For all tests tentative architecture to be provided by the end of JVET AG meeting (to appear in EE1 description draft). Final design for parameters must be announced during AhG11/14 2nd telco (Feb. 28, Wed).
6. In EE1-3.1 test (NNInter) study an effect of training data extension targeting better gain/complexity trade-off.
7. Test existing NNVC SR with default configuration of NNVC-8 (test to be conducted by AhG14).
8. Test adaptive resolution coding with RPR resampling filters with default configuration of NNVC-8.
9. Review proposed text of the call for training materials in JVET, create output document (if agreed).
10. New mandate to promote call for training materials, distribute and negotiate with content owners should be added to AhG11 mandates list.
11. Release up-dated SADL at the end of each meeting (do not change ‘main’ branch in between of meeting).

All recommendations were approved by JVET.

[JVET-AG0330](https://jvet-experts.org/doc_end_user/current_document.php?id=13903) BoG report on FGC subjective testing [M. Wien]

The document JVET-AG0290 was briefly presented by S. Paluri.

It was suggested by the proponent to study different denoisers and their efficiency for denoising the sequences prior to compression. The goal would be to determine a reference denoiser which could be used in the experiments. It was commented that potentially the ‘ground truth’ could be used for the purpose of achieving a denoised version (which would be the best possible one in the given scenario of material being generated by adding film grain to that ground truth sequence. It was argued that this might be more in line with the major focus of this activity of testing film grain synthesis, taking the ground truth as the output of a hypothetical reference denoiser. It was mentioned that this would provide an upper bound of what an encoder could do, and that this would need to be carefully regarded and mentioned in the test report. It was commented that adding the denoising task would give the opportunity to provide additional insight w.r.t the impact of the denoising process.

It was reported that a liaison from 3GPP has been received at the previous meeting which provides information on 3GPP film grain synthesis activities.

One expert reported on impressions from visual inspections of the content provided in JVET-AF0262. The range of film grain intensity was reported to range from very low to very strong film grain. It was suggested to consider the results of JVET-AG0290 as well as visual inspection when selecting candidate variants to be considered in a viewing test. It was reported that in the process of generating the material of JVET-AF0262 the processing tools used for adding the noise potentially induce artifacts such as color shifts in the resulting sequences.

It was mentioned that input contribution JVET-AG0228 suggests a further source for inclusion in the test set under consideration. It was commented that this content might not represent state-of-the-art material as e.g. provided by streaming services. It was argued that there is a lot of legacy content is available and hence, including some of such content is deemed useful.

It was commented that using the ground truth material as the starting point for encoding would be better than the approach of using the compression effect of the encoding algorithm as some type of denoiser.

It was suggested to develop a revised version of the test plan document in the AhG phase and prepare a selection which could be suitable for the test activity.

It was noted that currently, only the ToS sequences of JVET-AF0262 are available on the JVET ftp site. More sequences have been prepared but need to be made available. It was commented that using this subset for the start might be sufficient.

It was stated that as of now it is assumed the tests would be using VVC as the compression scheme.

Tasks for the AhG phase would be (transforming these into appropriate AhG mandates needed):

* JVET-AF0262 sequences
  + The sequence length in the set needs trimming, some have a duration of up to 20 s. A selection of suitable portions of the sequences is to be made
  + A small set of variants of film grain which should be viewed needs to be determined
  + Using the ground truth as virtual output of a hypothetical reference denoiser and run experiments with this
* JVET-AG0228
  + Identify a small set of candidate sequences, clarify licensing situation
* Determine the test protocol(s) to be applied
* Update the test plan document according to these findings
* Prepare input to the April meeting for trial viewings with experts.

[JVET-AG0331](https://jvet-experts.org/doc_end_user/current_document.php?id=13904) BoG on gaming content compression [S. Puri, J. Sauer, R. Chernyak, A. Duenas, L. Wang]

This report summarizes the discussion during the BoG on gaming content compression established during the JVET-AG meeting.

The first BoG session was conducted on Jan. 24, 11:20-12:20 (UTC).

**Introduction**

AHG15 on gaming content compression was established at the 32nd JVET meeting. Mandates of this ad hoc group are:

* Identify gaming content application scenarios and their requirements for codec operation.
* Identify and characterize required types of content; solicit contributions, collect, and make a variety of gaming content available, in coordination with AHG4 and AG 5.
* Propose test conditions appropriate for gaming applications.
* Evaluate JVET test models (such as ECM, VTM, NNVC, etc.) under the proposed test conditions.
* Investigate possibilities to enhance compression capability for gaming content.

**Testing conditions for gaming content**

This section describes simulations’ setup that was used within AhG15 prior to the meeting.

Anchor: VTM-11ecm11 (<https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tree/VTM-11.0ecm11.0>)

Test: ECM-11.0 (<https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tree/ECM-11.0>)

Configurations: AI/LDB/LDP

Testing conditions: JVET CTC with QP set [27, 32, 37, 42]

Content: 5s; TBD

Extra parameters: with and without ClassF cfg file

***Discussion on using classF config:***

* One expert suggested not to use CSGO sequence as it is relatively simple
* ClassF config has only high gains for very simple sequences
* One expert commented that classF runtime increase could be due to individual tools. Suggested to investigate each tool separately
* It was commented that classF seems to work least well for sequences needing high rates
* Surprising that CSGO works well with classF, as the selected segment of the sequence had a lot gameplay and not menus
* Proposed to take best and worst results out of consideration
* It was commented that we tested classF to see if sequences behave more like natural or screen content, not to investigate classF/SCC tools. Suggested not to use classF config
* In LDB results, some encoder runtimes seem decreasing with classF config. This is likely an artifact in result processing due to some missing simulation results
* It was commented that LDB is the most important config for gaming content compression and its desirable to be able to encode fast, i.e. avoid runtime increases
* It was commented that for sequences that one proponent simulated runtime increases were consistent around 20% when enabling classF config

Decision: Do not use classF config for further investigation of gaming content compression. Any tools that may bring gains including classF elements can be suggested to be included in test conditions in the future by bringing test results for said tools.

***Discussion on reporting results template:***

What to include in spreadsheet?

So far used spreadsheet based on JVET-U0127. It targets collection of info on many candidate sequences.

Build on this?

Could add:

* YUV PSNR 10-1-1
* Relative encoder / decoder runtime
* Others?

Switch to sheet based on JVET CTC sheet instead?

Two experts suggested to use CTC test sheets as basis.

Decision: use JVET CTC sheet as basis

***Discussion on tested QPs:***

Question was raised if quality range for current QPs (27-42) is adequate for this content.

One of the outputs of BoG should be mandates for next AhG group. Suggested to have new mandate: Solicit contribution from industry on typical bitrate/quality/resolution used for gaming content compression.

It was commented that Nvidia GeForce NOW uses transmission settings of

|  |  |  |  |
| --- | --- | --- | --- |
| Mbps | resolution | fps | Bit depth |
| 15 | 720p | 60 | 8 |
| 45 | 4K | 120 | 8 |

These are production numbers (tuned for lowest latency). It is expected that reference software default settings would require less rate, but is not suitable for real-time operation.

It was commented that results for some sequences have quite low PSNR numbers (26, 27), but there were no artifacts visible in a simple visual examination.

**Input documents to be discussed**

***JVET-AG0302: AHG15: Description of gaming content sequences proposed by Huawei***

This contribution proposes 5 new sequences to be used for the study of gaming content compression by AHG15. A short description for each sequence is provided in this document. A copyright notice is part of this contribution file and also attached ty o the file source.

The sequences have been uploaded to the jvet ftp server. They are available at:  
<ftp://jvet@ftp.ient.rwth-aachen.de/ahg/candidates/JVET-AG0302>

Particular format for auxiliary information for depth, optical flow and camera parameters was proposed.

Discussion:

It was commented that DesertTown1 seems the most interesting/challenging of these sequences

**Auxiliary information for gaming sequences**

Various additional information may be available from the gaming engine. This can include:

|  |  |  |
| --- | --- | --- |
| **Type** | **Format** | **Used in** |
| depth map | per-pixel, 14 bit | JVET-Y0041 |
| optical flow | quarter-pel | JVET-Y0041 |
| camera information | 4x4 array, 32bit | JVET-AF0187 |
| depth map | per-pixel, 32bit | JVET-AF0187 |
| motion vector | per-pixel, 32bit | JVET-AF0187 |
| camera information | 4x4 array, 32bit | DesertTown/ Sun\_Temple when released |
| depth map | per-pixel, 32bit (R32) | DesertTown/ Sun\_Temple when released |
| motion vector | per-pixel, 2x32bit (R32G32) | DesertTown/ Sun\_Temple when released |
| albedo texture | per-pixel, 4x8bit (R8G8B8A8) | DesertTown/ Sun\_Temple when released |
| normal map | per-pixel, 3x32bit (R32G32B32) | DesertTown/ Sun\_Temple when released |

Discussion:

***Which auxiliary data is desired?***

Of particular interest are:

* Depth
* Motion Vector
* Camera Information

InterDigital offered to convert provided depth map and optical flow to same format as Huawei. It is suggested to also convert Huawei format to InterDigital’s format and provide both at this point. InterDigital can also provide camera information.

Comment: How will 4x4 camera matrices be stored? It is suggested to use a json based format.

Other forms of auxiliary data that could be studied in the future are:

* Albedo
* Specular
* Normal
* Roughness

**Suggested sequence segments for collected sequences**

Of each sequence a segment of 5 seconds was selected for encoding for January meeting. This data should help identifying the most interesting sequences for a new class of gaming content sequences.

The table below lists the selected segments for each sequence. Note that for some sequences more than one potentially interesting segment were identified. The collected simulation results can be found in JVET-AG0015.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sequence name** | **Frame count** | **Frame rate** | **Bit depth** | **Start frame** | **End frame** |
| Level1 | 600 | 60 | 10 | 0 | 299 |
| Darktree | 600 | 60 | 10 | 300 | 599 |
| ArenaOfValor | 600 | 60 | 8 | 0 | 299 |
| ARPG | 600 | 60 | 8 | 100 | 399 |
| DesertTown1 | 600 | 60 | 8 | 200 | 499 |
| DesertTown2 | 600 | 60 | 8 | 150 | 449 |
| Sun\_Temple1 | 600 | 60 | 8 | 40 | 339 |
| Sun\_Temple2 | 600 | 60 | 8 | 50 | 349 |
| CSGO | 3600 | 60 | 8 | 1100 | 1399 |
| DOTA2 | 3600 | 60 | 8 | 1734 | 2033 |
| EuroTruckSimulator2 | 3600 | 60 | 8 | 1000 | 1299 |
| Fallout4 | 3600 | 60 | 8 | 600 | 899 |
| GTAV | 3600 | 60 | 8 | 300 | 599 |
| Hearthstone | 3600 | 60 | 8 | 3099 | 3398 |
| Minecraft | 3600 | 60 | 8 | 600 | 899 |
| Rust | 3600 | 60 | 8 | 3000 | 3299 |
| Starcraft | 3600 | 60 | 8 | 2900 | 3199 |
| Witcher3 | 3600 | 60 | 8 | 1300 | 1599 |
| Baolei-Man | 600 | 60 | 8 | 300 | 599 |
| Baolei-Balloon 4K | 600 | 60 | 8 | 300 | 599 |
| Baolei-Yard 4K | 600 | 60 | 8 | 60 | 359 |
| Baolei-Woman | 600 | 60 | 8 | 120 | 419 |
| Jianling-Temple | 600 | 60 | 8 | 0 | 299 |
| Jianling-Beach | 600 | 60 | 8 | 0 | 299 |
| Heroes of the Storm part 1 | 300 | 30 | 8 | 150 | 299 |
| Project CARS | 300 | 30 | 8 | 0 | 149 |
| WoW part 2 | 300 | 30 | 8 | 150 | 299 |

**Gaming content compression test set and testing conditions**

Comments:

Two experts commented: Reasonable to have two sets. Better to be able to distinguish between sequences that have or do not have auxiliary information.

Suggested to add RA, since its faster to simulate. Can be good if LDB not available in time. Also there are some applications, i.e. streaming gaming content.

Suggested to have AhG telco and choose sequences based on RA results. No need to do crosschecks at this point. If there are concrete proposals (e.g. tools) crosscheck will be needed.

Comment: last time used QP 27 to 42. Two proponents also tested QP22, it is not that slow. Suggested to test QPs 22-42. Only one or two rate points very high.

Comment: QPs 22-42 is also used by AhG11, can re-use excel template of them. Template also reports results for 4 lowest rate points (27-42). Can still use it if QP22 is not finished in time.

Comment: simulations should be split between several parties. As there are many candidates (27?). Not reasonable that everyone simulates all.

Question: Do we need LDP?

Question: do we run 4K sequence for RA. No results for LDB/LDP. Comments: better to run for RA, but not for low delay. Twitch recently also tests 4K streaming of content (RA)

Comments:

* for lowest latency important. In actual implementation LDP is more used than LDB.
* Good to use only one to save some resources
* Agreed to use both of LDP/LDB
* So far no tools exploiting auxiliary information. If tools are developed simulations results need to be shown with and without corresponding tool.

Decision:

* use QPs 22-42 for AI/RA/LDB/LDP (we also add RA).
* Simulation work should be split throughout the group.
* No testing of 4K sequences for low-delay.
* Use excel template by AhG11 as it already supports 5 QPs. It shows results for 5 QPs and 4 QPs.

Proposed testsets contains all candidate sequences, but two classes: with (class G1) and without (class G2) auxiliary information:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Class** | **Sequence name** | **Frame count** | **Frame rate** | **Bit depth** | **Start frame** | **End frame** | **Intra** | **Random access** | **Low-delay** |
| G1 | Level1 | 600 | 60 | 10 | 0 | 299 | M | M | M |
| G1 | Darktree | 600 | 60 | 10 | 300 | 599 | M | M | M |
| G1 | ARPG | 600 | 60 | 8 | 100 | 399 | M | M | M |
| G1 | DesertTown1 | 600 | 60 | 8 | 200 | 499 | M | M | M |
| G1 | DesertTown2 | 600 | 60 | 8 | 150 | 449 | M | M | M |
| G1 | Sun\_Temple1 | 600 | 60 | 8 | 40 | 339 | M | M | M |
| G1 | Sun\_Temple2 | 600 | 60 | 8 | 50 | 349 | M | M | M |
| G2 | ArenaOfValor | 600 | 60 | 8 | 0 | 299 | M | M | M |
| G2 | CSGO | 3600 | 60 | 8 | 1100 | 1399 | M | M | M |
| G2 | DOTA2 | 3600 | 60 | 8 | 1734 | 2033 | M | M | M |
| G2 | EuroTruckSimulator2 | 3600 | 60 | 8 | 1000 | 1299 | M | M | M |
| G2 | Fallout4 | 3600 | 60 | 8 | 600 | 899 | M | M | M |
| G2 | GTAV | 3600 | 60 | 8 | 300 | 599 | M | M | M |
| G2 | Hearthstone | 3600 | 60 | 8 | 3099 | 3398 | M | M | M |
| G2 | Minecraft | 3600 | 60 | 8 | 600 | 899 | M | M | M |
| G2 | Rust | 3600 | 60 | 8 | 3000 | 3299 | M | M | M |
| G2 | Starcraft | 3600 | 60 | 8 | 2900 | 3199 | M | M | M |
| G2 | Witcher3 | 3600 | 60 | 8 | 1300 | 1599 | M | M | M |
| G2 | Baolei-Man | 600 | 60 | 8 | 300 | 599 | M | M | M |
| G2 | Baolei-Balloon 4K | 600 | 60 | 8 | 300 | 599 | M | M |  |
| G2 | Baolei-Yard 4K | 600 | 60 | 8 | 60 | 359 | M | M |  |
| G2 | Baolei-Woman | 600 | 60 | 8 | 120 | 419 | M | M | M |
| G2 | Jianling-Temple | 600 | 60 | 8 | 0 | 299 | M | M | M |
| G2 | Jianling-Beach | 600 | 60 | 8 | 0 | 299 | M | M | M |
| G2 | Heroes of the Storm part 1 | 300 | 30 | 8 | 150 | 299 | M | M | M |
| G2 | Project CARS | 300 | 30 | 8 | 0 | 149 | M | M | M |
| G2 | WoW part 2 | 300 | 30 | 8 | 150 | 299 | M | M | M |

**Licenses**

Need to clarify.

Comments:

* Should ask lawyers to be sure. We are not legal experts
* Some sequence (e.g. 3GPP..) have license information. Could be added to report.
* Suggested to add table similar to the one below to BoG report. License information should be checked by some lawyers.
* Can make two tables: sequences + source. Source + license
* While some sequence have licenses we are not sure we understand the licenses
  + Difficult to do for us (non-legals).
  + make table stating existence of license information and clarity of license
  + licenses should be reviewed by some legal expert

|  |  |  |  |
| --- | --- | --- | --- |
| Sequence | Source | License information available | Clarity of license information |
| Level1 | InterDigital | yes |  |
| Darktree | InterDigital | yes |  |
| ARPG | Huawei | yes | yes |
| DesertTown1 | Huawei | yes | yes |
| DesertTown2 | Huawei | yes | yes |
| Sun\_Temple1 | Huawei | yes | yes |
| Sun\_Temple2 | Huawei | yes | yes |
| ArenaOfValor | Tencent |  |  |
| CSGO | Twitch/SA4 |  | no (see note below) |
| DOTA2 | Twitch |  | no (see note below) |
| EuroTruckSimulator2 | Twitch |  | no (see note below) |
| Fallout4 | Twitch |  | no (see note below) |
| GTAV | Twitch |  | no (see note below) |
| Hearthstone | Twitch |  | no (see note below) |
| Minecraft | Twitch/SA4 |  | no (see note below) |
| Rust | Twitch |  | no (see note below) |
| Starcraft | Twitch/SA4 |  | no (see note below) |
| Witcher3 | Twitch |  | no (see note below) |
| Baolei-Man | Tencent/SA4 | yes | yes |
| Baolei-Balloon 4K | Tencent/SA4 | yes | yes |
| Baolei-Yard 4K | Tencent/SA4 | yes | yes |
| Baolei-Woman | Tencent/SA4 | yes | yes |
| Jianling-Temple | Tencent/SA4 | yes | yes |
| Jianling-Beach | Tencent/SA4 | yes | yes |
| Heroes of the Storm part 1 | Kingston/SA4 | yes |  |
| Project CARS | Kingston/SA4 | yes |  |
| WoW part 2 | Kingston/SA4 | yes |  |

Note: The Twitch sequences from the table above come with the following copyright: “These video game recordings are placed in the public domain. The recordings may contain copyrighted work which can only be used in compliance with fair use”.

**Recommendations**

The BoG recommends to:

* Revise testing condition for gaming content compression
  + Add RA
  + Add QP22
  + Use AhG11 excel template (supports 5 QPs)
  + No longer test classF config. Initial tests indicate that gaming content behaves more like natural than like screen content.
  + Complete crosschecks not necessary at this point, but some partial cross-checks would be desirable to increase robustness of the results . But will be needed if new tools are proposed.
* Define new gaming content test set with two classes
  + Class G1: sequences with auxiliary data
  + Class G2: sequences without auxiliary data
* Work on auxiliary data
  + converting data to several formats such that all formats are available for each sequence that has auxiliary data
  + add camera matrices to auxiliary data
* Plan joint telco with AG5 until April meeting
* Work on clarifying license situation of all sequences
* Add new mandate: Solicit contributions from industry on typical bitrate/quality/resolution used for gaming content compression.

It is planned to further narrow down the number of test sequences. Most sequences are HD, two are 4K. Those would fall into class G2. Clarify if this class could become a mix of resolutions?

In the end, the classes should not have more than 4 cases each. Needs to be clarified later if the classes would be mandatory in CTC (considering the computational burden).

Licensing needs further checking, clarify if sequences can be copied to JVET ftp. Sequences which would be finally selected for the Gx classes shall be made available in the ftp, to guarantee that they can be accessed in longer term.

It was commented that sequences that would not be selected for classes Gx could later be used as training materials.

[JVET-AG0333](https://jvet-experts.org/doc_end_user/current_document.php?id=13906) BoG report on ECM expert viewing preparations [M. Wien]

## Liaison communications (1)

JPEG liaison m66811 provided …

A reply was drafted as WG 5 N XXX. It provided …

The liaison response was reviewed in JVET on Thursday 25 Jan. at 0725-0905. The draft reply was also presented in the MPEG AG 3 Communication meeting Thursday 25 Jan. at 1300-1500.

# Project planning

## Software timeline (update)

ECM 11 software (including all adoptions) was planned to be available 3 weeks after the meeting.

The NNVC 7.0 codebase software was planned to be available 2 weeks after the meeting. An update 7.1 (including potentially the reduced-complexity intra prediction, and an update ofHOP) was planned to be available 5 weeks after the meeting.

VTM23.0 software was planned to be available on 2023-11-24. (Note that further updates may be released later)

Updates on top of HM17.0 software were not planned, but might be released after merging pending requests, as appropriate.

As a general rule in software development, a person who is executing a merge shall not be from the same company as the person who submitted that merge request.

## Core experiment and exploration experiment planning

An EE on neural network-based video coding was established, as recorded in output document JVET-AG2023.

An EE on enhanced compression technology beyond VVC capability using techniques other than neural-network technology was also established, as recorded in output document JVET-AG2024.

Initial versions of these documents were presented and approved.

## Drafting of specification text, encoder algorithm descriptions, and software

The following agreement has been established: the editorial team has the discretion to not integrate recorded adoptions for which the available text is grossly inadequate (and cannot be fixed with a reasonable degree of effort), if such a situation hypothetically arises. In such an event, the text would record the intent expressed by the committee without including a full integration of the available inadequate text.

## Plans for improved efficiency and contribution consideration

The group considered it important to have the full design of proposals documented to enable proper study.

Adoptions need to be based on properly drafted working draft text (on normative elements) and HM/VTM encoder algorithm descriptions – relative to the existing drafts. Proposal contributions should also provide a software implementation (or at least such software should be made available for study and testing by other participants at the meeting, and software must be made available to cross-checkers in EEs).

Suggestions for future meetings included the following generally-supported principles:

* Normative contributions (relating to changes in bitstream/decoder) shall include draft specification text
* Proposals shall contain all details relevant for understanding and be self-contained. In cases where the document is a follow-up of a previous contribution, the overall concept and the novelties should be highlighted at minimum
* Coding tool and encoder optimization proposals shall contain Excel sheets that allow assessment on a per-sequence basis
* Algorithm description text is strongly encouraged for non-normative contributions that are intended to be included in model description documents (VTM, ECM, etc.), and that is required for inclusion in TR drafts.
* Early upload deadline to enable substantial study prior to the meeting
* Using a clock timer to ensure efficient proposal presentations (5 min) and discussions (not exercised currently)

As general guidance, it was suggested to avoid usage of company names in document titles, software modules etc., and not to describe a technology by using a company name.

## General issues for experiments

It was emphasized that those rules which had been set up or refined during the 12th JVET meeting should be observed. In particular, for some CEs of some previous meetings, results were available late, and some changes in the experimental setup had not been sufficiently discussed on the JVET reflector.

Group coordinated experiments have been planned as follows:

* “Core experiments” (CEs) are the coordinated experiments on coding tools which are deemed to be interesting but require more investigation and could potentially become part of a draft standard by the next meeting or in the near future.
* “Exploration experiments” (EEs) are also coordinated experiments. These are conducted on technology which is not foreseen to become part of a draft standard in the near future. The investigating methodology for assessment of such technology can also be an important part of an EE. (Further general rules for EEs, as far as deviating from the CE rules below, should be discussed in a future meeting. For the current meeting, procedures as described in the EE description document are deemed to be sufficient.)
* A CE is a test of a specific fully described technology in a specific agreed way. It is not a forum for thinking of new ideas (like an AHG). The CE coordinators are responsible for making sure that the CE description is complete and correct and has adequate detail. Reflector discussions about CE description clarity and other aspects of CE plans are encouraged.
* A description of each experiment is to be approved at the meeting at which the experiment plan is established. This should include the issues that were raised by other experts when the tool was presented, e.g., interference with other tools, contribution of different elements that are part of a package, etc. The experiment description document should provide the names of individual people, not just company names.
* Software for tools investigated in a CE will be provided in one or more separate branches of the software repository. Each CE will have a “fork” of the software, and within the CE there may be multiple branches established by the CE coordinator. The software coordinator will help coordinate the creation of these forks and branches and their naming. All JVET members will have read access to the CE software branches (using shared read-only credentials as described below).
* During the experiment, revisions of the experiment plans can be made, but not substantial changes to the proposed technology. Withdrawing parts of experiments that were intended to show the individual benefits of a tool or parts of a tool is strongly discouraged. Combination tests may not be considered in such cases. Any changes made to individual tools in a combination shall be documented.
* The CE description must match the CE testing that is done. The CE description needs to be revised if there has been some change of plans.
* The CE summary report must describe any changes that were made in the process of finalizing the CE.
* By the next meeting it is expected that at least one independent cross-checker will report a detailed analysis of each proposed feature that has been tested and confirm that the implementation is correct. Commentary on the potential benefits and disadvantages of the proposed technology in cross-checking reports is highly encouraged. Having multiple cross-checking reports is also highly encouraged (especially if the cross-checking involves more than confirmation of correct test results). The reports of cross-checking activities may (and generally should) be integrated into the CE report rather than submitted as separate documents.
* It is mandatory to report encoder optimizations made for the benefit of a tool, and if an equivalent optimization could be applied on the anchor, a comparison against the improved anchor shall be provided.
* A new proposal can be included in a CE based on group decision, regardless if an independent party has already performed a cross-check in the meeting when it was first proposed.

It is possible to define sub-experiments within particular CEs, for example designated as CEX.a, CEX.b, etc., where X is the basic CE number.

As a general rule, it was agreed that each CE should be run under the same testing conditions using one software codebase, which should be based on the group test model software codebase. An experiment is not to be established as a CE unless there is access given to the participants in (any part of) the CE to the software used to perform the experiments.

The general agreed common conditions for single-layer coding efficiency experiments for SDR video are described in the prior output document JVET-T2010.

Experiment descriptions should be written in a way such that it is understood as a JVET output document (written from an objective “third party perspective”, not a proponent perspective – e.g. not referring to methods as “improved”, “optimized”, “enhanced”, etc.). The experiment descriptions should generally not express opinions or suggest conclusions – rather, they should just describe what technology will be tested, how it will be tested, who will participate, etc. Responsibilities for contributions to CE work should identify individuals in addition to company names.

CE descriptions contain a basic description of the technology under test, but should not contain excessively verbose descriptions of a technology (at least not unless the technology is not adequately documented elsewhere). Instead, the CE descriptions should refer to the relevant proposal contributions for any necessary further detail. However, the complete detail of what technology will be tested must be available – either in the CE description itself or in documents that are referenced in the CE description that are also available in the JVET document archive.

Any technology must have at least one cross-check partner to establish a CE – a single proponent is not enough. It is highly desirable have more than just one proponent and one cross-checker.

The CE development workflow is described at:

<https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/wikis/Core-experiment-development-workflow>

CE read access is available using shared accounts: One account exists for MPEG members, which uses the usual MPEG account data. A second account exists for VCEG members with account information available in the TIES informal ftp area (IFA) system at:

<https://www.itu.int/ifa/t/2017/sg16/exchange/wp3/q06/vceg_account.txt>

Some agreements relating to CE activities were established as follows:

* Only qualified JVET members can participate in a CE.
* Participation in a CE is possible without a commitment of submitting an input document to the next meeting. Participation was requested by contacting the CE coordinator.
* All software, results, and documents produced in the CE should be announced and made available to JVET in a timely manner.
* A JVET CE reflector will be established and announced on the main JVET reflector. Discussion of logistics arrangements, exchange of data, minor refinement of the test plans, and preparation of documents shall be conducted on the JVET CE reflector, with subject lines prefixed by “[CEx: ]”, where “x” is the number of the CE. All substantial communications about a CE other than such details shall take place on main JVET reflector. In the case that large amounts of data are to be distributed, it is recommended to send a link to the data rather than the data itself, or upload the data as an input contribution to the next meeting.

General timeline for CEs

T1= 3 weeks after the JVET meeting: To revise the CE description and refine questions to be answered. Questions should be discussed and agreed on JVET reflector. Any changes of planned tests after this time need to be announced and discussed on the JVET reflector. Initially assigned description numbers shall not be changed later. If a test is skipped, it is to be marked as “withdrawn”.

T2 = Test model software release + 2 weeks: Integration of all tools into a separate CE branch of the VTM is completed and announced to JVET reflector.

* Initial study by cross-checkers can begin.
* Proponents may continue to modify the software in this branch until T3.
* 3rd parties are encouraged to study and make contributions to the next meeting with proposed changes

T3: 3 weeks before the next JVET meeting or T2 + 1 week, whichever is later: Any changes to the CE test branches of the software must be frozen, so the cross-checkers can know exactly what they are cross-checking. A software version tag should be created at this time. The name of the cross-checkers and list of specific tests for each tool under study in the CE plan description shall be documented in an updated CE description by this time.

T4: Regular document deadline minus 1 week: CE contribution documents including specification text and complete test results shall be uploaded to the JVET document repository (particularly for proposals targeting to be promoted to the draft standard at the next meeting).

The CE summary reports shall be available by the regular contribution deadline. This shall include documentation about crosscheck of software, matching of CE description and confirmation of the appropriateness of the text change, as well as sufficient crosscheck results to create evidence about correctness (crosscheckers must send this information to the CE coordinator at least 3 days ahead of the document deadline). Furthermore, any deviations from the timelines above shall be documented. The numbers used in the summary report shall not be changed relative to the description document.

CE reports may contain additional information about tests of straightforward combinations of the identified technologies. Such supplemental testing needs to be clearly identified in the report if it was not part of the CE plan.

New branches may be created which combine two or more tools included in the CE document or the VTM (as applicable).

It is not necessary to formally name cross-checkers in the initial version of the CE description document. To adopt a proposed feature at the next meeting, JVET would like to see comprehensive cross-checking done, with analysis of whether the description matches the software, and a recommendation of the value of the tool and given tradeoffs.

The establishment of a CE does not indicate that a proposed technology is mature for adoption or that the testing conducted in the CE is fully adequate for assessing the merits of the technology, and a favourable outcome of CE does not indicate a need for adoption of the technology into a standard or test model.

Availability of specification text is important to have a detailed understanding of the technology and also to judge what its impact on the complexity of the specification will be. There must also be sufficient time to study this in detail. CE contributions without sufficiently mature draft specification text in the CE input document should not be considered for adoption.

Lists of participants in CE documents should be pruned to include only the active participants. Read access to software will be available to all members.

# Establishment of ad hoc groups

The ad hoc groups established to progress work on particular subject areas until the next meeting are described in the table below. The discussion list for all of these ad hoc groups was agreed to be the main JVET reflector ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de)).

Review of AHG plans was conducted during the plenary on XXday XX Jan. 2024 at XXXX.

|  |  |  |
| --- | --- | --- |
| **Title and Email Reflector** | **Chairs** | **Mtg** |
| **Project Management (AHG1)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Coordinate overall JVET interim efforts. * Supervise AHG and experiment studies. * Report on project status to JVET reflector. * Provide a report to the next meeting on project coordination status. * Supervise processing and delivery of output documents | J.-R. Ohm (chair), G. J. Sullivan (vice‑chair) | N |
| **Draft text and test model algorithm description editing (AHG2)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Produce and finalize draft text outputs of the meeting (JVET-AF1006, JVET-AF1016, and JVET-AF2027. * Collect reports of errata for the VVC, VSEI, HEVC, AVC, CICP, and the published related technical reports and produce the JVET-AF1004 errata output collection. * Coordinate with the test model software development AhG to address issues relating to mismatches between software and text. * Collect and consider errata reports on the texts. | B. Bross, C. Rosewarne (co-chairs), F. Bossen, A. Browne, S. Kim, S. Liu, J.‑R. Ohm, G. J. Sullivan, A. Tourapis, Y.-K. Wang, Y. Ye (vice‑chairs) | N |
| **Test model software development (AHG3)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Coordinate development of test models (VTM, HM, SCM, SHM, HTM, MFC, MFCD, JM, JSVM, JMVM, 3DV-ATM, 360Lib, and HDRTools) software and associated configuration files. * Produce documentation of software usage for distribution with the software. * Enable software support for recently standardized additional SEI messages, and SEI messages in TuC (the latter in a separate branch of VTM). * Discuss and make recommendations on the software development process. * Perform comparative tests of test model behaviour using common test conditions, including HDR, high bit depth and high bit rate. * Suggest configuration files for additional testing of tools. * Investigate how to minimize the number of separate codebases maintained for group reference software. * Coordinate with AHG on Draft text and test model algorithm description editing (AHG2) to identify any mismatches between software and text, and make further updates and cleanups to the software as appropriate. * Prepare drafts of merged and updated CTC documents for HM and VTM, as applicable. | F. Bossen, X. Li, K. Sühring (co-chairs), E. François, Y. He, K. Sharman, V. Seregin, A. Tourapis (vice‑chairs) | N |
| **Test material and visual assessment (AHG4)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Consider plans for additional verification testing of VVC capability, particularly target conducting a first test for VVC multi-layer features by the next meeting, and update the test plan according to subsequent tests. * Coordinate with AHG13 on assessing new test material and investigating metrics that could be used to assess quality of synthesized film grain; improve and update the draft test plan for subjective quality testing of the FGC SEI message. * Maintain the video sequence test material database for testing the VVC and HEVC standards and potential future extensions, as well as exploration activities. * Study coding performance and characteristics of available and proposed video test material. * Identify and recommend appropriate test material for testing the VVC standard and potential future extensions, as well as exploration activities. * Identify and characterize missing types of video material, solicit contributions, collect, and make available a variety of video sequence test material, in coordination with other AHGs, as appropriate. * Maintain and update the directory structure for the test sequence repository, as necessary. * Collect information about test sequences that have been made available by other organizations. * Prepare and conduct expert viewing for purposes of subjective quality evaluation. * Coordinate with AG 5 in studying and developing further methods of subjective quality evaluation, e.g. based on crowd sourcing. * Coordinate with AHG15 on investigating sequences with gaming content, and make such sequences available for study. * Prepare availability of viewing equipment and facilities arrangements for future meetings. | V. Baroncini, T. Suzuki, M. Wien (co-chairs), W. Husak, S. Iwamura, P. de Lagrange, S. Liu, X. Meng, S. Puri, A. Segall, S. Wenger (vice-chairs) | Y (tel., 2 weeks notice) |
| **Conformance testing (AHG5)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the draft of additional conformance bitstreams for VVC multilayer configurations JVET-AE2028, and investigate the need for future improvements of conformance testing specifications. * Study the conformance needs for HEVC multi-view profiles, and develop a set of conformance bitstreams as appropriate. * Study the requirements of VVC, HEVC, and AVC conformance testing to ensure interoperability. * Maintain and update the conformance bitstream database, and contribute to report problems, and suggest actions to resolve these. * Study additional testing methodologies to fulfil the needs for VVC conformance testing. | I. Moccagatta (chair), F. Bossen, K. Kawamura, P. de Lagrange, T. Ikai, S. Iwamura, H.-J. Jhu, S. Paluri, K. Sühring, Y. Yu (vice‑chairs) | N |
| **ECM software development (AHG6)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Coordinate development of the ECM software and associated configuration files. * Produce documentation of software usage for distribution with the software. * Prepare and deliver ECM-11.0 software version and the reference configuration encodings according to the ECM common test conditions. * Investigate encoder speedup and other software optimization such as reduction of memory consumption. * Coordinate with ECM algorithm description editors to identify any mismatches between software and text, make further updates and cleanups to the software as appropriate. | V. Seregin (chair), J. Chen, R. Chernyak, F. Le Léannec, K. Zhang (vice-chairs) | N |
| **ECM tool assessment (AHG7)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Investigate methodology of tool assessment. * Coordinate with AHG6 on resolving tool-off test related software issues (missing tool controls and software bugs). * Prepare configuration files and generate bitstreams and results of tool-on/tool-off testing. * Prepare reporting of tool assessment results. * Collect simulation results on non-CTC sequences (e.g., those used in previous verification tests), and identify a set of non-CTC sequences that would be appropriate for additional testing. * Investigate the possibility of conducting subjective tests on subsets of tools in coordination with AHG4 and AG 5. * Develop methodology of more reliable runtime measurement | X. Li (chair), L.-F. Chen, Z. Deng, J. Gan, E. François, H.-J. Jhu, X. Li, H. Wang (vice‑chairs) | N |
| **Optimization of encoders and receiving systems for machine analysis of coded video content (AHG8)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Solicit and study non-normative encoder and receiving systems technologies that enhance performance of machine analysis tasks on coded video content. * Identify and collect test materials that are suitable to be used by JVET for machine analysis tasks. * Generate anchors according to the common test conditions JVET-AF2031. * Discuss improvements on the evaluation framework, including evaluation procedures and methodologies. * Coordinate software development, and continue to migrate the software basis used in AHG8 to newest VTM version. * Coordinate experiments on optimization of encoders and receiving systems for machine analysis of coded video content. * Maintain the software implementation example algorithms in the repository, including sufficient documentation in terms of operation and performance. * Evaluate proposed technologies and their suitability for machine analysis applications. * Propose improvements to the draft technical report JVET-AE2030 on optimization of encoders and receiving systems for machine analysis of coded video content. * Study the potentials of using SEI messages for the purpose of machine analysis in coordination with AHG9. * Investigate variety of models for machine task … * Coordinate with WG 4 VCM AHG on aspects such as common test conditions, evaluation metrics, test and training materials, usage of SEI messages, and on studying characteristics and requirements of targeted machine analysis tasks, etc. | C. Hollmann, S. Liu, S. Wang, M. Zhou (AHG chairs) | N |
| **SEI message studies (AHG9)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the SEI messages in VSEI, VVC, HEVC and AVC. * Discuss the document for technologies under consideration for VSEI JVET-AF2032, and propose improvements as appropriate. * Collect software and showcase information for SEI messages, including encoder and decoder implementations and bitstreams for demonstration and testing. * Mandate about preparing a test of visual quality of MPI in the next meeting, and coordinate with WG 4 how that could be compared to MIV. * Identify potential needs for additional SEI messages, including the study of SEI messages defined in HEVC and AVC for potential use in the VVC context. * Study the alignments of the same SEI messages in different standards * Coordinate with AHG8 and WG 4 to study mechanisms for signalling metadata in the context of machine analysis of coded video content. * Coordinate with AHG3 for software support of SEI messages. | S. McCarthy, Y.-K. Wang (co-chairs), T. Chujoh, S. Deshpande, C. Fogg, M. M. Hannuksela, Hendry, P. de Lagrange, G. J. Sullivan, A. Tourapis, S. Wenger (vice-chairs) | N |
| **Encoding algorithm optimization (AHG10)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the impact of using techniques such as tool adaptation and configuration, and perceptually optimized adaptive quantization for encoder optimization. * Study the impact of non-normative techniques of preprocessing for the benefit of encoder optimization. * Study encoding techniques of optimization for objective quality metrics and their relationship to subjective quality. * Study optimized encoding for reference picture resampling and scalability modes in VTM. * Study optimized encoding and tool combinations for low latency and low complexity. * Consider neural network-based encoding optimization technologies for video coding standards. * Investigate other methods of improving objective and/or subjective quality, including adaptive coding structures and multi-pass encoding. * Study methods of rate control and rate-distortion optimization and their impact on performance, subjective and objective quality. * Study the potential of defining default or alternate software configuration settings and test conditions optimized for either subjective quality, or higher objective quality, and coordinate such efforts with AHG3 and AHG6. * Study the effect of varying configuration parameters depending on temporal layer, such as those related to deblocking, partitioning, chroma QP. | P. de Lagrange, A. Duenas, R. Sjöberg, A. Tourapis (AHG chairs) | N |
| **Neural network-based video coding (AHG11)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Evaluate and quantify the performance improvement potential of NN-based video coding technologies compared to existing video coding standards such as VVC, including both individual coding tools, architectures and content adaptation with NN parameters overfitting. * Discuss potential refinements of the test conditions for NN-based video coding in JVET-AF2016. Generate and distribute anchor encoding, and develop supporting software as needed. * Study the impact of training (including the impact of loss functions) on the performance of candidate technologies and identify suitable material for testing and training. * Analyse complexity characteristics for technologies under study, including transformers, perform complexity analysis, and develop complexity reductions of candidate technology. * Finalize and discuss the EE on neural network-based video coding. * New mandate to promote call for training materials, distribute and negotiate with content owners * Coordinate with other groups, including SC29/AG5 on the evaluation and assessment of visual quality, and AHG12 on the interaction with ECM coding tools. If possible, prepare encodings with combinations of tools included in the NNVC software for visual quality assessment at the next meeting. * Coordinate with AHG14 on items related to NNVC software development. | E. Alshina, F. Galpin, S. Liu, A. Segall (co-chairs), J. Li, R.-L. Liao, D. Rusanovskyy, M. Santamaria, T. Shao, M. Wien, P. Wu (vice chairs) | Y (tel., 2 weeks notice), first on Feb. 13 |
| **Enhanced compression beyond VVC capability (AHG12)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Solicit and study non-neural-network video coding tools with enhanced compression capabilities beyond VVC. * Discuss and propose refinements to the ECM11 algorithm description JVET-AF2025. * Coordinate with AHG7 to study the performance and complexity tradeoff of these video coding tools. * Coordinate with AHG6 on ECM software development. * Support AHG6 in generating anchors according to the test conditions in JVET-AF2017. * Analyse the results of exploration experiments described in JVET-AF2024 in coordination with the EE coordinators. * Coordinate with AHG11 to study the interaction with neural network-based coding tools. | M. Karczewicz, Y. Ye, L. Zhang (co-chairs), B. Bross, R. Chernyak, X. Li, K. Naser, Y. Yu (vice-chairs) | N |
| **Film grain technologies (AHG13)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the benefits and characteristics of film grain technologies, including autoregressive and frequency-filtering technologies. * Study alternative film grain models and their associated documentation. * In consultation with AHG4, study and define content characteristics and test conditions that are desirable for the study and testing of film grain technologies, and perform an assessment of newly available test materials in that regard. * Investigate metrics for measuring film grain fidelity in itself, or as present in a video. * Discuss the potential need for film grain conformance guidelines. * Given the study of desirable content characteristics, solicit or create new test material for further determining the operational characteristics of, testing, and developing any related technologies. * Study preprocessing and encoder technologies for determining values for FGC (Film Grain Characteristics) SEI message syntax elements. * Identify potential need for additional film grain technology and signalling, if needed. * Coordinate development of film grain technology software and configuration files. * Coordinate with AG 5 on improving the draft plan for subjective quality testing of the FGC SEI message JVET-AD2022, and conduct preparations for such testing. * Coordinate with AHG3 for software support of the FGC SEI message. | W. Husak, P. de Lagrange (co-chairs), A. Duenas, D. Grois, Y. He, X. Meng, M. Radosavljević, A. Segall, G. Teniou, A. Tourapis (vice-chairs) | Y (tel., 2 weeks notice) |
| **NNVC software development (AHG14)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Coordinate development of the NNVC software and associated configuration files. * Prepare and deliver NNVC-7.0 and NNVC-7.1 software versions and the reference configuration encodings according to the NNVC common test conditions as described in JVET-AF2016. * Investigate combinations of tools included in the NNVC software, prepare and release anchor data for all configurations of the software, including anchors for High and Low Operation Point (HOP/LOP) configurations. * Study and maintain the SADL (Small Adhoc Deep-Learning Library). Identify gaps in functionality and develop improvements as needed. * Coordinate with NNVC algorithm and software description (JVET-AF2019) editors to identify any mismatches between software and description document, suggest further updates to the description document as appropriate. * Coordinate with AHG11 on items related to NNVC activities. | F. Galpin (chair), Y. Li, Y. Li, J. Shingala, L. Wang, Z. Xie (vice chairs) | Y (tel., 2 weeks notice), first on Nov. 16 |
| **Gaming content compression (AHG15)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Identify gaming content application scenarios and their requirements for codec operation. * Identify and characterize required types of content; solicit contributions, collect, and make a variety of gaming content available, in coordination with AHG4 and AG 5. * Propose test conditions appropriate for gaming applications. * Evaluate JVET test models (such as ECM, VTM, NNVC, etc.) under the proposed test conditions. * Investigate possibilities to enhance compression capability for gaming content. | S. Puri, J. Sauer (co-chairs), R. Chernyak, A. Duenas, L. Wang (vice chairs) | Y (tel., 2 weeks notice) |
| **Generative face video compression (AHG16)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Establish testing conditions for evaluating the compression performance of generative face video compression (GFVC). * Identify and develop software tools for experimentation on GFVC, and make a software package available. * Study interoperability requirements, including study of the compression performance impact due to GFVC parameter translation. * Study compression performance using the VVC Main 10 profile. * Develop a document summarizing GFVC technologies. | Y. Ye (chair), H.-B. Teo, Z. Lyu, S. McCarthy, S. Wang (vice chairs) | Y (tel., 2 weeks notice) |

It was confirmed that the rules which can be found in document ISO/IEC JTC 1/‌SC 29/‌AG 2 [N 046](https://www.mpegstandards.org/wp-content/uploads/2022/01/ISO-IECJTC1-SC29-AG2_N0046_AhG.pdf) “Ad hoc group rules for MPEG AGs and WGs” (available at <https://www.mpegstandards.org/adhoc/>), are consistent with the operation mode of JVET AHGs. It is pointed out that JVET does not maintain separate AHG reflectors, such that any JVET member is implicitly a member of any AHG. This shall be mentioned in the related WG Recommendations. The list above was also issued as a separate WG 5 document (ISO/IEC JTC 1/‌SC 29/‌WG 5 N 270) in order to make it easy to reference.

# Output documents

The following documents were agreed to be produced or endorsed as outputs of the meeting. Names recorded below indicate the editors responsible for the document production. Where applicable, dates of planned finalization and corresponding parent-body document numbers are also noted.

It was reminded that in cases where the JVET document is also made available as a WG 5 output document, a separate version under the WG 5 document header should be generated. This version should be sent to GJS and JRO for upload.

The list of JVET ad hoc groups was also issued as a WG 5 output document WG 5 N 270, as noted in section 9.

[JVET-AG1000](https://jvet-experts.org/doc_end_user/current_document.php?id=13582) Meeting Report of the 33rd JVET Meeting [J.-R. Ohm] [WG 5 N 257] (2024-02-23)

Initial versions of the meeting notes (d0 … d8) were made available on a daily basis during the meeting.

Remains valid – not updated: [JVET-AC1001](https://jvet-experts.org/doc_end_user/current_document.php?id=12566) Guidelines for HM-based software development [K. Sühring, F. Bossen, X. Li (software coordinators)]

Remains valid – not updated: [JVET-Y1002](https://jvet-experts.org/doc_end_user/current_document.php?id=11463) High Efficiency Video Coding (HEVC) Test Model 16 (HM 16) Encoder Description Update 16 [C. Rosewarne (primary editor), K. Sharman, R. Sjöberg, G. J. Sullivan (co-editors)] [WG 5 [N 103](https://dms.mpeg.expert/doc_end_user/current_document.php?id=82085&id_meeting=189)]

Remains valid – not updated: [JVET-AD1003](https://jvet-experts.org/doc_end_user/current_document.php?id=12970) Coding-independent code points for video signal type identification (Draft 2 of 3rd edition) [WG 5 preliminary FDIS N 206] [G. J. Sullivan, A. Tourapis]

The technical content was submitted for ITU consent (but will not be published until ST 2128 is available); ISO FDIS was to be delayed until it is available.

Post-meeting note: Expected *de facto* primary editor for ITU consent text: G. J. Sullivan.

[JVET-AG1004](https://jvet-experts.org/doc_end_user/current_document.php?id=13583) Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP [Y.-K. Wang, B. Bross, I. Moccagatta, C. Rosewarne, G. J. Sullivan] (2024-03-29, near next meeting)

Primary editor: Y.-K. Wang.

1. …

These items should also be included into the AVC DIS and the HEVC DAM1 texts and their corresponding JVET output documents, when applicable.

No output: JVET-Axx1005

Remains valid – not updated: [JVET-AF1006](https://jvet-experts.org/doc_end_user/current_document.php?id=13584) New profiles, colour descriptors, and SEI messages for HEVC (draft 2) [WG 5 DAM N 244] [Y.-K. Wang, B. Bross, T. Ikai, G. J. Sullivan, A. Tourapis]

Primary editor for this document and WG 5 N 244: Y.-K. Wang.

A DoCR on ISO/IEC 23008-2/CDAM1 was issued as WG 5 N 243.

Remains valid – not updated: [JCTVC-V1007](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=10312) SHVC Test Model 11 (SHM 11) Introduction and Encoder Description [G. Barroux, J. Boyce, J. Chen, M. M. Hannuksela, Y. Ye] [WG 11 N 15778]

No output: JVET-Axx1008

Remains valid – not updated: [JCTVC-AC1009](https://jvet-experts.org/doc_end_user/current_document.php?id=12569) Common test conditions for SHVC [K. Sühring]

Remains valid – not updated [JCTVC-O1010](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=8511) Guidelines for Conformance Testing Bitstream Preparation [T. Suzuki, W. Wan]

No output: JVET-Axx1011

Remains valid – not updated: JVET-[AD1012](https://jvet-experts.org/doc_end_user/current_document.php?id=12973) Overview of IT systems used in JVET [J.-R. Ohm, I. Moccagatta, K. Sühring, M. Wien]

Remains valid – not updated: [JCT3V-G1003](http://phenix.int-evry.fr/jct3v/doc_end_user/current_document.php?id=1884) 3D-AVC Test Model 9 [ D. Rusanovskyy, F. C. Chen, L. Zhang, T. Suzuki] [WG 11 N 14239]

Remains valid – not updated: [JCT3V-K1003](http://phenix.int-evry.fr/jct3v/doc_end_user/current_document.php?id=2499) Test Model 11 of 3D-HEVC and MV-HEVC [Y. Chen, G. Tech, K. Wegner, S. Yea] [WG 11 N 15141]

Remains valid – not updated: [JVET-AE1013](https://jvet-experts.org/doc_end_user/current_document.php?id=13268) Common test conditions of 3DV experiments [K. Sühring, M. Wien]

Remains valid – not updated [JCTVC-V1014](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=10316) Screen Content Coding Test Model 7 Encoder Description (SCM 7) [R. Joshi, J. Xu, R. Cohen, S. Liu, Y. Ye] [WG 11 N 16049]

Remains valid – not updated: [JVET-AC1015](https://jvet-experts.org/doc_end_user/current_document.php?id=12571) Common test conditions for SCM-based screen content coding [K. Sühring]

Remains valid – not updated: [JVET-AF1016](https://jvet-experts.org/doc_end_user/current_document.php?id=13585) AVC with extensions and corrections (draft 2) [WG5 DIS of 11th ed. N 241] [B. Bross, T. Ikai, G. J. Sullivan, A. Tourapis, Y.-K. Wang]

Primary editor of this document and WG 5 N 241: B. Bross.

A DoCR on ISO/IEC CD 14496-10 was issued as WG 5 N 240.

No output: JVET-Axx1017 through JVET-Axx1099

Remains valid – not updated [JVET-AA1100](https://jvet-experts.org/doc_end_user/current_document.php?id=11944) Common Test Conditions for HM Video Coding Experiments [K. Sühring, K. Sharman]

This specifies only the CTC for non-4:2:0 colour formats. The corresponding document for VVC is JVET-T2013, with no unification yet.

**No output: JVET-Axx2001**

Remains valid – not updated: [JVET-AF2002](https://jvet-experts.org/doc_end_user/current_document.php?id=13586) Algorithm description for Versatile Video Coding and Test Model 21 (VTM 21) [A. Browne, Y. Ye, S. Kim] [WG 5 N 245]?

New elements from notes elsewhere in this report:

* …

It is noted that the list above may not be complete; if some adoption is missing that is recorded somewhere else in the meeting notes it shall also be considered included.

It was suggested that editorial improvements submitted as input to the next meeting would be welcome.

Remains valid – not updated: [JVET-AC2003](https://jvet-experts.org/doc_end_user/current_document.php?id=12573) Guidelines for VTM-based software development [F. Bossen, X. Li, K. Sühring]

Remains valid – not updated: [JVET-T2004](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10542) Algorithm descriptions of projection format conversion and video quality metrics in 360Lib (Version 12) [Y. Ye, J. Boyce]

Remains valid – not updated: [JVET-AE2005](https://jvet-experts.org/doc_end_user/current_document.php?id=13270) New level and systems-related supplemental enhancement information for VVC (Draft 6) [B. Bross, E. François, M. M. Hannuksela, A. Tourapis, Y.-K. Wang]

Primary editor for this document, the corresponding ITU consent text, and the corresponding FDIS text WG 5 N 228: B. Bross.

Remains valid – not updated: [JVET-AE2006](https://jvet-experts.org/doc_end_user/current_document.php?id=13271) Additional SEI messages for VSEI (Draft 5) [S. McCarthy, T. Chujoh, M. M. Hannuksela, G. J. Sullivan, Y.-K. Wang]

Primary editor for this document, the corresponding ITU consent text, and the corresponding FDIS text WG 5 N 220: Y.-K. Wang.

Remains valid – not updated: [JVET-AD2007](https://jvet-experts.org/doc_end_user/current_document.php?id=12977) Guidelines for NNVC software development [F. Galpin, S. Eadie, L. Wang, Z. Xie, Y. Li]

Remains valid – not updated: [JVET-X2008](https://jvet-experts.org/doc_end_user/current_document.php?id=11228) Conformance testing for versatile video coding (Draft 7) [J. Boyce, F. Bossen, K. Kawamura, I. Moccagatta, W. Wan]

Remains valid – not updated: [JVET-Y2009](https://jvet-experts.org/doc_end_user/current_document.php?id=11470) Reference software for versatile video coding (Draft 3) [F. Bossen, K. Sühring, X. Li]

Remains valid – not updated [JVET-AB2010](https://jvet-experts.org/doc_end_user/current_document.php?id=12216) VTM and HM common test conditions and software reference configurations for SDR 4:2:0 10 bit video [F. Bossen, X. Li, V. Seregin, K. Sharman, K. Sühring]

Remains valid – not updated: [JVET-AC2011](https://jvet-experts.org/doc_end_user/current_document.php?id=12575) VTM and HM common test conditions and evaluation procedures for HDR/WCG video [A. Segall, E. François, W. Husak, S. Iwamura, D. Rusanovskyy]

Remains valid – not updated: [JVET-U2012](https://jvet-experts.org/doc_end_user/current_document.php?id=10681) JVET common test conditions and evaluation procedures for 360° video [Y. He, J. Boyce, K. Choi, J.-L. Lin]

Remains valid – not updated: [JVET-T2013](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10546) VTM common test conditions and software reference configurations for non-4:2:0 colour formats [Y.-H. Chao, Y.-C. Sun, J. Xu, X. Xu]

Remains valid – not updated: [JVET-Q2014](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9683) JVET common test conditions and software reference configurations for lossless, near lossless, and mixed lossy/lossless coding [T.-C. Ma, A. Nalci, T. Nguyen]

Remains valid – not updated: [JVET-Q2015](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9684) JVET functionality confirmation test conditions for reference picture resampling [J. Luo, V. Seregin]

Remains valid – not updated: [JVET-AF2016](https://jvet-experts.org/doc_end_user/current_document.php?id=13587) Common test conditions and evaluation procedures for neural network-based video coding technology [E. Alshina, R.-L. Liao, S. Liu, A. Segall]

Remains valid – not updated: [JVET-AF2017](https://jvet-experts.org/doc_end_user/current_document.php?id=13588) Common test conditions and evaluation procedures for enhanced compression tool testing [M. Karczewicz, Y. Ye]

Remains valid – not updated: [JVET-AA2018](https://jvet-experts.org/doc_end_user/current_document.php?id=11949) Common test conditions for high bit depth and high bit rate video coding [A. Browne, T. Ikai, D. Rusanovskyy, X. Xiu, Y. Yu]

[JVET-AG2019](https://jvet-experts.org/doc_end_user/current_document.php?id=13580) Description of algorithms and software in neural network-based video coding (NNVC) version 6 [F. Galpin, Y. Li, D. Rusanovskyy, J. Ström, L. Wang] [WG 5 N 265] (2024-03-15)

New elements from notes elsewhere in this report:

* ….
* Adoption to NNVC-8.0 software
  + …

It is noted that the list above may not be complete; if some adoption is missing that is recorded somewhere else in the meeting notes it shall also be considered included.

[JVET-AG2020](https://jvet-experts.org/doc_end_user/current_document.php?id=13275) Film grain synthesis technology for video applications (Draft 6) [D. Grois, Y. He, W. Husak, P. de Lagrange, A. Norkin, M. Radosavljević, A. Tourapis, W. Wan] [WG 5 DTR N 260] (2024-02-16)

Remains valid – not updated: [JVET-AF2021](https://jvet-experts.org/doc_end_user/current_document.php?id=13589) Verification test plan for VVC multilayer coding (update 2) [S. Iwamura, P. de Lagrange, M. Wien]

[JVET-AG2022](https://jvet-experts.org/doc_end_user/current_document.php?id=12982) Draft plan for subjective quality testing of FGC SEI message [P. de Lagrange, W. Husak, M. Radosavljević, M. Wien] (2024-XX-XX)

See BoG JVET-AG0330 for details.

[JVET-AG2023](https://jvet-experts.org/doc_end_user/current_document.php?id=13578) Exploration experiment on neural network-based video coding (EE1) [E. Alshina, F. Galpin, Y. Li, D. Rusanovskyy, M. Santamaria, J. Ström, R. Chang, Z. Xie] [WG 5 N 264] (2024-02-09)

An initial draft of this document was reviewed and approved at XXXX-XXXX on Friday 26 Jan.

This round of EE1 tests will include:

…

[JVET-AG2024](https://jvet-experts.org/doc_end_user/current_document.php?id=13579) Exploration experiment on enhanced compression beyond VVC capability (EE2) [V. Seregin, J. Chen, R. Chernyak, K. Naser, J. Ström, F. Wang, M. Winken, X. Xiu, K. Zhang] [WG 5 N 267] (2024-02-16)

An initial draft of this document was reviewed and approved at XXXX-XXXX on Friday 26 Jan.

Categories and experiments are listed in the subsequent table:

…

[JVET-AG2025](https://jvet-experts.org/doc_end_user/current_document.php?id=13581) Algorithm description of Enhanced Compression Model 12 (ECM 12) [M. Coban, R.-L. Liao, K. Naser, J. Ström, L. Zhang] [WG 5 N 268] (2024-03-15)

New elements from notes elsewhere in this report:

* …

It is noted that the list above may not be complete; if some adoption is missing that is recorded somewhere else in the meeting notes it shall also be considered included.

Remains valid – not updated: [JVET-AC2026](https://jvet-experts.org/doc_end_user/current_document.php?id=12581) Conformance testing for VVC operation range extensions (Draft 4) [D. Rusanovskyy, T. Hashimoto, H.-J. Jhu, I. Moccagatta, Y. Yu] (2023-04-14)

Can be removed after publication of VVC conf. v2.

[JVET-AG2027](https://jvet-experts.org/doc_end_user/current_document.php?id=13590) SEI processing order and processing order nesting SEI messages in VVC (draft 7) [G. J. Sullivan, M. M. Hannuksela, Y.-K. Wang] [WG 5 preliminary WD 5 N 261] (2024-03-15)

Primary editor: G. J. Sullivan

Updated from …

Remains valid – not updated: [JVET-AE2028](https://jvet-experts.org/doc_end_user/current_document.php?id=13279) Additional conformance bitstreams for VVC multilayer configurations [S. Iwamura, P. de Lagrange, I. Moccagatta]

Remains valid – not updated: [JVET-AB2029](https://jvet-experts.org/doc_end_user/current_document.php?id=12225) Visual quality comparison of ECM/VTM encoding [V. Baroncini, J.-R. Ohm, M. Wien] [AG 5 N 75]

[JVET-AG2030](https://jvet-experts.org/doc_end_user/current_document.php?id=13280) Optimization of encoders and receiving systems for machine analysis of coded video content (draft 4) [J. Chen, C. Hollmann, S. Liu] [WG 5 WD N 263)] (2024-03-15)

Primary editor for this document: XXX.

A request for subdivision (part 3 of MPEG-AI, ISO/IEC 23888-3) was issued as WG 5 document N 262.

Remains valid – not updated: [JVET-AF2031](https://jvet-experts.org/doc_end_user/current_document.php?id=13591) Common test conditions for optimization of encoders and receiving systems for machine analysis of coded video content [S. Liu, C. Hollmann] (2023-11-17)

[JVET-AG2032](https://jvet-experts.org/doc_end_user/current_document.php?id=13592) Technologies under consideration for future extensions of VSEI (version 3) [M. M. Hannuksela, J. Chen, S. Deshpande, Hendry, S. McCarthy] [WG 5 N 259)] (2024-03-29)

New elements from notes elsewhere in this report:

* …

It is noted that the list above may not be complete; if some adoption is missing that is recorded somewhere else in the meeting notes it shall also be considered included.

Remains valid – not updated: [JVET-AF2033](https://jvet-experts.org/doc_end_user/current_document.php?id=13593) Report of verification test on VVC multi-layer coding: Content layering [S. Iwamura, P. de Lagrange, M. Wien] [AG 5 N 105)] (2023-12-22)

[JVET-AG2034](https://jvet-experts.org/doc_end_user/current_document.php?id=13271) SEI messages for VSEI version 4 (Draft 1) [S. McCarthy, T. Chujoh, M. M. Hannuksela, G. J. Sullivan, Y.-K. Wang] [WG 5 preliminary WD N 258)] (2024-03-15)

Primary editor for this document: XXX.

[JVET-AG2035](https://jvet-experts.org/doc_end_user/current_document.php?id=13271) Test conditions for Generative Face Video Coding [XXXX, XXXX] (2024-XX-XX)

[JVET-AG2036](https://jvet-experts.org/doc_end_user/current_document.php?id=13271) Call for training materials for neural network-based video coding tool development [XXXX, XXXX] [WG 5 N 266)] (2024-02-09)

# Future meeting plans, expressions of thanks, and closing of the meeting

Future meeting plans were established according to the following guidelines (assuming face-to-face meetings):

* Meeting under ITU-T SG16 auspices when it meets (ordinarily starting meetings on the Tuesday or Wednesday of the first week and closing it on the Wednesday of the second week of the SG16 meeting – a total of 8-9 meeting days), and
* Otherwise meeting under ISO/IEC JTC 1/‌SC 29 auspices when its MPEG WGs meet (ordinarily starting meetings on the Thursday or Friday prior to the main week of such meetings and closing it on the same day as other MPEG WGs – a total of 8–9 meeting days).

In cases where an exceptionally high workload is expected for a meeting, an earlier starting date may be defined. In cases of online meetings, no sessions should be held on weekend days, such that meetings would typically start two days earlier.

Some specific future meeting plans (to be confirmed) were established as follows:

* During Wed. 17 – Wed. 24 April 2024, 34th meeting under ITU-T SG16 auspices in Rennes, FR,
* During Fri. 12 – Fri. 19 July 2024, 35th meeting under ISO/IEC JTC 1/‌SC 29 auspices in Sapporo, JP,
* During Fri. 1 – Fri. 8 November 2024, 36th meeting under ISO/IEC JTC 1/‌SC 29 auspices in Antalya, TR (confirmed by host one week after the closing of the JVET meeting),
* During Tue. 14 – Wed. 22 January 2025, 37th meeting under ITU-T SG16 auspices in Geneva, CH,
* During Wed. 26 March – Fri. 4 April 2025, 38th meeting under ISO/IEC JTC 1/‌SC 29 auspices, to be conducted as teleconference meeting,
* During Thu. 26 June – Fri. 4 July 2025, 39th meeting under ISO/IEC JTC 1/‌SC 29 auspices in Daejeon, KR,
* During October 2025, 40th meeting under ITU-T SG16 auspices, date and location t.b.d.
* During January 2026, 41st meeting under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.,

The agreed document deadline for the 34th JVET meeting was planned to be Wednesday 10 April 2024.

The JVET chair notes that contributions registered as JVET documents which do not directly relate to JVET agenda items, nor to potential future work items of JVET or joint work with other WGs, may not be considered for presentation in future meetings or joint meetings.

Huawei was thanked for providing video test materials with gaming content.

Marius Preda was thanked for maintaining the document site jvet-experts.org, as well as the document sites of JCT-VC and JCT-3V. Institut Mines-Télécom was thanked for hosting the sites.

The 33rd JVET meeting was closed at approximately XXXX hours UTC on XXday 2X Jan. 2024.

# Annex A to JVET report: List of documents

(Dates and times in the table below are in Paris/Geneva time.)

# Annex B to JVET report: List of meeting participants

The participants of the thirty-third meeting of the JVET, according to the participation records from the Zoom teleconferencing tool used for the meeting sessions (approximately XXX people in total, not including those who attended only the joint sessions with other groups), were as follows:

# Annex C to JVET report: Recommendations of the 14th meeting of ISO/IEC JTC 1/SC 29/WG 5 MPEG Joint Video Experts Team with ITU-T SG 16

**ISO/IEC JTC 1/SC 29/WG 5 N XXX**