|  |  |
| --- | --- |
| **Joint Video Experts Team (JVET)**  **of ITU-T SG21 WP3/21 and ISO/IEC JTC 1/SC 29**  37th Meeting, Geneva, CH, 14–22 January 2025 | Document: JVET-AK\_notes\_d8 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Meeting Report of the 37th Meeting of the Joint Video Experts Team (JVET), Geneva, 14–22 January 2025** | | |
| *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 WP3/21 and ISO/IEC JTC 1/‌SC 29 held its thirty-seventh meeting during 14–22 January 2025 meeting at the ITU headquarters facilities in Geneva, Switzerland. The meeting was held as a hybrid meeting, where remote participation was provided on best-effort basis for experts who were unable to travel.

For ISO/IEC purposes, JVET is alternatively designated ISO/IEC JTC 1/‌SC 29/‌WG 5, and this was the eighteenth 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 0930 CET on Tuesday 14 January 2025. Meeting sessions were held on all days including the weekend days of Saturday and Sunday 18 and 19 January 2025, until the meeting was closed at approximately XXXX hours CET on Wednesday 22 January 2025. Approximately XXX people attended the JVET meeting (XXX in person and XXX remotely), and approximately XXX input documents (not counting crosschecks, reports, and summary documents), 17 AHG reports, 2 EE summary reports, X BoG reports, and X incoming liaison document(s) were discussed. The meeting took place in a collocated fashion with a meeting of SG21 – one of the two parent bodies of the JVET, under whose auspices this JVET meeting was held. Various SC29 Working Groups and Advisory Groups were also meeting in Geneva with partial temporal overlap – where WG 5 is representing the Joint Video Experts Team and its activities from the perspective of the SC 29 parent body. 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-sixth JVET meeting in producing the following documents:

a) JVET documents

* JVET-AJ1004 Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP
* JVET-AJ1006 HEVC extensions and corrections
* JVET-AJ1011 HEVC white paper, also issued as AG 3 N 174
* JVET-AJ1012 Overview of IT systems used in JVET
* JVET-AJ2003 Guidelines for VTM-based software development
* JVET-AJ2005 Additions and corrections for VVC version 4 (Draft 10), also issued as WG 5 preliminary DAM N 330
* JVET-AJ2006 Additional SEI messages for VSEI version 4 (Draft 4), also issued as WG 5 preliminary DAM N 319
* JVET-AJ2007 Guidelines for NNVC software development
* JVET-AJ2009 Reference software for versatile video coding 2nd edition (Draft 2), also issued as WG 5 DIS N 322
* JVET-AJ2016 Common test conditions and evaluation procedures for neural network-based video coding technology
* JVET-AJ2019 Description of algorithms version 9 and software version 11 in neural network-based video coding (NNVC)
* JVET-AJ2020 Film grain synthesis technology for video applications ed. 2 (Draft 1)
* JVET-AJ2021 Verification test plan for VVC multilayer coding (update 5)
* JVET-AJ2022 Plan for subjective quality testing of the FGC SEI message (update 4)
* JVET-AJ2023 Exploration experiment on neural network-based video coding (EE1)
* JVET-AJ2024 Exploration experiment on enhanced compression beyond VVC capability (EE2)
* JVET-AJ2025 Algorithm description of Enhanced Compression Model 15 (ECM 15)
* JVET-AJ2026 Testing of video coding technology beyond conditions of exploration experiments
* JVET-AJ2027 Common test conditions for gaming applications
* JVET-AJ2030 Optimization of encoders and receiving systems for machine analysis of coded video content (draft 7), also issued as WG 5 CDTR N 323
* JVET-AJ2032 Technologies under consideration for future extensions of VSEI (version 6)
* JVET-AJ2035 Test conditions and evaluation procedures for generative face video coding
* JVET-AJ2037 Report on subjective quality testing of the FGC SEI message, also issued as AG 5 N 140

b) documents produced as WG 5 documents only:

* WG 5 N 318 Draft disposition of comments received on ISO/IEC 23002-7:202x/CDAM1
* WG 5 N 320 Request for ISO/IEC 23008-2:202x (6th ed.) Amd.1 Additional profiles and SEI messages
* WG 5 N 321 Disposition of comments received on ISO/IEC CD 23090-16:202x
* WG 5 N 324 Liaison statement to ISO/IEC JTC 1/SC 29/WG 1 (JPEG) on JPEG AI and explorations on video coding
* WG 5 N 325 Liaison statement to DVB on film grain synthesis
* WG 5 N 326 Liaison statement to ARIB, ATSC, SBTVD, and SCTE on film grain synthesis evaluation
* WG 5 N 327 Liaison statement to ATSC on multi-layer coding
* WG 5 N 329 Draft disposition of comments received on ISO/IEC 23090-3:202x/CDAM1

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

* JVET-AJ1004 Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP
* JVET-AJ1006 HEVC extensions and corrections
* JVET-AJ1011 HEVC white paper, also issued as AG 3 N 174
* JVET-AJ1012 Overview of IT systems used in JVET
* JVET-AJ2003 Guidelines for VTM-based software development
* JVET-AJ2005 Additions and corrections for VVC version 4 (Draft 10), also issued as WG 5 preliminary DAM N 330
* JVET-AJ2006 Additional SEI messages for VSEI version 4 (Draft 4), also issued as WG 5 preliminary DAM N 319
* JVET-AJ2007 Guidelines for NNVC software development
* JVET-AJ2009 Reference software for versatile video coding 2nd edition (Draft 2), also issued as WG 5 DIS N 322
* JVET-AJ2016 Common test conditions and evaluation procedures for neural network-based video coding technology
* JVET-AJ2019 Description of algorithms version 9 and software version 11 in neural network-based video coding (NNVC)
* JVET-AJ2020 Film grain synthesis technology for video applications ed. 2 (Draft 1)
* JVET-AJ2021 Verification test plan for VVC multilayer coding (update 5)
* JVET-AJ2022 Plan for subjective quality testing of the FGC SEI message (update 4)
* JVET-AJ2023 Exploration experiment on neural network-based video coding (EE1)
* JVET-AJ2024 Exploration experiment on enhanced compression beyond VVC capability (EE2)
* JVET-AJ2025 Algorithm description of Enhanced Compression Model 15 (ECM 15)
* JVET-AJ2026 Testing of video coding technology beyond conditions of exploration experiments
* JVET-AJ2027 Common test conditions for gaming applications
* JVET-AJ2030 Optimization of encoders and receiving systems for machine analysis of coded video content (draft 7), also issued as WG 5 CDTR N 323
* JVET-AJ2032 Technologies under consideration for future extensions of VSEI (version 6)
* JVET-AJ2035 Test conditions and evaluation procedures for generative face video coding
* JVET-AJ2037 Report on subjective quality testing of the FGC SEI message, also issued as AG 5 N 140

The following X documents were produced as WG 5 or SG21 documents only, without a corresponding JVET output document or direct repetition of their content in this meeting report:

* WG 5 N 275 Disposition of comments received on ISO/IEC DIS 14496-10:202x
* WG 5 N 277 Request for ISO/IEC 23002-7:2023/Amd.1
* WG 5 N 280 Disposition of comments received on ISO/IEC 23008-2:2023 (5th ed.) DAM 1
* WG 5 N 282 Request for ISO/IEC 23090-3:2023/Amd.1
* WG 5 N 285 Request for ISO/IEC 23090-16:202x
* WG 5 N 287 Disposition of comments received on ISO/IEC DIS 23091-3:202X
* WG 5 N 295, included in SG16-TD237/WP3, Liaison statement to ISO/IEC JTC 1/SC 29/WG 1 (JPEG) on JPEG AI and explorations on video coding
* WG 5 N 296, included in SG16-TD237/WP3, Liaison statement to 3GPP SA 4 on feasibility study on film grain synthesis
* WG 5 N 297, included in SG16-TD237/WP3, Liaison statement to ITU-R WP 6B on use cases of VVC multilayer profiles for broadcasting applications
* WG 5 N 298, included in SG16-TD237/WP3, Liaison statement to ARIB, ATSC, DVB, SBTVD, and SCTE on film grain synthesis investigations
* WG 5 N 299 List of AHGs established at the 15th WG 5 meeting

The following X draft revised ITU-T Recommendations were forwarded by JVET and Q6/21 for ITU-T Consent:

* [TD276/Plen](https://www.itu.int/md/meetingdoc.asp?lang=en&parent=T22-SG16-240415-TD-PLEN-0276) ITU-T H.264 (V15) "*Advanced video coding for generic audiovisual services*" (Rev.)
* [TD278/Plen](https://www.itu.int/md/meetingdoc.asp?lang=en&parent=T22-SG16-240415-TD-PLEN-0278) ITU-T H.266.2 (V2) "*Reference software for ITU-T H.266 versatile video coding*" (Rev.)
* [TD279/Plen](https://www.itu.int/md/meetingdoc.asp?lang=en&parent=T22-SG16-240415-TD-PLEN-0279) ITU-T H.265 (V10) "*High efficiency video coding*" (Rev.)
* [TD283/Plen](https://www.itu.int/md/meetingdoc.asp?lang=en&parent=T22-SG16-240415-TD-PLEN-0283) ITU-T H.273 (V4) "*Coding-independent code points for video signal type identification*" (Rev.)

For the organization and planning of its future work, the JVET established 18 “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 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 7 – 15 October 2025 under ITU-T SG21 auspices in Geneva, CH; during 14 – 23 January 2026 under ISO/IEC JTC 1/‌SC 29 auspices, to be conducted as teleconference meeting; during 24 April – 1 May 2026 under ISO/IEC JTC 1/‌SC 29 auspices in Santa Eulària, ES; during 7 – 15 July 2026 under ITU-T SG21 auspices in Geneva, CH; during October 2026 under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.; and during January 2027 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/> were shut down, but the most recent version of JCT-VC and JCT-3V documents can now be accessed directly via the JVET site.

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 21 Visual Coding Experts Group (VCEG) and ISO/IEC JTC 1/‌SC 29/‌WG 5. The parent bodies of the JVET are ITU-T WP3/21 and ISO/IEC JTC 1/‌SC 29.

The Joint Video Experts Team (JVET) of ITU-T WP3/21 and ISO/IEC JTC 1/‌SC 29 held its thirty-seventh meeting during 14–22 January 2025 meeting at the ITU headquarters facilities in Geneva, Switzerland. The meeting was held as a hybrid meeting, where remote participation was provided on best-effort basis for experts who were unable to travel.

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/2025_01_AK_Geneva/>.

## Primary goals

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

a) JVET documents

* JVET-AJ1004 Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP
* JVET-AJ1006 HEVC extensions and corrections
* JVET-AJ1011 HEVC white paper, also issued as AG 3 N 174
* JVET-AJ1012 Overview of IT systems used in JVET
* JVET-AJ2003 Guidelines for VTM-based software development
* JVET-AJ2005 Additions and corrections for VVC version 4 (Draft 10), also issued as WG 5 preliminary DAM N 330
* JVET-AJ2006 Additional SEI messages for VSEI version 4 (Draft 4), also issued as WG 5 preliminary DAM N 319
* JVET-AJ2007 Guidelines for NNVC software development
* JVET-AJ2009 Reference software for versatile video coding 2nd edition (Draft 2), also issued as WG 5 DIS N 322
* JVET-AJ2016 Common test conditions and evaluation procedures for neural network-based video coding technology
* JVET-AJ2019 Description of algorithms version 9 and software version 11 in neural network-based video coding (NNVC)
* JVET-AJ2020 Film grain synthesis technology for video applications ed. 2 (Draft 1)
* JVET-AJ2021 Verification test plan for VVC multilayer coding (update 5)
* JVET-AJ2022 Plan for subjective quality testing of the FGC SEI message (update 4)
* JVET-AJ2023 Exploration experiment on neural network-based video coding (EE1)
* JVET-AJ2024 Exploration experiment on enhanced compression beyond VVC capability (EE2)
* JVET-AJ2025 Algorithm description of Enhanced Compression Model 15 (ECM 15)
* JVET-AJ2026 Testing of video coding technology beyond conditions of exploration experiments
* JVET-AJ2027 Common test conditions for gaming applications
* JVET-AJ2030 Optimization of encoders and receiving systems for machine analysis of coded video content (draft 7), also issued as WG 5 CDTR N 323
* JVET-AJ2032 Technologies under consideration for future extensions of VSEI (version 6)
* JVET-AJ2035 Test conditions and evaluation procedures for generative face video coding
* JVET-AJ2037 Report on subjective quality testing of the FGC SEI message, also issued as AG 5 N 140

b) documents produced as WG 5 documents only:

* WG 5 N 318 Draft disposition of comments received on ISO/IEC 23002-7:202x/CDAM1
* WG 5 N 320 Request for ISO/IEC 23008-2:202x (6th ed.) Amd.1 Additional profiles and SEI messages
* WG 5 N 321 Disposition of comments received on ISO/IEC CD 23090-16:202x
* WG 5 N 324 Liaison statement to ISO/IEC JTC 1/SC 29/WG 1 (JPEG) on JPEG AI and explorations on video coding
* WG 5 N 325 Liaison statement to DVB on film grain synthesis
* WG 5 N 326 Liaison statement to ARIB, ATSC, SBTVD, and SCTE on film grain synthesis evaluation
* WG 5 N 327 Liaison statement to ATSC on multi-layer coding
* WG 5 N 329 Draft disposition of comments received on ISO/IEC 23090-3:202x/CDAM1

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.

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 Tuesday, 7 January 2025. Any documents uploaded after 1159 hours Paris/Geneva time on Wednesday 8 January 2025 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-AK0239 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-AK0XXX (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-AK0XXX (a document describing …), uploaded 01-XX,
* … .

All cross-verification reports at this meeting 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-AK0046, JVET-AK0051, JVET-AK0116, JVET-AK0120, JVET-AK0143, JVET-AK0144, JVET-AK0145, JVET-AK0163, JVET-AK0174, and JVET-AK0182.

The following cross-verification reports were still missing by the end of the meeting, but were uploaded later: JVET-AK0XXX, …. (not applicable for current meeting, kept for future use) The following reports had not become available yet three weeks after the end of the meeting: JVET-AK0XXX, … . These were 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-AK0077, JVET-AK0078, and JVET-AK0220 which were empty of results and were flagged as late in the list above, based on the time of the first reasonable document upload.

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-AJ1000, the Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP JVET-AJ1004, the HEVC extensions and corrections (draft 1) JVET-AJ1006, the White paper on HEVC JVET-AJ1011 (not delivered yet by beginning of the 37th meeting), the Overview of IT systems used in JVET JVET-AJ1012, the Guidelines for VTM-based software development JVET-AJ2003, the Additions and corrections for VVC version 4 (Draft 10) JVET-AJ2005, the Additional SEI messages for VSEI version 4 (Draft 4) JVET-AJ2006, the Guidelines for NNVC software development JVET-AJ2007, the Reference software for versatile video coding 2nd edition (Draft 2) JVET-AJ2009, the Common test conditions and evaluation procedures for neural network-based video coding technology JVET-AJ2016, the Description of algorithms version 9 and software version 11 in neural network-based video coding (NNVC) JVET-AJ2019, the Film grain synthesis technology for video applications ed. 2 (Draft 1) JVET-AJ2020 (not delivered yet by beginning of the 37th meeting), the Verification test plan for VVC multilayer coding (update 5) JVET-AJ2021, the Plan for subjective quality testing of the FGC SEI message (update 4) JVET-AJ2022, the Description of the EE on Neural Network-based Video Coding JVET-AJ2023, the Description of the EE on Enhanced Compression beyond VVC capability JVET-AJ2024, the Algorithm description of Enhanced Compression Model 15 (ECM 15) JVET-AJ2025, the Testing of video coding technology beyond conditions of exploration experiments JVET-AJ2026, the Common test conditions for gaming applications JVET-AJ2027, the Optimization of encoders and receiving systems for machine analysis of coded video content (draft 7) [JVET-AJ2030](https://jvet-experts.org/doc_end_user/current_document.php?id=12584), the Technologies under consideration for future extensions of VSEI (version 6) JVET-AJ2032, the Test conditions and evaluation procedures for generative face video coding JVET-AJ2035, and the Report on subjective quality testing of the FGC SEI message JVET-AJ2037 (not delivered yet by beginning of the 37th meeting), had been completed and those which were available 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 version 23.5 and 23.6, ECM version 15.0, and NNVC version 11.0 were also approved.

Only minor editorial issues were found in the meeting report JVET-AJ1000; 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 Annexes B1 and B2 of this report.

The meeting was open to those qualified to participate either in ITU-T WP3/21 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 the Q6/21 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. On-site participants furthermore were required to register via the ITU-T website in order to get access to the headquarter facilities.

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 times of meeting sessions followed the needs of the face-to-face meeting, with highest priority given to the aim of achieving the goals of the meeting. Typical meeting hours were expected to be 0900-2000 CET with coffee breaks and lunch breaks as appropriate, however some early morning or late-night sessions were anticipated to be necessary. Sessions were announced in the online JVET calendar in advance as far as possible, but it was anticipated that some activities (such as breakout sessions) could be held at short notice.

## ISO and IEC Code of Conduct reminders

Participants were reminded of the ISO and IEC Codes of Conduct, found at

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

<https://www.iec.ch/basecamp/iec-code-conduct-technical-work>

These include points relating to:

* Behave ethically
* Promote and enable all voices to be heard
* Engage constructively in ISO and IEC activities
* Respect others (in meetings, when writing, on social media)
* Declare actual and potential conflicts of interest and manage them appropriately
* Protect confidential information
* Protect ISO and IEC assets
* Avoid and prevent any form of bribery or corruption
* Uphold the consensus process
* When disputes arise, escalate, resolve and uphold agreed resolution

## 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](https://dms.mpeg.expert/doc_end_user/documents/89_London/wg11/w10791.zip) 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/>. The site contains links for creating a user account for document uploading and for contacting an administrator – e.g., to update the email address associated with an account, and experts were reminded to keep this information up to date. It was reminded to send a notice to the chairs in cases of changes to document titles, authors, etc.

It was noted that the previous document sites <http://phenix.int-evry.fr/jvet/>, <http://phenix.int-evry.fr/jct/>, and <http://phenix.int-evry.fr/jct3v/> were shut down, but the most recent version of JCT-VC and JCT-3V documents can now be accessed directly via the JVET site. A mirror for JCT-VC, JCT-3V and JVET document access (including all versions with original upload dates) will shortly be provided via the ITU ftp site, providing a “documents” subfolder in the directory of the respective meeting.

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 1277 (as of 13 January 2025). 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.

It is further emphasized that the document JVET-AJ1012 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)
* **EIP**: Extrapolation based intra prediction
* **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
* **GFV**: Generative face video
* **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
* **NNPF**: Neural network post filter
* **NNVC**: Neural network-based video coding (experimental software package)
* **NSQT**: Non-square quadtree
* **NSPT**: Non-separable primary transform
* **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
* **ROI**: Region of interest.
* **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.
* **SPO**: SEI processing order.
* **SPS**: Sequence parameter set (as in AVC and HEVC).
* **SR**: Super resolution .
* **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/public.
* **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 WGs of ISO/IEC JTC 1/SC 29).
* **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 (update)

* 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 (Ed. 3) 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 (Ed. 2) 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 (Ed. 2) 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 IPT-PQ-C2 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 closed 2024-04-15.
    - H.264 V15 Consented 2024-04-26, last call began 2024-07-16, approved 2024-08-13, published 2024-11-13 (after the current meeting)
    - ISO/IEC 14496-10 (Ed. 11) FDIS issued at 34th meeting 2024-04, DIS approved for registration as FDIS 2024-09-17, pending FDIS ballot
  + 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 23008-2:2023 (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, published 2023-10-30
  + 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 with new profiles and SEI messages at 32nd meeting 2023-10, DAM ballot closed 2024-04-08
    - H.265 V10 Consented 2024-04-26, last call opened 2024-07-01, closed 2024-07-28, approved 2024-07-29, pre-published 2024-08-06, published 2024-10-07
    - ISO/IEC 23008-2 (Ed. 6) FDIS issued from 35th meeting 2024-04, DIS approved for registration as FDIS 2024-10-03, FDIS registered for formal approval 2024-10-13, pending FDIS ballot
  + 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, FDIS ballot opened 2022-06-29, closed 2022-08-24, published 2022-09-25
  + ISO/IEC 23090-3:2024 (Ed. 3), initated as (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 at WG level 2023-07, FDIS ballot opened 2024-05-11, closed 2024-06-26, published 2024-07-17
  + ITU-T H.266 V3 Consented 2023-07, approved 2023-09-29 and pre-published 2023-09, published 2023-11-29
  + ISO/IEC 23090-3:2024/CDAM 1 Request & CDAM issued 2024-04, consultation deferred due to meeting timing, updated text issued 2024-07, consultation initiated 2024-09-05, closed 2024-10-31, ready for action at the current meeting
  + 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 (Ed. 1) 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 23090-15:2024 (Ed. 2) began as Amd.1 Operation range extensions – CDAM 1 issued from 24th meeting 2021-10, 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, ballot opened 2024-04-08, closed 2024-06-03, published 2024-07-04.
    - 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
    - H.266.2 V2 Consented 2024-04-26, last call opened 2024-05-16, closed 2024-06-12, approved 2024-06-13, pre-published 2024-08-06, published 2024-10-09
    - ISO/IEC 23090-16:202x (Ed. 2) Request & CD issued 2024-04, consultation deferred due to meeting timing, consultation initiated 2024-09-06, closed 2024-11-01, ready for action at the current meeting (Roughly corresponding H.266.2 V2 already approved and published in ITU-T)
* 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 23002-7:2024 (Ed. 3) began as (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, FDIS ballot began 2024-07-25, closed 2024-09-20, published 2024-10-30
  + ITU-T H.274 V3 Consent 2023-07, approved 2023-09-29, pre-published 2023-10-11, published 2024-03-12.
  + ISO/IEC 23002-7:202x (Ed. 3) Request & CDAM 1 issued 2024-04, consultation deferred due to meeting timing, updated text issued 2024-07, CD consultation initiated 2024-09-05, closed 2024-10-31, ready for action at the current meeting
* 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 issued 2024-04 (after waiting for publication of SMPTE ST 2128, then proceeding), FDIS registered for formal approval 2024-10-24, pending FDIS ballot.
  + ITU-T H.273 V3 Consent 2023-07, approved 2023-09, not published due to waiting for publication of SMPTE ST 2128.
  + ITU-T H.273 V4 Consented 2024-04-26, last call opened 2024-06-16, closed 2024-07-13, approved 2024-07-14, pre-published 2024-08-06, published 2024-10-07.
* 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 (Ed. 1) published 2018-08-06
* 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 (Ed. 1) 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 23002-9:2024 (Ed. 1) 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, DTR text was issued from the 31st meeting in July 2023, put on hold by ISO staff editors, DTR ballot opened 2024-03-05, closed 2024-04-30, NB comments handled by ISO staff editor in consultation with project editors, published 2024-07-24
  + ITU-T provisional name H.Sup-FGST
* Optimization of encoders and receiving systems for machine analysis of coded video content (twin text)
  + ISO/IEC 23888-3 (Ed. 1) Request for subdivision issued from 33rd JVET meeting 2024-01, CDTR issued 2024-04, consultation deferred due to meeting timing, updated text issued 2024-07, consultation further deferred due to meeting timing, ready for action at the current meeting
  + ITU-T provisional name H.Sup-MACVC
* The following freely available standards are published here in ISO/IEC:  
  <https://standards.iso.org/ittf/PubliclyAvailableStandards/index.html> as of the time of the current meeting:
  + 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 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 the time of the current meeting. (These should be checked for previously issued requests for free availability.)
  + ISO/IEC 23002-7:2024 (Ed. 3) – VSEI
  + ISO/IEC 23008-2:2023 (Ed. 5) HEVC, published 2023-10-30
  + ISO/IEC 23008-5:2017 (Ed. 2) Reference software for HEVC, published 2017-03-01
  + ISO/IEC 23008-5:2017/AMD 1:2017 Reference software for HEVC screen content coding extensions, published 2017-11-09
  + 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
  + ISO/IEC 23090-3:2024 (Ed. 3) VVC, published 2024-07-17
  + ISO/IEC 23090-15:2024 (Ed. 2) Conformance specification for VVC, published 2024-07-04.
* 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 (update)

* VVC ISO/IEC 23090-3:2024/CDAM 1, Request & CDAM issued 2024-04, consultation deferred due to meeting timing, updated text issued 2024-07, consultation initiated 2024-09-05, closed 2024-10-31, ready for action at the current meeting.
* VVC reference software ISO/IEC 23090-16:202x (Ed. 2) Request & CD issued 2024-04, consultation deferred due to meeting timing, consultation initiated 2024-09-06, closed 2024-11-01, ready for action at the current meeting. (Roughly corresponding H.266.2 V2 already approved and published in ITU-T.)
* VSEI ISO/IEC 23002-7:202x (Ed. 3) / CDAM 1 Request & CDAM issued 2024-04, consultation deferred due to meeting timing, updated text issued 2024-07, CD consultation initiated 2024-09-05, closed 2024-10-31, ready for action at the current meeting.
* 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, put on hold by ISO staff editors, DTR ballot opened 2024-03-05, closed 2024-04-30, NB comments handled by ISO staff editor in consultation with project editors, published 2024-07-24. ITU-T provisional name H.Sup-FGST. ITU-T approval delayed to January 2025 due to the delay on the ISO/IEC side.
* Optimization of encoders and receiving systems for machine analysis of coded video content – ISO/IEC 23888-3 Request for subdivision issued from 33rd JVET meeting 2024-01, CDTR issued 2024-04, consultation deferred due to meeting timing, updated text issued 2024-07, consultation further deferred due to meeting timing, ready for action at the current meeting. ITU-T provisional name H.Sup-MACVC.
* 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 Tuesday 14 January at 0930 CET were as follows.

* Timing and organization of the meeting and online access and calendar posting of session plans were reviewed
  + The initial number of documents was slightly lower than in last meeting (approximately 205 vs. 220 by the time of opening the meeting) – parallel sessions were announced to be necessary.
  + Evening sessions may also be necessary, typically expecting to meet until around 2000.
  + Parallel sessions starting the afternoon of the first day (EE and HLS), and probably continuing until the weekend.
  + Friday 1400-1830: Joint ITU/ISO workshop on “Future video coding – advanced signal processing, AI and standards” – no JVET meeting in parallel with that, but sessions may resume thereafter. Separate registration is necessary for the workshop.
* In the context of registration for JVET in the ITU website, it was found that some experts (also some who had participated in-person in the previous meeting) were neither listed as WG 5 nor as ITU members. In most cases this could be clarified, confirming that the process of becoming member of one of the parent bodies was experiencing delays. It will be necessary to enforce this check for qualification more severely in the future, also for online participants.
* It was noted that the JVET document template has a new header with WP3/21 as ITU parent body. Roughly half of the uploaded inputs obviously used own edits of old documents with the old header.
* Plans for subsequent hybrid meetings (with best-effort remote access) were reviewed:June/July 2025 (Daejeon), Oct. 2025 (Geneva), April 2026 (Santa Eulària), June/July 2026 (Geneva, dates to be clarified).
* March/April 2025 and January 2026 meetings will be virtual. Considering the rigid time constraints of virtual meetings, and considering the increased number of input documents since the last virtual meeting in January 2024, it may be necessary to extend them by one or two days, or meet over weekends.
* Depending on the status of preparing future standardization activities, it may be necessary to extend the duration of meetings, anyway (both for hybrid and virtual)
* Significant workload can be expected at this meeting for AHG17 activities – possible need of having a hybrid AHG meeting in Aachen during second week of March
* The meeting logistics, agenda, working practices, policies, and document allocation considerations were reviewed.
  + Access to the meeting was provided using Zoom. The meeting notes by the session chair were to be continually shared via zoom screen sharing. In the meeting room C, it is also possible to show them on a separate projection screen in parallel with another presentation.
  + Having text and software available is crucial (and not just arriving at the end of the meeting).
* The results of the previous meeting and the meeting report JVET-AJ1000 were reviewed. Only minor issues in the meeting report were noted and were not considered sufficient to warrant issuing a revision.
* At the current meeting, on-site attendance will be recorded via the traditional sign-in sheet. Participants were asked to correct their affiliation and email in case these changed.
* All cross-check documents were, but number of late non-cross-check documents seemed to have slightly increased relative to the last meeting. Several contributions did not report any results initially, and were also flagged as late. There were no objections voiced in the opening plenary to the consideration of late contributions.
* There were again a few documents registered where authors’ given names were not abbreviated according to the JVET custom (which helps produce shorter headings in the JVET meeting report), 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. Reserving numbers without filling a precise title shall also be avoided – the chair may flag such documents as withdrawn, as they cannot be allocated to a certain category in the meeting notes.
* Experts were asked to always register JVET documents via the “jvet-experts.org” site, not via the MPEG DMS site, as WG 5 docs (as that feature of the DMS site has not been working properly).
* 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. Provisional titles such as “EEx contribution” shall be avoided.
* The practice introduced in the last meeting, submitting WG 5 N-numbered output documents only in cases of standards text submitted for ballot, DoCs, standards/parts requests, and meeting reports has turned out to be useful. It is basically the same practice that had always been exercised on the ITU side.
* Document JVET-AJ1012 was produced, which includes the changes in the sites for hosting the JVET test sequences, as well as the new location of JCT-VC and JCT-3V documents. Only the newest versions of the JCT-VC and JCT-3V documents are available from the links in the JVET site, but it is still planned to provide all approved versions in the future (although this may take some time). During or after the current meeting, It is further planned to set up a second source for JVET, JCT-VC and JCT-3V docs in the ITU ftp directories of each meeting. This will include all versions of documents with original upload times.
* The following ballot results had become available through the SC 29 secretariat (kept for future use).
  + …
* DAM/DIS for VVC and VSEI could be done at the current meeting, but no ballot results could become available in time for the March/April meeting. The FDAM/FDIS could be in July 2025 at the earliest. This might be achieved with a very long editing period (potentially reaching into the next meeting). Another option would be to target the 4th editions for October 2025, both for ITU-T and ISO/IEC, and send them to ballot early in April, such that still improvements might be made during the March/April meeting.
* The CDTR on machine analysis was submitted, but it would be necessary that the VSEI v4 Amd. reaches DAM/DIS stage before it is promoted to TR (as SEI messages in VSEI v4 are referred to in the draft of the TR).
* DIS of VVC software was submitted for ballot – matching the ITU edition that was issued in April. A next edition could be targeted for October, considering implementations of new SEI messages.
* The primary goals of the meeting were:
  + Film grain TR to be submitted for ITU approval – which version? V1 with corrections/improvements, to be taken from JVET-AJ2020
  + New versions of VVC and VSEI (JVET-AK2005/2006, <preliminary> DAM texts, not for ITU consent yet)
  + In the last meeting, a new edition of HEVC software was considered to be possible at the current meeting. This would however depend on the implementation status of the software for the new multiview profiles (provided they would be complete).
  + Any action items on reference software JM?
  + Lot of activity in AHG17 – planning of viewing sessions needed.
  + Any viewing for film grain and multi-layer?
  + Exploration Experiments
    - Neural network-based video coding
    - Enhanced compression beyond VVC
  + Any liaison communication (through ITU)?
* New training materials for NNVC – extended BVI set.
* It was suggested that the VVC white paper could benefit from an update, and generating a separate one on VSEI would be useful – first draft in the current meeting?
* Joint meetings were expected with MPEG AG 5 (on matters involving visual quality assessment), with MPEG WG 2 Requirements and ITU-T VCEG on future video standardization. Joint meetings will be held in ITU premises, not in Crowne Plaza Hotel where other MPEG WGs will meet. (see section 7.3).
* JVET cannot meet during plenaries of the ITU parent body (and will need to finish by Wednesday next week around lunchtime). Due to time constraints, JVET sessions will be held during MPEG WGs’ information exchange meetings on Monday morning and Wednesday morning.
* As a follow-up to liaison communication after the April meeting, parent bodies should conduct discussion about future JVET management structures. SC 29 had requested all MPEG WGs to appoint convener support teams by January 2025 at latest.
* Principles of standards development were discussed.
* Scheduling of sessions was discussed – see under sections 2.6 and 2.12.

## Scheduling of discussions

The times of the meeting sessions followed the needs of the onsite meeting arrangements, with highest priority given to the aim of achieving the goals of the meeting. Typical meeting hours were expected to be in the range of 0900-2000 CET with coffee breaks and lunch breaks as appropriate. Sessions were announced in the JVET calendar and the ITU schedule system as far as possible in advance, although it was acknowledged that some activities (such as breakout sessions) might be held at short notice.

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:

* Tue. 14 Jan., 1st day
  + Morning session:
    - 0930–1030 Opening remarks, review of practices, agenda, IPR policy reminder
    - 1030–1315 Reports of AHGs 1-11
  + Afternoon sessions:
    - 1430–1545 Reports of AHGs 12-17
    - 1600–2015 EE1/EE2 review (room C)
    - 1530–2000 HLS (room A, chaired by Jill Boyce)
* Wed. 15 Jan., 2nd day
  + Morning sessions:
    - 0900–1315 EE2 review (room C)
    - 0900–1300 HLS (room A, chaired by Jill Boyce)
  + Afternoon sessions:
    - 1430–1915 EE2 review, EE2 related, non-EE2 (room C)
    - 1400–2000 HLS (room A, chaired by Jill Boyce)…
* Thu. 16 Jan., 3rd day
  + Morning sessions:
    - 0900–1300 non-EE2 (room C)
    - 0900–1300 HLS (room A, chaired by Jill Boyce)
  + Afternoon sessions:
    - 1400–2000 non-EE2, EE1 related (room C)
    - 1400–2000 HLS (room A, chaired by Jill Boyce)
* Fri. 17 Jan., 4th day
  + Morning sessions:
    - 0900–1300 5.1.4 NNVC/SADL, 4.9 ECM tool complexity, 4.10 Opt. for MA, 4.11 encoder optimization (room C)
    - …
  + 1400–1830 Workshop on “Future video coding – advanced signal processing, AI and standards” – announcement/registration [on the ITU website](https://www.itu.int/en/ITU-T/Workshops-and-Seminars/2025/0117/Pages/default.aspx)
* Sat. 18 Jan., 5th day
  + Morning sessions:
    - 0900–1300 Review 4.11-4.16, (room C)
    - 0900–1300 HLS (room A, chaired by Jill Boyce)
  + Afternoon sessions:
    - 1500–1600 remaining docs 5.2.4 (room C)
    - 1600–1800 (joint with AG 5) 4.17 + viewing results + further planning
    - 1400–1800 HLS (room A, chaired by Jill Boyce)
    - 1800–1930 4.10 (room C)
* Sun. 19 Jan., 6th day
  + Morning sessions:
    - 0900–1345 JVET plenary: Reports from tracks; coordination/planning, 4.1, 4.7
    - 1100–1300 Remaining EE2, revisits EE1/EE2 (room C)
  + Afternoon sessions:
    - 1500–1730 (joint with AG 5) Remaining 4.17 + further discussion/planning from viewing results (room C)
* 1400–XXXX HLS (room A, chaired by Jill Boyce)Monday. 20 Jan., 7th day
  + Morning sessions:
    - 0900–1200 JVET session: HLS (room C)
  + Afternoon sessions:
    - 1300–1605 EE1/EE2 remaining doc review & revisits, EE planning (room C)
    - 1615–1730 Joint meeting: Next gen video standard (room Popov)
    - Viewing of new sequences t.b.a.
* Tuesday. 21 Jan., 8th day
  + Morning sessions:
    - 0900–1300 JVET session: Any remaining doc review & revisits, output & AHG planning, output doc review (room C)
    - …
  + Afternoon sessions:
    - 1400–XXXX further selection of sequences from viewing, planning of next actions in preparing CfE (room C)
    - …
* Wed. 22 Jan., 9th day
  + 0900–XXX Wrapup plenary:
    - Remaining revisits
    - EE descriptions, DoCRs
    - Establishment of AHGs
    - Review and approval of output docs
    - Software timelines
    - Approval of WG 5 meeting recommendations
    - Future planning, a.o.b.

## Contribution topic overview (update categories and doc counts)

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: Development, deployment and advertisement of standards (6)
  + AHG2: Text development and errata reporting (2)
  + AHG3: Test conditions (0)
  + AHG3: Software development (2)
  + AHG4: Subjective quality testing and verification testing (2)
  + AHG4: Test and training material (3)
  + AHG4: Codec performance with alternative test material and non-CTC conditions (0)
  + AHG5: Conformance test development (0)
  + AHG7: ECM tool assessment (2)
  + AHG8: Optimization of encoders and receiving systems for machine analysis of coded video content (7)
  + AHG10: Encoding algorithm optimization (3)
  + AHG13: Film grain synthesis (0)
  + Implementation studies (3)
  + Profile/tier/level specification (3)
  + Gaming content compression (5)
  + Generative face video (1)
* 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 (26) (section 5.1)
  + AHG6/AHG12 and EE2: Enhanced compression beyond VVC capability (94) (section 5.2)
* AHG9: High-level syntax (HLS) proposals (section 6) with subtopics
  + SEI messages in VSEIv4 (32) (section 6.1)
  + SEI messages in TuC doc (29) (section 6.2)
  + SEI messages on other topics (4) (section 6.3)
  + Non-SEI HLS aspects (0) (section 6.4)
* Joint meetings, plenary discussions, BoG reports (2) liaison (3), summary of actions (section 7)
* 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 1030–1315 and 1430–1545 on Tuesday 14 Jan. 2025 (chaired by JRO).

[JVET-AK0001](https://jvet-experts.org/doc_end_user/current_document.php?id=15168) JVET AHG report: Project Management (AHG1) [J.-R. Ohm (chair), G. J. Sullivan (vice chair)]

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

The number of subscribers (on the day before the beginning of the current meeting) was 1277 (compared to 1270 by the time of the previous meeting). Furthermore, the previous lists of joint teams (which were still kept open as archives) had the following number of subscribers:

* JCT-VC – 1158 subscribers
* JCT-3V – 680 subscribers
* JVT-experts – 2072 subscribers

It is likely that many subscriptions of these three reflectors have become obsolete, as no emails are being sent over them, and therefore automatic removal of participants causing permanent bounces through invalid email addresses (as with the main JVET reflector) cannot be executed by the postorius management system.

1. **Goals and activity**

The work of the JVET overall had proceeded well in the interim period with higher number of input documents (as compared to the previous meeting) submitted to the current meeting. Intense discussion had been carried out on the group email reflector, and most output documents from the preceding meeting had been produced as planned.

Output documents from the preceding meeting had been made initially available at the JVET web site (<https://jvet-experts.org/>) or the ITU-based JVET ftp site (<http://wftp3.itu.int/av-arch/jvet-site/2024_11_AJ_Kemer/>). It is noted that the previous document sites http://phenix.int-evry.fr/jvet/, http://phenix.int-evry.fr/jct/, and http://phenix.int-evry.fr/jct3v/ were shut down, but JCT-VC and JCT-3V documents can now be accessed directly via the JVET site. During or shortly after the 37th meeting, all documents of JCT-VC, JCT-3V and JVET will also be made available from the ITU-based ftp site, where sub-folders ‘./documents/’ are planned to be created in the folders of the respective meetings. This will provide the files with the original time stamps when the original versions of documents were uploaded.

The list of output documents produced since the last meeting included the following, particularly:

* JVET-AJ1004 Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP [Posted 2025-01-06]
* JVET-AJ1006 HEVC extensions and corrections [Posted 2024-12-02]
* JVET-AJ1011 HEVC white paper, also issued as AG 3 N 174 [Posted 2025-01-XX]
* JVET-AJ1012 Overview of IT systems used in JVET [Posted 2025-01-06]
* JVET-AJ2003 Guidelines for VTM-based software development [Posted 2025-01-12]
* JVET-AJ2005 Additions and corrections for VVC version 4 (Draft 10), also issued as WG 5 preliminary DAM N 330 [Posted 2024-11-26, last update 2024-12-30]
* JVET-AJ2006 Additional SEI messages for VSEI version 4 (Draft 4), also issued as WG 5 preliminary DAM N 319 [Posted 2025-12-07, last update 2024-12-18]
* JVET-AJ2007 Guidelines for NNVC software development [Posted 2024-12-04]
* JVET-AJ2009 Reference software for versatile video coding 2nd edition (Draft 2), also issued as WG 5 DIS N 322 [Posted 2024-12-30]
* JVET-AJ2016 Common test conditions and evaluation procedures for neural network-based video coding technology [Posted 2024-12-04]
* JVET-AJ2019 Description of algorithms version 9 and software version 11 in neural network-based video coding (NNVC) [Posted 2025-01-08]
* JVET-AJ2020 Film grain synthesis technology for video applications ed. 2 (Draft 1) [Posted 2025-01-XX]
* JVET-AJ2021 Verification test plan for VVC multilayer coding (update 5) [Posted 2025-01-13]
* JVET-AJ2022 Plan for subjective quality testing of the FGC SEI message (update 4) [Posted 2025-01-13]
* JVET-AJ2023 Exploration experiment on neural network-based video coding (EE1) [Posted 2024-11-08, last update 2024-12-12]
* JVET-AJ2024 Exploration experiment on enhanced compression beyond VVC capability (EE2) [Posted 2024-11-08, last update 2024-12-11]
* JVET-AJ2025 Algorithm description of Enhanced Compression Model 15 (ECM 15) [Posted 2025-01-09]
* JVET-AJ2026 Testing of video coding technology beyond conditions of exploration experiments [Posted 2024-12-03]
* JVET-AJ2027 Common test conditions for gaming applications [Posted 2024-11-22]
* JVET-AJ2030 Optimization of encoders and receiving systems for machine analysis of coded video content (draft 7), also issued as WG 5 CDTR N 323 [Posted 2024-12-31]
* JVET-AJ2032 Technologies under consideration for future extensions of VSEI (version 6) [Posted 2024-12-06, last update 2024-12-07]
* JVET-AJ2035 Test conditions and evaluation procedures for generative face video coding [Posted 2024-11-26, last update 2024-12-09]
* JVET-AJ2037 Report on subjective quality testing of the FGC SEI message, also issued as AG 5 N 140 [Posted 2025-01-XX]

The seventeen *ad hoc* groups had made progress, and reports from those activities had been submitted. Furthermore, two exploration experiments (EE) on neural network-based video coding and on enhanced compression beyond VVC capability were conducted.

The arrangements for the 37th meeting had been announced in the JVET reflector, in the JVET logistics document (<https://www.itu.int/wftp3/av-arch/jvet-site/2025_01_AK_Geneva/JVET-AK_Logistics.docx>), and in the WG 5 calling notice (N 331) and agenda (N 332) for the 18th WG 5 meeting.

Software integration was finalized approximately according to the plan.

Various problem reports relating to asserted bugs in the software, draft specification text, and reference encoder description had been submitted to an informal "bug tracking" system. That system is not intended as a replacement of our ordinary contribution submission process. However, the bug tracking system was considered to have been helpful to the software coordinators and text editors. The bug tracker reports had been automatically forwarded to the group email reflector, where the issues were discussed – and this is reported to have been helpful.

Roughly 205 input contributions (not counting the AHG reports, summary reports and crosschecks) had been registered for consideration at the current meeting.

It is further noted that, starting from the twentieth JVET 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 minor enhancement work on the *Advanced Video Coding* (AVC) standard, associated conformance test sets and reference software.

To retain a consistent numbering scheme, the number range of output documents starting from 1001 was reserved for the previous JCT-VC topic items listed above, whereas the number range starting from 2001 was retained for VVC, VSEI and exploration activities.

A preliminary basis for the document subject allocation and meeting notes for the 37h meeting had been made publicly available on the ITU-hosted ftp site as <http://wftp3.itu.int/av-arch/jvet-site/2025_01_AK_Geneva/JVET-AK_notes_d0.docx>.

1. **Recommendations**

* The AHG recommends its continuation.
* The AHG recommends communication between JVET’s parent bodies about appointment of an extended management support team.

[JVET-AK0002](https://jvet-experts.org/doc_end_user/current_document.php?id=15169) 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)]

**Ad hoc group activity**

**Output documents produced**

**JVET-AJ1004 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-AJ meeting:

* For HEVC:
  + Removed three items that have been addressed by the following integrations in JVET-AJ1006:
    - JVET-AJ0213 On Multiview Main 10 profiles
    - Bug fix: The variable subBitstream was used without being defined for the Multiview Extended or Multiview Extended 10 profiles.
    - Bug fix: Change the general\_profile\_idc value for the multiview format range extension profiles from 10 to 14, to differ from that for the multiview format range extension profiles.
* For VVC:
  + Spec ticket [#1650](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1650)

**JVET-AJ2005 Additions and corrections for VVC version 4 (Draft 10)**

This document contains the draft text for changes to the Versatile Video Coding (VVC) standard (Rec. ITU-T H.266 | ISO/IEC 23090-3).

**Changes to be integrated or checked:**

1. Elements from JVET-AG1004 (errata), and bug fixes from JVET-AH0002 for tickets [#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), [#1624](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1624) (Incorrect indexing in computation of motion vector offset), [#1628](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1635) (Derivation of ModeTypeCondition should say “one or more”), [#1630](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1630) (Missing equations for applying AmvrShift), [#1631](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1631) (Should “Motion vector storing process for geometric partitioning mode” store HpelIfIdx? BBross: HpekIfIdx should always be written since it might be accessed later so it is definitely missing in "Motion vector storing process for geometric partitioning mode".), [#1632](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1632) (Incorrect indexing used for choosing matrix intra sample prediction), [#1634](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1634) (Matrices QStateTransTable,levelScale,AlfFixFiltCoeff,AlfClassToFiltMap are incorrectly transposed).

**Changes that have been integrated for the 36th JVET meeting in November 2024:**

1. JVET-AJ0048 On SeiProcessingOrderSeiList and SpoProcessSeiList
2. JVET-AJ0049 NNPFC and NNPFA interface text in VVC
3. JVET-AJ0129 Miscellaneous comments on the SPO SEI message
4. JVET-AJ0151 Move digitally signed content SEI messages to VSEI draft
5. JVET-AJ0207 Move GFV and GFVE to VSEI draft
6. JVET-AJ0114 On signalling input pictures shift in NNPFA SEI message (the aspect regarding the establishment of input picture lists for NNPF)

**JVET-AJ2006 Additional SEI messages for VSEI version 4 (Draft 4)**

This document contains the draft text for changes to the versatile supplemental enhancement information messages for coded video bitstreams (VSEI) standard (Rec. ITU-T H.274 | ISO/IEC 23002-7), to specify additional SEI messages, including encoder optimization information, source picture timing information , object mask information, modality information, text description information, generative face video, generative face video enhancement, digitally signed content initialization, digitally signed content selection, and digitally signed content verification SEI messages and updates to the neural-network post-filter characteristics SEI message.

**Editors’ notes and changes yet to be integrated:**

* All automatically generated numbering and cross-reference fields will need to be updated after the changes are integrated into the basis text.Check the use of “pertain”; the word is not used in a similar way anywhere in the existing standard.payloadType values are to be specified the corresponding VVC, HEVC or AVC draft text. The use of prefix SEI NAL unit was enabled for all new SEI messages, whereas the use of suffix SEI NAL unit was enabled for encoder optimization information and object mask information SEI messages, since their content may be determined while the encoder encodes the associated picture.
* (JVET-AH0121 item 5) check/fix the ChromaFormatIdc derivation for the use of the colour transform information SEI message as a part of a processing chain in a similar manner as done for the film grain characteristics SEI message in JVET-AG2027 (to carefully check and communicate offline for correctness, e.g., w.r.t. JVET-AH0047).
* (JVET-AH0121 item 6) check/fix the interface variable derivation for the use of the NNPFC and NNPFA SEI messages as a part of a processing chain in a similar manner as done for the film grain characteristics SEI message in JVET-AG2027 (to carefully check and communicate offline for correctness, e.g., w.r.t. JVET-AH0047).

To be integrated:

* + General editorial improvementsJVET-AJ0185 Editorial updates for NNPFC, SPO and TDI SEI messages
    - Item 2

**Changes that have been integrated:**

*A description of the contributions reflected in this document relative to JVET-AI2006:*

* General editorial improvements
  + JVET-AJ0050 Editorial changes to VSEI draft text
  + JVET-AJ0065 Comments On VSEI Version 4
  + JVET-AJ0070 Editorial and technical problems in the draft of VSEI v4 and their solutions
  + JVET-AJ0120 Miscellaneous modifications for SEI messages in the VSEI draft
  + JVET-AJ0185 Editorial updates for NNPFC, SPO and TDI SEI messages
    - Item 1, 3, 4, 5, and 6.
* SEI Processing Order SEI
  + JVET-AJ0045 Handling of processing chains - operations
  + JVET-AJ0046 Handling of processing chains - constraints and special cases
  + JVET-AJ0047 Merging NNPF filtering process into processing chain handling
  + JVET-AJ0048 On SeiProcessingOrderSeiList and SpoProcessSeiList
  + JVET-AJ0128 Adding functionality to po\_processing\_degree\_flag[ i ] of the SPO SEI message
  + JVET-AJ0129 Miscellaneous comments on the SPO SEI message
  + JVET-AJ0105 On signalling of complexity information in SEI processing order SEI message
* Neural Network Post Filter
  + JVET-AJ0049 NNPFC and NNPFA interface text in VVC
  + JVET-AJ0073 A clean-up for the spatial resampling NNPF design
  + JVET-AJ0058 Generalizing text prompt usage in the NNPFC SEI message
  + JVET-AJ0131 On text prompts for NNPF
  + JVET-AJ0234 On Spatial Extrapolation for NNPF
  + JVET-AJ0104 On signalling of prompt in NNFPA SEI message
  + JVET-AJ0114 On signalling input pictures shift in NNPFA SEI message
* Encoder optimization information SEI
  + JVET-AJ0063 On Encoder Optimization Information SEI
  + JVET-AJ0183 Signaling source and added pictures in EOI SEI message
* Source picture timing SEI
  + JVET-AJ0308 Leading pictures and the SPTI SEI message
  + JVET-AJ0252 On SPTI SEI message
* Text description information SEI
  + JVET-AJ0184 On signaling of cancellation, persistency and id in TDI SEI messages
  + JVET-AJ0241 Text description SEI purpose for encoder description
* Generative Face Video and Generative Face Video Enhancement SEI message
  + JVET-AJ0207 Move GFV and GFVE to VSEI draft
  + JVET-AJ0051 On the GFV SEI message
  + JVET-AJ0101 Miscellaneous modifications for SEI messages in the TuC for futhre exensions of VSEI (GFV aspects)
  + JVET-AJ0108 On signalling of instance count in generative face video SEI message
  + JVET-AJ0135 Refined Methodology for Pupil Position SEI Message for Generaive Face Video
* Digitally signed content SEIs
  + Move to VSEI draft (see notes for JVET-AJ0151)
* Neural Network Post Filter
  + JVET-AJ0142 On potential conflicts between NNPFs and other post-processings

**Related input contributions**

Input contributions were noted as relevant to the work of this ad hoc group:

* (none)

**Remaining VVC bug tickets**

Closed since JVET-AJ0002 was reported:

* [#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.
* [#1637](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1637) The wording of H.266 B.3 seems to require byte streams to start with at least three zeros.

Carried over (strikethough means addressed in the indicated errata report and can be closed):

* [#1609](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1609) NoBackwardPredFlag derivation ambiguity (Addressed in JVET-AI1004)
* [#1617](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1617) Not initialized NumCtusInSlice[0] to 0. (Addressed in JVET-AI1004)
* [#1624](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1624) Incorrect indexing in computation of motion vector offset. (Addressed in JVET-AI1004)
* [#1627](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1627) "Decoding process for palette mode" does not say what to do with output samples. (Discussed in JVET-AI1004)
* [#1628](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1628) Derivation of ModeTypeCondition should say "one or more". (Addressed in JVET-AI1004)
* [#1629](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1629) mtt\_split\_cu\_vertical\_flag context uses undefined variable chType. (Discussed in JVET-AI1004)
* [#1630](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1630) Missing equations for applying AmvrShift. (Discussed in JVET-AI1004)
* [#1631](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1631) Should "Motion vector storing process for geometric partitioning mode" store HpelIfIdx? (addressed in JVET-AI1004)
* [#1632](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1632) Incorrect indexing used for choosing matrix intra sample prediction. (Addressed in JVET-AI1004)
* [#1634](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1634) Matrices QStateTransTable, levelScale, AlfFixFiltCoeff, AlfClassToFiltMap are incorrectly transposed. (Addressed in JVET-AI1004)
* [#1635](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1635) Incorrect inference for tu\_y\_coded\_flag. (Discussed in JVET-AI1004)
* [#1644](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1644) Correction of wrong reference on VVC Annex C.4 regarding DpbOutputInterval constraint. Applicable to and corrected in HEVC as well. (Addressed in JVET-AI1004)

New (since JVET-AJ0002 was reported)

* [#1650](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1650) Incorrect indexing for the h location component in ALF process. (Discussed in JVET-AJ1004)
* [#1651](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1651) References to various non-existent syntax elements relating to sps\_partition\_constraints\_override\_enabled\_flag.

**Remaining HEVC bug tickets**

Closed since JVET-AJ0002 was reported:

* (none)

Carried over:

* [#1427](https://hevc.hhi.fraunhofer.de/trac/hevc/ticket/1427) (8-155) and (8-157) do not seem to be used
* [#1491](https://hevc.hhi.fraunhofer.de/trac/hevc/ticket/1491) Duplicate invocation of 9.3.4.3 arithmetic decoding process
* [#1498](https://hevc.hhi.fraunhofer.de/trac/hevc/ticket/1498) Typos in the Table 9-43
* [#1500](https://hevc.hhi.fraunhofer.de/trac/hevc/ticket/1500) Typo in equation (8-69),(8-70)
* [#1504](https://hevc.hhi.fraunhofer.de/trac/hevc/ticket/1504) Small typos in profile\_tier\_level syntax in tabular form (7.3.3)
* [#1505](https://hevc.hhi.fraunhofer.de/trac/hevc/ticket/1505) Misleading bitstream requirement related to EOB NAL unit
* [#1507](https://hevc.hhi.fraunhofer.de/trac/hevc/ticket/1507) Duplicate row entries for CU QP delta syntax elements in Table 9-48
* [#1520](https://hevc.hhi.fraunhofer.de/trac/hevc/ticket/1520) Some smaller errors in the multiview spec
* [#1522](https://hevc.hhi.fraunhofer.de/trac/hevc/ticket/1522) Offset issue in clause 8.5.4.3
* [#1644](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1644) Correction of wrong reference on VVC Annex C.4 regarding DpbOutputInterval constraint. Applicable to and corrected in HEVC as well. (Addressed in JVET-AI1004)

New (since JVET-AI0002 was reported):

* (none)

**Recommendations**

The AHG recommends to:

* Approve JVET-AJ1004, JVET-AJ2005 and JVET-AJ2006 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.

It was agreed to do fixes to HEVC in the next editions (version 11 in ITU targeted for October), and in the ISO amendment on top of edition 6 requested in last meeting (other items are related to new SEI messages in VSEI). It was agreed to issue CDAM at this meeting (short editing period) and DAM in April. This may require leaving out interfaces with some VSEI parts, which could be added at DAM.

This would also have the advantage of keeping timing of all three standards (HEVC, VVC, VSEI) aligned.

Edition 6 FDIS at this moment is aligned with ITU H.265 version 10 that was approved after the April meeting. Both have some errors in the multi-view part.

[JVET-AK0003](https://jvet-experts.org/doc_end_user/current_document.php?id=15170) 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 23.6](https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/-/releases/VTM-23.6) (Nov. 2024)
* [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.

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 version 23.5 was tagged on Nov. 8, 2024. VTM 23.7 is expected during the 37th JVET meeting.

VTM 23.5 was tagged on Nov. 8, 2024. Changes include:

* Fix Scalable Nesting SEI parsing
* Fix Multilayer decoding with wraparound
* Fix invalid vector indexing
* Explicit configuration of inter-layer reference pictures
* FIX input file with rescaling (multilayer scalable) when using "SourceScalingRatioHor/Ver" encoder option
* Fix GCC 14 warning: use of uninitialized variable
* Fix uninitialized and cleaning
* JVET-AH2006: Text Description Information SEI message
* Make RdCost members non-`static` to use correct weighting tables for each layer in multilayer encoding
* JVET-AI0073: On the handling of a processing chain specified by an SPO SEI message
* Override SPS using --UpscaledOutputWidth and --UpscaledOutputHeight when upscaling output...
* JVET-AI0071: Indication of NNPFs and processing chains for human viewing and/or machine consumption
* JVET-AI0180: Adding original source picture dimensions to EOI SEI
* Move macro definition into appropriate location
* Add NNPFC SEI message support of spatial extrapolation in JVET-AH0174 and JVET-AI0202.
* add MS-SSIM2 for multilayer with reference layer (scalable)
* JVET-AI0153: Object Mask Information SEI
* optimization of RdCost::getWeightedMSE() in case of constant weight
* Workaround for VLCReader::xReadString for Visual C++
* JVET-AI0059: Define persistence scope for text description information SEI to be purpose specific
* JVET-AI0061-PROPOSAL2: Ensure spatial extrapolation works correctly (Option 1)
* Fix #1643 on different pps.conformanceWindows when using --UpscaledOutput=1 with RPR
* Clean up unregistered user data SEI
* Fix #1647 on adding HMVP candidates to AMVP list (multilayer with >2 layers)
* JVET-AG0322: Modality Information SEI
* JVET AI0061 Proposal 1: signal a NNPFC prompt for spatial extrapolation
* Fix #1646 on the condition of using tempMV for SbTMVP
* Fix #1645 on sbtmvp enabling condition in CU-level in case of multilayer profile.
* Fix #1649 on rprConstraintsActiveFlag
* Fix a regression in RPL parsing in Multilayer case
* Fix #1648 on sps\_chroma\_horizontal/vertical\_collocated\_flag usage in case of multilayer
* Define default value for Picture::m\_numSubPic
* Initialize m\_numSubPic for unavailable pictures.

VTM 23.6 was tagged on Nov. 11, 2024. Changes only include removal of macros from previous meeting cycles.

VTM 23.7 is expected to be tagged during the 37th JVET meeting. Changes are expected to include:

* Take bit depth into account when applying temporal prefilter
* JVET-AJ0067: AHG10: Suggested updates of settings for scalable VVC coding
* JVET-AJ0225: Changes to RDO for enhancement layer
* JVET-AJ0129: Add more SEI types to the list of allowed SEI types in SPO SEI
* JVET-AJ0241: Text description SEI purpose for encoder description
* Fix GCC 14 warning
* JVET-AJ0207: GFV SEI message
* JVET-AJ0063: Proposal 1,2,4,5
* Port ET with millisecond precision (from ECM MR 803)
* JVET-AJ0151: Digitally Signed Content SEI messages
* fix ScalingList for multilayer, when used for only one layer that is not the...
* JVET-AJ0131: Add nnpfc\_inband\_prompt\_flag to gate NNPF text prompt
* Fix for decoder when DSC SEI is not present
* JVET-AJ0128: Adding functionality to po\_processing\_degree\_flag[i] of the SPO SEI message
* JVET-AJ0234: Allow for patch-wise operation when using spatial extrapolation

In addition, there are pending MRs on the fixes of multi-layer coding. It may be good to also have some tests to verify the fixes.

***CTC Performance***

VTM 23.6 shows identical performance compared to VTM 23.4 in SDR CTC.

For the HDR CTCs, coding performance of VTM 23.6 and VTM 23.4 are identical. Encoding/decoding run times are very close.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | |
|  | **Over VTM-23.5** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | 0.00% | 0.00% | 0.00% | 100% | 98% |
| Class A2 | 0.00% | 0.00% | 0.00% | 100% | 98% |
| Class B | 0.00% | 0.00% | 0.00% | 100% | 98% |
| Class C | 0.00% | 0.00% | 0.00% | 100% | 103% |
| Class E | 0.00% | 0.00% | 0.00% | 100% | 100% |
| **Overall** | 0.00% | 0.00% | 0.00% | 100% | 99% |
| Class D | 0.00% | 0.00% | 0.00% | 100% | 100% |
| Class F | 0.00% | 0.00% | 0.00% | 103% | 100% |
| Class TGM | #VALUE! | #VALUE! | #VALUE! | #NUM! | #NUM! |
|  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | |
|  | **Over VTM-23.5** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | 0.00% | 0.00% | 0.00% | 99% | 99% |
| Class A2 | 0.00% | 0.00% | 0.00% | 99% | 100% |
| Class B | 0.00% | 0.00% | 0.00% | 99% | 101% |
| Class C | 0.00% | 0.00% | 0.00% | 99% | 101% |
| Class E |  |  |  |  |  |
| **Overall** | 0.00% | 0.00% | 0.00% | 99% | 100% |
| Class D | 0.00% | 0.00% | 0.00% | 99% | 102% |
| Class F | 0.00% | 0.00% | 0.00% | 100% | 103% |
| Class TGM | #VALUE! | #VALUE! | #VALUE! | #NUM! | #NUM! |
|  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | |
|  | **Over VTM-23.5** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | 0.00% | 0.00% | 0.00% | 99% | 99% |
| Class C | 0.00% | 0.00% | 0.00% | 100% | 100% |
| Class E | 0.00% | 0.00% | 0.00% | 99% | 102% |
| **Overall** | 0.00% | 0.00% | 0.00% | 99% | 100% |
| Class D | 0.00% | 0.00% | 0.00% | 99% | 100% |
| Class F | 0.00% | 0.00% | 0.00% | 100% | 101% |
| Class TGM | #VALUE! | #VALUE! | #VALUE! | #NUM! | #NUM! |
|  |  |  |  |  |  |
|  | **Low delay P Main10** | | | | |
|  | **Over VTM-23.5** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | 0.00% | 0.00% | 0.00% | 100% | 100% |
| Class C | 0.00% | 0.00% | 0.00% | 100% | 102% |
| Class E | 0.00% | 0.00% | 0.00% | 100% | 103% |
| **Overall** | 0.00% | 0.00% | 0.00% | 100% | 101% |
| Class D | 0.00% | 0.00% | 0.00% | 100% | 100% |
| Class F | 0.00% | 0.00% | 0.00% | 100% | 102% |
| Class TGM | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! |

With a bug fix of temporal filtering (merged shortly after VTM 23.6), VTM shows good improvement with option --InternalBitDepth=12. Note that the runtime is not reliable.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | |
|  | **Over VTM-23.6-d25697cc5** | | | | | |
|  | Y | U | V | EncT | DecT | mPeakR |
| Class A1 | -0.52% | -0.84% | -0.61% | Not reliable |  | 101% |
| Class A2 | -0.37% | 0.39% | 0.68% |  |  | 100% |
| Class B | -0.21% | 0.29% | 0.22% |  |  | 100% |
| Class C | -0.13% | 0.73% | 0.67% |  |  | 100% |
| Class E | -0.24% | 0.29% | -0.02% |  |  | 100% |
| **Overall** | -0.28% | 0.22% | 0.22% |  |  | 100% |
| Class D | 0.08% | 0.95% | 0.73% |  |  | 100% |
| Class F | -0.14% | -0.03% | 0.19% |  |  | 100% |
| Class TGM | -0.34% | 0.21% | 0.19% |  |  | 100% |
|  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | |
|  | **Over VTM-23.6-d25697cc5** | | | | | |
|  | Y | U | V | EncT | DecT | mPeakR |
| Class A1 | -0.85% | -1.14% | -0.32% | Not reliable |  | 100% |
| Class A2 | -0.55% | -0.31% | -0.64% |  |  | 100% |
| Class B | -0.40% | 0.33% | 0.01% |  |  | 100% |
| Class C | -0.36% | 1.42% | 0.93% |  |  | 100% |
| Class E |  |  |  |  |  |  |
| **Overall (Ref)** | -0.51% | 0.20% | 0.06% |  |  | 100% |
| Class D | -0.19% | 0.74% | 0.41% |  |  | 100% |
| Class F | -0.47% | 0.17% | 0.30% |  |  | 100% |
| Class TGM | -0.46% | 0.25% | 0.33% |  |  | 100% |
|  |  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | | |
|  | **Over VTM-23.6-d25697cc5** | | | | | |
|  | Y | U | V | EncT | DecT | mPeakR |
| Class A1 |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |
| Class B | -0.52% | -0.80% | -1.01% | Not reliable |  | 105% |
| Class C | -0.35% | 0.57% | 0.35% |  |  | 99% |
| Class E | -0.64% | -0.75% | -0.67% |  |  | 100% |
| **Overall (Ref)** | -0.49% | -0.33% | -0.48% |  |  | 102% |
| Class D | -0.26% | 1.16% | 0.53% |  |  | 100% |
| Class F | -0.35% | -0.24% | -0.18% |  |  | 101% |
| Class TGM | -0.29% | 0.06% | 0.30% |  |  | 99% |
|  |  |  |  |  |  |  |
|  | **Low delay P Main10** | | | | | |
|  | **Over VTM-23.6-d25697cc5** | | | | | |
|  | Y | U | V | EncT | DecT | mPeakR |
| Class A1 |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |
| Class B | -0.44% | -0.86% | -1.32% | Not reliable |  | 89% |
| Class C | -0.31% | 0.68% | 0.19% |  |  | 101% |
| Class E | -0.90% | -0.79% | -1.16% |  |  | 100% |
| **Overall (Ref)** | -0.51% | -0.33% | -0.78% |  |  | 96% |
| Class D | -0.27% | 0.29% | -0.18% |  |  | 100% |
| Class F | -0.30% | -0.73% | 0.14% |  |  | 100% |
| Class TGM | -0.34% | 0.08% | 0.14% |  |  | 97% |

For the high bit depth CTCs, VTM 23.6 also shows no changes in coding performance compared to VTM 23.4.

***Issues in VTM affecting conformance***

The following issues in VTM master branch may affect conformance:

* Missing HLS features (see sections below)

***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.

***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-2.0 was tagged on July 17, 2024. Changes include:

* Remove unused variables
* JVET-AG0044: Copyright information SEI message
* JVET-AG0045: AI information SEI message
* JVET-AF0141: EXIF, JFIF and XMP metadata SEI messages

VTM-22.2-TuC-3.0 was tagged on August 2nd, 2024. Changes include:

* JVET-AH0162: Constituent Rectangles SEI
* JVET-AH0161: Packed Regions Info SEI
* JVET-AH0164: Quality metric SEI

VTM-22.2-TuC-4.0 was expected to be tagged during the 36th JVET meeting. The release was delayed due to pending merge request. It is now expected to be tagged during the 37th JVET meeting. The following merge requests were submitted:

* JVET-AI0099-SW: Implementation of adaptive film grain models in different regions (No change in TuC software. Experiment was run using external software. To be closed.)
* JVET-AI0195: GFV SEI message (comments pending)
* JVET-AI0182: Bitdepth range information SEI (review pending)
* Fix typos in constituent rectangles implementation (review pending)
* JVET-AI0194: GFV chroma key (based on JVET-AI0195 – review pending)
* JVET\_AI0184\_0189\_0190\_0191\_GFV\_SEI (based on JVET-AI0195 – review pending)
* JVET-AI0127: Digitally Signed Content SEI messages (review pending)
* JVET-AI0181: Support for Display Overlays Information (DOI) SEI (minor comments pending)

The MRs for JVET-AI0195, JVET-AI0194, JVET-AI0184 and JVET-AI0127 were closed after adoption of the related technologies into VTM and the merge requests for VTM are available.

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:

* 44 tickets for “HM”, (most of which were created before 2019; 1 created in last year),
* 1 ticket for “HM RExt” (created in 2020),
* 9 tickets for “HM SCC” (most of which were created before 2018),
* 1 ticket for “RExt Text” (created in 2015),
* 1 ticket for “SCC Text” (created in 2016),
* 8 tickets for text (6 created before 2021; 2 created in the last year),
* 2 tickets for encoder description (created in 2020 and 2014)

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 was merged:

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

The next release will also 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**

An updated version of VTM-based software guidelines is proposed in JVET-AJ0187.

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.

It was noted that no merge request was received on the new MV-HEVC profiles in combination with RExt tools, though there was an input report in the last meeting. This also has impact on development of conformance for these profiles.

It was commented that the origin of this 12-bit change was a normative ECM proposal, which also fixed the MCTF bug in ECM.

Coffee break until 11:35

[JVET-AK0004](https://jvet-experts.org/doc_end_user/current_document.php?id=15171) 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)]

**Activities**

***Verification tests for VVC multilayer coding***

The verification test plan has been updated to output JVET-AJ2021 of the previous JVET meeting.

***Subjective quality testing of FGC SEI message***

A report on subjective quality testing of FGC SEI message has been produced as output document JVET-AJ2037. The report includes the results and findings from expert viewing activities at the 35th and 36th JVET meetings.

The plan for subjective quality testing has been updated as output JVET-AJ2022 of the previous JVET meeting.

***Test sequences***

The server providing JVET test sequences has been updated. The new site is <https://vqa.lfb.rwth-aachen.de>. A mirror of this site is available at <https://datacloud.hhi.fraunhofer.de> with the same login credentials. The directory structure of the previous ftp server has been maintained. The test sequences used for CfP/CTC are available in directory “/ctc”.

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). Members of JVET may contact the JVET chair for login information. It is suggested to update the password for the site after each meeting with the Zoom password of that meeting.

**Related contributions**

|  |  |  |
| --- | --- | --- |
| [JVET-AK0066](https://jvet-experts.org/doc_end_user/current_document.php?id=15037) | AHG4/AHG17: 8K HDR sequences as a response to the call for new HDR test materials for future video coding | [Y. Li (CMG)](mailto:liyan@cctv.com), [Q. Zhang](mailto:zhangqian@abp2003.cn), [J. Ning (ABP)](mailto:ningjinhui@abp2003.cn) |
| [JVET-AK0282](https://jvet-experts.org/doc_end_user/current_document.php?id=15271) | AHG4/AHG17: Additional HDR sequences for non-CTC testing | [A. Filippov](mailto:alexey.filippov@tcl.com), [J. Konieczny](mailto:jacek.k@tcl.com), [Z. Zhi](mailto:zhou.zhi@tcl.com), [V. Rufitskiy](mailto:vasily.rufitskiy@tcl.com), [H. Qin](mailto:hongdong.qin@tcl.com), [T. Dong](mailto:tianyu.dong@tcl.com), [X. Tang (TCL)](mailto:xiaoyong.tang@tcl.com), [S. Shen](mailto:s.shen@nercdtv.org), [Y. Qin](mailto:yf.qin@nercdtv.org), [Y. Wang](mailto:y.wang@nercdtv.org), [X. Wang](mailto:xs.wang@nercdtv.org), [Y. Guan (NERC-DTV)](mailto:yf.guan@nercdtv.org) |

**Recommendations**

The AHG recommends:

* To review the input contributions related to AHG4 during the meeting.
* To progress the work in AHG17 towards setting up a test set suitable for performance assessment of video compression technology beyond the JVET CTC.
* To continue to discuss and to update the non-finalized categories of the verification test plan for multilayer VVC and the subjective quality test plan for FGC, including those which have not been addressed yet.
* To collect volunteers to conduct further verification tests and subjective quality tests.
* 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-AK0005](https://jvet-experts.org/doc_end_user/current_document.php?id=15172) JVET AHG report: Conformance testing (AHG5) [I. Moccagatta (chair), F. Bossen, T. Ikai, S. Iwamura, H.-J. Jhu, K. Kawamura, P. de Lagrange, S. Paluri, K. Sühring, Y. Yu (vice chairs)]

1. **Activities**

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 36th and 37th 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, FDIS ballot opened 2024-04-08, FDIS ballot closed 2024-06-03, published 2024-07-04
  + H.266.1 V2 Consent 2023-07, Last Call began 2023-08-16, Approved 2023-09-13, pre-published 2023-10-06, published 2023-10-19

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
  + Three streams have been re-generated
  + No changes between 36th and 37th 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 36th and 37th meeting.
* additional conformance bitstreams for VVC Multilayer:
  + 3 bitstream categories have been identified
  + At least one bitstream has been submitted in each identified category
  + 7 total bitstreams have been provided, cross-checked, and made available
  + No changes between 36th and 37th meeting.
* conformance bitstreams for new HEVC Multiview profiles
  + 2 HEVC Multiview Extended 10 bitstreams have been provided and cross-checked.
  + 2 HEVC Multiview Extended bitstreams have been provided and cross-checked.
  + No changes between 36th and 37th meeting.

1. **Activities and Discussion**

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

VVC activities:

The new version of RAP\_B (RAP\_B\_2) reported at the 36th meeting is available at ​<https://www.itu.int/wftp3/av-arch/jvet-site/bitstream_exchange/VVC2ndEd_regenerated/under_test>. No change in the other bitstreams and/or packages.

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:

No change in bitstreams and/or packages.

HEVC Multiview supporting extended bit depth activities:

No change in bitstreams and/or packages, the 4 packages are available at <https://www.itu.int/wftp3/av-arch/jvet-site/bitstream_exchange/HEVCMultiview/under_test/>.

No update on the implementation of the HEVC Multiview monochrome profiles (Multiview Monochrome, Multiview Monochrome 10, Multiview Monochrome 12, and Multiview Monochrome 16 profiles) in JVET-AH1006 and in WG5 N0281. No update on conformance streams for the new HEVC Multiview Main 10 profile in JVET-AJ0213 added at the 36th meeting and in output document AJ1006.

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**

No contributions.

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 was previously necessary to set the minimum TLS level to 1.0. Since around January 2024, this configuration issue went away, and ITU disabled the use of TLS 1.0 and 1.1 on its servers.

1. **Recommendations**

The AHG recommends the following:

* 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 and the new HEVC Multiview Main 10 profile
* Prepare for a new edition of VVC conformance with the new multi-layer streams and the corrections of the existing streams.

It was reported that software for generating standalone bitstreams for profiles with bitdepth >10 should work OK, but was not merged yet (see AHG3 report). This should be the next step. For bitstreams using combinations of layers with different bit depth, more updates of software are necessary.

On the last bullet item, and considering that the document JVETAI2028 is stable, it was concluded that a new edition of VVC conformance should be started (request in ISO at the current meeting, CD in April, DIS and ITU consent in October).

[JVET-AK0006](https://jvet-experts.org/doc_end_user/current_document.php?id=15173) 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-14.1:

Fix issue #74 TM tools off for JVET-AI0103 [MR 726]

Fix Clang 16.0.6 compiling issues [MR 724]

Set GeoInerIbc in config file explicitly to enable it for SCC [MR 725]

Fix: compilation when macro JVET\_AI0136\_ADAPTIVE\_DUAL\_TREE is set to 0[MR 726]

Fix: wrong memory access in case of DebugBitstream feature usage [MR 728]

Correction related to adaptive dual treee (JVET\_AI0136) to make feature DebugBitstream work [MR 727]

Fix DebugBitstream issue caused by copying values for adaptive clipping [MR 730]

Fix: RPR encoder crash when JVET\_AH0314\_LIC\_INHERITANCE\_FOR\_MRG=1 [MR 729]

Fix useless copies of TempCtx objects to Ctx dest [MR 731]

Fix RPR w.r.t. temporal prediction of partitioning information [MR 734]

Fix: 2 valgrind erros at encoder side [MR 735]

Fix DebugBitstream issue caused by PDP [MR 733]

Fix in addBIOAvgN\_SSE for non-AVX2 with JVET\_AI0046\_HIGH\_PRECISION\_BDOF\_SAMPLE [MR 736]

Add check whether writing to bitstream succeeded [MR 738]

Fix a bug of tracing blocks stats on decoder [MR 740]

The following changes were integrated into ECM-14.2:

Fix: always build hash info for I slices [MR 801]

The following changes were integrated into ECM-15.0:

JVET-AJ0096: SATD-based reordering for Intra/Inter (Test3.10a/b) [MR 752]

JVET-AJ0081: Chroma TMRL (Test 2.18) [MR 742]

JVET-AJ0161: OBMC extension with intra prediction (Test 3.3b) [MR 748, 798]

JVET-AJ0274: GPM-affine with TM and Regression-based GPM with TM (Test 3.12) [MR 753]

JVET-AJ0172: Unified reference area of IBC and IntraTMP (Test 2.16) [MR 743]

JVET-AJ0061: TIMD Merge Mode (Test 2.5) [MR 750, 769]

JVET-AJ0226: Using CABAC cost for MTT split modes selection [MR 758]

JVET-AJ0268: Combination of DIMD with 2x2 edge operator applied to small blocks (Tests 2.11) and adaptive HoG for DIMD (Test 2.10a) (Test 2.13a) [MR 760, 763]

JVET-AJ0126: Enhanced inter AMVP (Test 3.2b) [MR 751]

JVET-AJ0175: NSPT for non-regular modes (Test 4.1) [MR 749]

JVET-AJ0057: High-level intra tool control (Test 2.14) [MR 744]

JVET-AJ0188: Coding Info based Classification for ALF (Test 5.4) [MR 754, 781]

JVET-AJ0082: Multi-model EIP (Test 2.19a) [MR 745, 786]

JVET-AJ0107: GPM mode extension (Test 3.4d) [MR 741, 776, 783]

JVET-AJ0097: BDOF for lowdelay pictures (Test 3.5) [MR 755]

JVET-AJ0085: Subblock merge mode extension (Test 3.7a) [MR 770]

JVET-AJ0112: Regression-based SGPM blending (Test 2.15a\*) [MR 762, 785]

JVET-AJ0158: Sub-block inter mode improvement (Test 3.9d) [MR 746, 775, 778, 795]

JVET-AJ0260: SBT corner mode (Test 4.3a) [MR 771]

JVET-AJ0249: Neural network-based intra prediction with DIMD mode derivation (Test 2.20/21) [MR 764, 791]

JVET-AJ0146: Additional TIMD mode with different cost metric (Test 2.6) [MR 773, 780]

JVET-AJ0257: Improved Implicit MTS (Test 4.2) [MR 779]

JVET-AJ0079: Disabling timd-isp and timd-mrl combinations in a non-normative way (Test 2.7b) [MR 761]

JVET-AJ0237: 12-bit internal bit depth modifications for ECM [MR 757]

DebugBitstream usage without –DebugPOC [MR 772]

Fix ClassF.cfg lowdelay [MR 765, 774]

Fix for resolving enc/dec mismatch from JVET-AG0098(getAmvpSbTMVP) [MR 777]

Fix memory leak in allocated g\_pdpFiltersMip [MR 782]

Fix the compiling issues of clang 16.0.6 [MR 784]

Fix for bvFilter copying of FIBC [MR 788]

Fix for IPM storage of SGPM [MR 789]

Fix for SeparateTree [MR 790]

Fix KEEP\_PRED\_AND\_RESI\_SIGNALS [MR 792]

The following changes were integrated into VTM-11ecm anchor:

JVET-AJ0226: Using CABAC cost for MTT split modes selection [MR 768]

Update class F low delay config [MR 793]

ECM-14.1 was tagged on November 7, 2024.

ECM-15.0, ECM-14.2, and VTM-11ecm15.0 were tagged on December 11, 2024.

***CTC Performance***

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

ECM-15.0 performance over ECM-14.0 anchor is summarized in the tables below.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | | | |
|  | **Over ECM-14.0** | | | | | | |
|  | Y | U | V | EncT | DecT | EncVmPeak | DecVmPeak |
| Class A1 | -1.24% | -1.09% | -1.26% | 107.0% | 105.6% | 117.8% | 109.1% |
| Class A2 | -0.72% | -0.69% | -0.68% | 107.7% | 110.4% | 117.0% | 109.4% |
| Class B | -0.82% | -0.60% | -0.57% | 108.4% | 115.4% | 110.7% | 106.1% |
| Class C | -0.64% | -0.64% | -0.72% | 103.8% | 118.7% | 105.6% | 104.2% |
| Class E | -1.23% | -1.33% | -1.14% | 108.2% | 116.6% | 105.0% | 104.7% |
| **Overall** | -0.90% | -0.83% | -0.83% | 107.0% | 113.8% | 110.7% | 106.5% |
| Class D | -0.64% | -0.56% | -0.76% | 104.1% | 126.0% | 104.1% | 103.5% |
| Class F | -0.93% | -0.97% | -0.89% | 106.5% | 115.9% | 102.6% | 104.6% |
| Class TGM | -1.54% | -1.73% | -1.67% | 106.6% | 115.6% | 101.0% | 105.3% |
|  |  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | | |
|  | **Over ECM-14.0** | | | | | | |
|  | Y | U | V | EncT | DecT | EncVmPeak | DecVmPeak |
| Class A1 | -0.73% | -0.89% | -0.83% | 104.3% | 103.5% | 101.4% | 104.3% |
| Class A2 | -0.84% | -0.66% | -0.78% | 103.7% | 103.4% | 101.0% | 104.5% |
| Class B | -0.74% | -0.74% | -0.66% | 104.1% | 106.9% | 101.2% | 103.6% |
| Class C | -0.89% | -0.69% | -0.50% | 106.3% | 110.0% | 103.0% | 102.8% |
| Class E |  |  |  |  |  |  |  |
| **Overall** | -0.79% | -0.74% | -0.68% | 104.7% | 106.3% | 101.7% | 103.7% |
| Class D | -0.85% | -0.43% | -0.72% | 106.4% | 110.7% | 104.2% | 103.1% |
| Class F | -0.88% | -0.92% | -0.93% | 106.4% | 115.5% | 101.7% | 103.3% |
| Class TGM | -1.23% | -1.13% | -1.18% | 106.3% | 111.4% | 100.4% | 103.6% |
|  |  |  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | | | |
|  | **Over ECM-14.0** | | | | | | |
|  | Y | U | V | EncT | DecT | EncVmPeak | DecVmPeak |
| Class A1 |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |
| Class B | -0.79% | -0.57% | -0.62% | 110.5% | 109.0% | 102.1% | 102.7% |
| Class C | -1.05% | -0.12% | -0.98% | 111.2% | 113.8% | 104.1% | 102.2% |
| Class E | -1.14% | 0.61% | -1.08% | 113.2% | 107.2% | 101.4% | 102.0% |
| **Overall** | -0.96% | -0.13% | -0.86% | 111.4% | 110.1% | 102.6% | 102.4% |
| Class D | -1.22% | -0.03% | -0.22% | 112.9% | 111.8% | 104.2% | 103.2% |
| Class F | -1.10% | -1.35% | -0.21% | 108.2% | 110.7% | 102.4% | 103.2% |
| Class TGM | -1.19% | -1.15% | -1.14% | 106.7% | 116.1% | 100.7% | 103.5% |
|  |  |  |  |  |  |  |  |
|  | **Low delay P Main 10** | | | | | | |
|  | **Over ECM-14.0** | | | | | | |
|  | Y | U | V | EncT | DecT | EncVmPeak | DecVmPeak |
| Class A1 |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |
| Class B | -0.73% | -0.49% | -0.70% | 107.6% | 109.4% | 101.4% | 105.1% |
| Class C | -0.82% | -0.46% | -0.57% | 105.7% | 106.3% | 106.8% | 104.3% |
| Class E | -1.01% | -2.31% | -1.10% | 106.6% | 105.0% | 102.5% | 104.7% |
| **Overall** | -0.83% | -0.93% | -0.76% | 106.7% | 107.2% | 103.4% | 104.7% |
| Class D | -1.32% | -0.73% | -1.17% | 104.9% | 100.4% | 105.4% | 103.5% |
| Class F | -0.92% | -0.66% | -0.94% | 106.3% | 114.3% | 102.0% | 104.6% |
| Class TGM | -0.94% | -0.88% | -1.15% | 107.0% | 113.0% | 102.3% | 104.6% |

The below tables show ECM-15.0 performance comparing to VTM-11.0ecm15.0 anchor.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | | | |
|  | **Over VTM-11.0ecm15.0** | | | | | | |
|  | Y | U | V | EncT | DecT | EncVmPeak | DecVmPeak |
| Class A1 | -14.31% | -15.16% | -26.55% | 1136.4% | 528.0% |  |  |
| Class A2 | -20.82% | -23.47% | -27.99% | 1131.2% | 568.7% |  |  |
| Class B | -14.38% | -21.88% | -19.73% | 1052.3% | 601.4% |  |  |
| Class C | -14.30% | -11.21% | -12.34% | 1007.8% | 551.3% |  |  |
| Class E | -18.63% | -21.83% | -20.07% | 995.7% | 623.0% |  |  |
| **Overall** | -16.13% | -18.65% | -20.66% | 1058.7% | 575.3% |  |  |
| Class D | -12.16% | -8.19% | -8.92% | 990.4% | 585.2% |  |  |
| Class F | -29.93% | -33.46% | -33.65% | 744.5% | 672.5% |  |  |
| Class TGM | -43.12% | -48.77% | -48.02% | 576.4% | 704.0% |  |  |
|  |  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | | |
|  | **Over VTM-11.0ecm15.0** | | | | | | |
|  | Y | U | V | EncT | DecT | EncVmPeak | DecVmPeak |
| Class A1 | -26.92% | -23.37% | -35.46% | 1148.5% | 1017.6% |  |  |
| Class A2 | -30.10% | -33.29% | -38.74% | 1103.2% | 1216.1% |  |  |
| Class B | -24.46% | -31.49% | -28.80% | 933.8% | 1077.1% |  |  |
| Class C | -26.33% | -21.66% | -22.37% | 1003.4% | 1179.8% |  |  |
| Class E |  |  |  |  |  |  |  |
| **Overall** | -26.58% | -27.61% | -30.41% | 1025.7% | 1117.9% |  |  |
| Class D | -27.14% | -22.16% | -23.42% | 949.0% | 1295.2% |  |  |
| Class F | -32.70% | -34.86% | -35.64% | 870.5% | 827.1% |  |  |
| Class TGM | -42.33% | -47.83% | -47.63% | 735.6% | 656.0% |  |  |
|  |  |  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | | | |
|  | **Over VTM-11.0ecm15.0** | | | | | | |
|  | Y | U | V | EncT | DecT | EncVmPeak | DecVmPeak |
| Class A1 |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |
| Class B | -21.69% | -35.54% | -31.98% | 964.9% | 896.7% |  |  |
| Class C | -24.13% | -24.52% | -26.26% | 917.1% | 969.2% |  |  |
| Class E | -21.40% | -25.57% | -24.70% | 861.1% | 592.9% |  |  |
| **Overall** | -22.43% | -29.37% | -28.26% | 922.1% | 829.8% |  |  |
| Class D | -25.45% | -25.45% | -26.27% | 933.8% | 1110.6% |  |  |
| Class F | -30.25% | -38.19% | -37.94% | 824.3% | 729.5% |  |  |
| Class TGM | -40.56% | -50.17% | -50.03% | 740.4% | 622.4% |  |  |
|  |  |  |  |  |  |  |  |
|  | **Low delay P Main 10** | | | | | | |
|  | **Over VTM-11.0ecm15.0** | | | | | | |
|  | Y | U | V | EncT | DecT | EncVmPeak | DecVmPeak |
| Class A1 |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |
| Class B | -19.47% | -44.58% | -41.77% | 836.4% | 875.7% |  |  |
| Class C | -22.13% | -33.99% | -34.62% | 774.1% | 895.5% |  |  |
| Class E | -19.91% | -35.54% | -35.94% | 778.7% | 627.7% |  |  |
| **Overall** | -20.47% | -38.79% | -37.93% | 800.7% | 811.8% |  |  |
| Class D | -24.78% | -36.48% | -37.19% | 783.5% | 1014.5% |  |  |
| Class F | -28.36% | -43.60% | -44.08% | 815.3% | 742.8% |  |  |
| Class TGM | -38.65% | -52.57% | -52.14% | 790.4% | 613.0% |  |  |

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) | 9 | 15 |  |
| class B | 4 | 6 | 5 |
| class C | 2 | 3 | 2 |
| class D | 2 | 2 | 2 |
| class E | 3 |  | 3 |
| class F | 4 | 6 | 5 |
| class TGM | 3 | 8 | 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.

[JVET-AK0007](https://jvet-experts.org/doc_end_user/current_document.php?id=15174) JVET AHG report: ECM tool assessment (AHG7) [X. Li (chair), L.-F. Chen, Z. Deng, J. Gan, E. François, R. Ishimoto, H.-J. Jhu, X. Li, J. Pardo, H. Wang (vice chairs)]

1. **AHG7 Conference call on 12/11/2024**

An AHG7 conference call was held on 12/11/2024. Two agenda items were discussed: 1) ECM software implementation on tool controls for the group 5; 2) Complexity analysis of the tools in the group 5. Most items were fully discussed during the call. Several follow-up items were offline addressed. The details including the complexity assessment may be found in the conference call report JVET-AK0048.

1. **Group off tests**

***Test settings and crosschecking***

The following five groups of tools were defined. Among them the group 5 was newly introduced for tools with large memory access during the 36th JVET meeting. However, the implementation of the control flags for the group 5 have not been completed before the testing. Therefore, the first four groups below were tested by AHG7 during this meeting cycle. There are input documents on tools of group 5.

* 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
* Group 5: Tools with large memory access

The testers and crosscheckers are summarized in the table below.

|  |  |  |
| --- | --- | --- |
| Tests | Testers | Crosscheckers |
| Group 1 off | Johan Pardo ([johan.esprit.pardo@huawei.com](mailto:johan.esprit.pardo@huawei.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 | Ryo Ishimoto ([ishimoto.ryo@mail.sharp](mailto:ishimoto.ryo@mail.sharp)) | 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) |

ECM-15.0 and a commit before it were used in the AHG7 tool off tests, as summarized in the table below. The cfg files used are included in the ECM software package.

|  |  |  |
| --- | --- | --- |
| Tests | Testers | Crosscheckers |
| Group 1 off | ECM-15.0 | ECM-15.0 |
| Group 2 off | ECM-15.0 | ECM-15.0 |
| Group 3 off | ECM-15.0 | ECM-15.0 |
| Group 4 off | ECM-15.0 | ECM-15.0pre-1e54f849f |
| Group 1-4 off | ECM-15.0 | ECM-15 |

***Group 1 off***

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

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | | | | |
|  | **Over ECM-15 (698389d6)** | | | | | | **Over VTM-11ecm15 (0113b3c0)** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 | 0.00% | 0.00% | 0.00% | 102.0% | 100.6% | 100.0% | -14.32% | -15.16% | -26.55% | 2101.4% | 1153.2% |
| Class A2 | 0.00% | 0.00% | 0.00% | 101.6% | 100.0% | 100.0% | -20.82% | -23.47% | -27.99% | 2044.5% | 1190.0% |
| Class B | 0.00% | 0.00% | 0.00% | 101.8% | 101.6% | 100.9% | -14.39% | -21.89% | -19.73% | 1894.9% | 1240.7% |
| Class C | -0.01% | -0.01% | -0.01% | 101.1% | 102.1% | 99.1% | -14.31% | -11.22% | -12.35% | 1865.3% | 1166.0% |
| Class E | -0.02% | -0.02% | -0.02% | 101.7% | 101.9% | 100.0% | -18.65% | -21.84% | -20.08% | 1787.6% | 1349.0% |
| **Overall** | -0.01% | -0.01% | -0.01% | 101.6% | 101.3% | 100.1% | -16.14% | -18.65% | -20.66% | 1926.8% | 1217.3% |
| Class D | -0.04% | -0.04% | -0.04% | 101.2% | 99.6% | 99.9% | -12.20% | -8.23% | -8.96% | 1828.5% | 1214.9% |
| Class F | -0.02% | -0.02% | -0.02% | 102.8% | 103.0% | 99.7% | -29.94% | -33.47% | -33.66% | 1286.5% | 1255.4% |
| Class TGM | -0.01% | 0.00% | 0.00% | 101.8% | 101.6% | 102.4% | -43.12% | -48.77% | -48.02% | 1055.6% | 1249.2% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | | | | | | |
|  | **Over ECM-15 (698389d6)** | | | | | | **Over VTM-11ecm15 (0113b3c0)** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 | 4.94% | 4.49% | 4.84% | 88.5% | 73.8% | 100.1% | -23.28% | -20.09% | -32.18% | 1816.9% | 1704.0% |
| Class A2 | 5.82% | 5.67% | 5.74% | 86.6% | 66.0% | 99% | -26.04% | -29.56% | -35.21% | 1761.3% | 1810.5% |
| Class B | 4.63% | 4.57% | 4.70% | 80.6% | 67.4% | 99.9% | -20.93% | -28.34% | -25.41% | 1414.1% | 1634.1% |
| Class C | 4.90% | 4.95% | 5.24% | 77.1% | 63.5% | 100.0% | -22.67% | -17.78% | -18.29% | 1419.1% | 1635.0% |
| Class E |  |  |  |  |  |  |  |  |  |  |  |
| **Overall (Ref)** | 5.00% | 4.87% | 5.08% | 82.3% | 67.3% | 100% | -22.88% | -24.12% | -26.83% | 1555.0% | 1682.2% |
| Class D | 3.97% | 3.95% | 4.25% | 75.5% | 60.1% | 100.0% | -24.19% | -19.04% | -20.09% | 1320.3% | 1738.3% |
| Class F | 3.71% | 3.69% | 3.80% | 88.2% | 75.6% | 99.9% | -30.11% | -32.34% | -33.09% | 1297.2% | 1177.9% |
| Class TGM | 4.63% | 4.48% | 4.53% | 89.2% | 80.5% | 100.3% | -39.67% | -45.45% | -45.21% | 1164.4% | 1130.0% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | | | | | | | |
|  | **Over ECM-15 (698389d6)** | | | | | | **Over VTM-11ecm15 (0113b3c0)** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |  |  |  |  |
| Class B | 5.75% | 6.11% | 6.04% | 70.6% | 59.0% | 100% | -17.15% | -31.62% | -28.02% | 1248.8% | 1243.1% |
| Class C | 5.84% | 6.03% | 6.53% | 70.8% | 54.7% | 100% | -19.69% | -19.98% | -21.41% | 1155.8% | 1202.9% |
| Class E | 5.94% | 5.77% | 6.12% | 77.4% | 71.6% | 100% | -16.73% | -21.34% | -20.15% | 1188.1% | 1129.0% |
| **Overall (Ref)** | 5.82% | 6.00% | 6.22% | 72.3% | 60.4% | 100% | -17.90% | -25.17% | -23.85% | 1201.9% | 1200.3% |
| Class D | 4.95% | 4.91% | 5.30% | 69.5% | 54.2% | 100% | -21.74% | -21.87% | -22.32% | 1182.9% | 1262.1% |
| Class F | 5.51% | 5.17% | 4.83% | 82.8% | 66.3% | 100% | -26.31% | -34.83% | -34.85% | 1183.0% | 998.5% |
| Class TGM | 6.61% | 6.23% | 6.34% | 88.8% | 73.5% | 99% | -36.60% | -46.90% | -46.69% | 1121.2% | 915.0% |

***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 offgroup2.cfg was used in addition to ECM CTC settings.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | | | | |
|  | **Over ECM15** | | | | | | **Over VTM11ecm15** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 | 0.16% | 0.27% | 0.63% | 94.1% | 95.5% | 100% | -14.17% | -14.94% | -26.04% | 1068.1% | 544.9% |
| Class A2 | 0.26% | 0.43% | 0.48% | 93.4% | 97.4% | 101% | -20.61% | -23.17% | -27.65% | 1040.6% | 577.0% |
| Class B | 0.20% | 0.44% | 0.45% | 92.7% | 96.2% | 100% | -14.21% | -21.55% | -19.35% | 984.1% | 613.7% |
| Class C | 0.16% | 0.38% | 0.41% | 91.5% | 94.1% | 99% | -14.16% | -10.88% | -11.99% | 963.0% | 548.7% |
| Class E | 0.32% | 0.58% | 0.86% | 92.9% | 96.3% | 98% | -18.37% | -21.38% | -19.37% | 930.5% | 640.9% |
| **Overall** | 0.22% | 0.42% | 0.54% | 92.8% | 95.8% | 99% | -15.95% | -18.32% | -20.21% | 992.8% | 585.1% |
| Class D | 0.12% | 0.23% | 0.44% | 91.1% | 92.1% | 99% | -12.05% | -7.99% | -8.52% | 944.6% | 585.0% |
| Class F | 0.62% | 0.59% | 0.56% | 92.8% | 98.1% | 100% | -29.52% | -33.08% | -33.23% | 576.9% | 625.9% |
| Class TGM | 0.41% | 0.54% | 0.48% | 93.3% | 99.9% | 100% | -42.88% | -48.49% | -47.76% | 496.4% | 664.0% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | | | | | | |
|  | **Over ECM15** | | | | | | **Over VTM11ecm15** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 | 2.47% | 1.18% | 1.96% | 87.7% | 98.4% | 100% | -25.12% | -22.38% | -34.22% | 984.7% | 1151.3% |
| Class A2 | 1.44% | 0.80% | 1.57% | 86.7% | 100.6% | 100% | -29.08% | -32.72% | -37.74% | 945.5% | 1417.7% |
| Class B | 1.46% | 1.07% | 1.16% | 86.6% | 97.2% | 99% | -23.34% | -30.72% | -27.88% | 812.7% | 1197.8% |
| Class C | 1.04% | 1.14% | 1.01% | 85.7% | 95.0% | 99% | -25.56% | -20.79% | -21.61% | 881.0% | 1253.2% |
| Class E |  |  |  |  |  |  |  |  |  |  |  |
| **Overall (Ref)** | 1.55% | 1.06% | 1.36% | 86.6% | 97.5% | 100% | -25.44% | -26.81% | -29.45% | 889.4% | 1244.0% |
| Class D | 0.50% | -0.11% | 0.25% | 84.7% | 94.9% | 99% | -26.77% | -22.26% | -23.24% | 850.3% | 1485.7% |
| Class F | 1.97% | 2.11% | 1.84% | 88.2% | 95.8% | 99% | -31.37% | -33.50% | -34.46% | 668.5% | 736.9% |
| Class TGM | 3.47% | 5.39% | 5.63% | 90.5% | 98.6% | 99% | -40.46% | -45.30% | -45.01% | 608.5% | 685.5% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | | | | | | | |
|  | **Over ECM15** | | | | | | **Over VTM11ecm15** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |  |  |  |  |
| Class B | 1.84% | 3.78% | 3.49% | 80.8% | 94.2% | 99% | -20.24% | -33.19% | -29.66% | 788.2% | 1016.4% |
| Class C | 1.52% | 2.53% | 3.79% | 81.4% | 90.3% | 98% | -23.03% | -22.61% | -23.43% | 781.3% | 1042.6% |
| Class E | 1.27% | 3.17% | 3.22% | 83.8% | 94.1% | 101% | -20.40% | -23.25% | -22.27% | 732.3% | 721.4% |
| **Overall (Ref)** | 1.59% | 3.21% | 3.52% | 81.7% | 92.9% | 99% | -21.21% | -27.18% | -25.73% | 771.6% | 940.8% |
| Class D | 0.73% | 3.31% | 3.53% | 79.4% | 92.4% | 98% | -24.88% | -22.98% | -23.61% | 793.0% | 1268.1% |
| Class F | 2.46% | 3.06% | 3.30% | 85.9% | 95.1% | 97% | -28.56% | -36.45% | -35.99% | 644.3% | 711.4% |
| Class TGM | 2.43% | 7.51% | 8.57% | 89.1% | 100.2% | 101% | -39.24% | -46.80% | -46.40% | 588.2% | 651.5% |

***Group 3 off***

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

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | | | | |
|  | **Over ECM15** | | | | | | **Over VTM11ecm15** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 | 1.01% | 3.88% | 6.28% | 72.4% | 90.9% | 100% | -13.45% | -12.15% | -22.35% | 774.4% | 487.7% |
| Class A2 | 2.29% | 6.17% | 6.09% | 71.9% | 80.8% | 100% | -18.95% | -19.12% | -24.05% | 749.0% | 457.9% |
| Class B | 2.21% | 4.85% | 5.03% | 70.1% | 79.3% | 99% | -12.52% | -18.48% | -15.92% | 698.3% | 489.5% |
| Class C | 1.77% | 2.48% | 2.73% | 70.2% | 79.9% | 99% | -12.81% | -9.09% | -10.01% | 706.8% | 474.1% |
| Class E | 3.06% | 4.71% | 4.20% | 70.5% | 79.6% | 102% | -16.14% | -18.32% | -16.75% | 676.5% | 522.8% |
| **Overall** | 2.07% | 4.36% | 4.77% | 70.8% | 81.6% | 100% | -14.41% | -15.42% | -17.17% | 716.9% | 485.7% |
| Class D | 1.17% | 1.96% | 2.23% | 71.3% | 81.4% | 100% | -11.13% | -6.43% | -6.90% | 694.5% | 543.8% |
| Class F | 6.05% | 6.86% | 7.14% | 71.3% | 69.8% | 100% | -25.91% | -29.31% | -29.28% | 437.9% | 421.5% |
| Class TGM | 11.53% | 13.48% | 13.88% | 75.8% | 57.5% | 100% | -36.70% | -42.03% | -40.99% | 389.2% | 349.5% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | | | | | | |
|  | **Over ECM15** | | | | | | **Over VTM11ecm15** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 | 0.98% | 2.60% | 6.53% | 92.0% | 96.9% | 100% | -26.21% | -21.45% | -32.08% | 1010.1% | 1106.4% |
| Class A2 | 1.39% | 3.73% | 3.70% | 93.5% | 97.7% | 99% | -29.13% | -30.88% | -36.51% | 999.4% | 1317.1% |
| Class B | 1.16% | 4.42% | 4.81% | 92.0% | 98.7% | 100% | -23.59% | -28.52% | -25.42% | 842.5% | 1179.6% |
| Class C | 0.81% | 1.42% | 1.76% | 92.0% | 95.3% | 100% | -25.73% | -20.54% | -20.98% | 892.3% | 1270.4% |
| Class E |  |  |  |  |  |  |  |  |  |  |  |
| **Overall (Ref)** | 1.08% | 3.12% | 4.12% | 92.3% | 97.2% | 100% | -25.79% | -25.45% | -27.79% | 917.9% | 1214.3% |
| Class D | 0.64% | 1.46% | 2.05% | 93.1% | 94.0% | 100% | -26.68% | -20.97% | -21.76% | 822.8% | 1284.3% |
| Class F | 4.05% | 4.77% | 4.93% | 89.7% | 93.1% | 100% | -30.19% | -32.15% | -32.84% | 672.2% | 689.6% |
| Class TGM | 7.21% | 8.87% | 9.34% | 92.6% | 91.7% | 100% | -38.19% | -43.26% | -42.80% | 627.2% | 617.7% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | | | | | | | |
|  | **Over ECM15** | | | | | | **Over VTM11ecm15** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |  |  |  |  |
| Class B | 0.43% | 1.61% | 1.06% | 95.2% | 99.6% | 100% | -21.35% | -34.64% | -31.33% | 880.5% | 1013.0% |
| Class C | 0.54% | 0.16% | 1.35% | 94.3% | 97.5% | 100% | -23.72% | -24.38% | -25.22% | 864.3% | 1126.4% |
| Class E | 0.66% | 2.69% | 2.54% | 96.2% | 96.0% | 101% | -20.87% | -23.66% | -22.84% | 814.6% | 724.1% |
| **Overall (Ref)** | 0.52% | 1.40% | 1.53% | 95.1% | 98.0% | 100% | -22.02% | -28.47% | -27.17% | 858.2% | 965.0% |
| Class D | 0.26% | 0.54% | 0.39% | 93.2% | 98.1% | 100% | -25.25% | -25.10% | -25.89% | 859.2% | 1245.6% |
| Class F | 3.48% | 4.91% | 4.95% | 92.5% | 103.1% | 100% | -27.95% | -35.49% | -35.07% | 683.8% | 683.5% |
| Class TGM | 5.10% | 6.23% | 6.51% | 92.1% | 97.7% | 100% | -37.55% | -47.22% | -46.92% | 624.9% | 575.7% |

***Group 4 off***

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

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

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | | | | |
|  | **Over ECM-15** | | | | | | **Over VTM-11ecm15** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 | 1.67% | 6.72% | 9.34% | 86.7% | 89.5% | 100% | -12.89% | -9.94% | -20.12% | 702.2% | 323.1% |
| Class A2 | 2.56% | 6.88% | 7.95% | 82.5% | 90.3% | 101% | -18.89% | -18.56% | -22.87% | 674.0% | 361.0% |
| Class B | 1.09% | 7.21% | 6.08% | 84.7% | 90.8% | 98% | -13.46% | -16.94% | -15.11% | 814.9% | 525.8% |
| Class C | 1.19% | 3.56% | 3.56% | 79.7% | 85.1% | 100% | -13.30% | -8.16% | -9.31% | 791.1% | 442.2% |
| Class E | 1.58% | 7.01% | 5.79% | 85.8% | 87.1% | 100% | -17.35% | -16.76% | -15.56% | 783.3% | 520.9% |
| **Overall** | 1.54% | 6.23% | 6.33% | 83.7% | 88.6% | 100% | -14.88% | -14.06% | -16.03% | 760.1% | 437.5% |
| Class D | 1.11% | 3.34% | 3.24% | 86.6% | 87.0% | 100% | -11.19% | -5.16% | -5.97% | 835.4% | 491.9% |
| Class F | 1.49% | 4.43% | 4.32% | 88.7% | 90.7% | 100% | -28.89% | -30.70% | -30.88% | 482.5% | 471.8% |
| Class TGM | 2.27% | 4.09% | 4.16% | 87.5% | 93.2% | 100% | -41.88% | -46.76% | -45.96% | 418.4% | 573.9% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | | | | | | |
|  | **Over ECM-15** | | | | | | **Over VTM-11ecm15** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 | 1.65% | 3.43% | 5.95% | 92.1% | 99.1% | 100% | -25.72% | -20.84% | -32.13% | 734.9% | 659.1% |
| Class A2 | 1.20% | 3.49% | 3.74% | 93.8% | 100.3% | 99% | -29.26% | -31.03% | -36.53% | 719.8% | 833.1% |
| Class B | 0.99% | 6.08% | 5.50% | 89.4% | 93.8% | 100% | -23.72% | -27.50% | -25.05% | 760.8% | 942.4% |
| Class C | 1.13% | 2.08% | 2.13% | 91.9% | 96.5% | 100% | -25.49% | -20.04% | -20.70% | 845.4% | 1037.2% |
| Class E |  |  |  |  |  |  |  |  |  |  |  |
| **Overall (Ref)** | 1.20% | 3.97% | 4.34% | 91.5% | 96.8% | 100% | -25.70% | -24.88% | -27.60% | 768.5% | 878.2% |
| Class D | 0.95% | 1.78% | 2.16% | 93.2% | 96.9% | 99% | -26.44% | -20.81% | -21.79% | 850.8% | 1139.7% |
| Class F | 1.33% | 3.25% | 3.39% | 92.0% | 95.8% | 99% | -31.87% | -32.96% | -33.63% | 656.3% | 616.2% |
| Class TGM | 0.87% | 1.72% | 1.66% | 90.8% | 94.6% | 100% | -41.87% | -46.96% | -46.79% | 568.9% | 543.6% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | | | | | | | |
|  | **Over ECM-15** | | | | | | **Over VTM-11ecm15** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |  |  |  |  |
| Class B | 1.05% | 1.63% | 1.62% | 98.4% | 99.8% | 100% | -20.86% | -34.56% | -31.01% | 875.2% | 835.4% |
| Class C | 1.29% | 1.03% | 1.11% | 93.2% | 95.8% | 99% | -23.15% | -23.76% | -25.42% | 796.9% | 862.8% |
| Class E | 0.97% | 3.29% | 2.38% | 91.4% | 91.8% | 99% | -20.63% | -23.35% | -22.93% | 758.4% | 568.2% |
| **Overall (Ref)** | 1.11% | 1.84% | 1.64% | 94.9% | 96.4% | 99% | -21.57% | -28.15% | -27.13% | 818.4% | 766.9% |
| Class D | 1.08% | 0.71% | 1.33% | 95.3% | 92.5% | 98% | -24.60% | -24.97% | -25.34% | 821.3% | 1011.7% |
| Class F | 1.31% | 2.48% | 1.29% | 95.0% | 96.7% | 99% | -29.33% | -36.87% | -37.18% | 604.8% | 519.1% |
| Class TGM | 0.54% | 1.33% | 1.40% | 100.0% | 102.9% | 102% | -40.26% | -49.53% | -49.38% | 637.6% | 509.1% |

***Group 1-4 off***

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

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | | | | |
|  | **Over ECM-15** | | | | | | **Over VTM-11ecm15** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 | 3.09% | 10.53% | 15.85% | 56.6% | 72.6% | 100% | -11.66% | -6.92% | -15.90% | 458.2% | 262.2% |
| Class A2 | 5.64% | 13.95% | 14.47% | 52.9% | 62.8% | 102% | -16.37% | -13.52% | -18.49% | 432.1% | 251.1% |
| Class B | 3.98% | 12.74% | 11.55% | 52.9% | 68.5% | 98% | -11.01% | -12.96% | -10.92% | 509.0% | 396.4% |
| Class C | 3.46% | 6.60% | 6.93% | 50.7% | 62.8% | 99% | -11.38% | -5.53% | -6.42% | 504.0% | 326.0% |
| Class E | 6.09% | 12.35% | 11.25% | 55.5% | 65.5% | 102% | -13.68% | -12.68% | -11.20% | 506.9% | 391.9% |
| **Overall** | 4.34% | 11.14% | 11.68% | 53.4% | 66.4% | 100% | -12.54% | -10.35% | -12.06% | 485.3% | 327.7% |
| Class D | 2.70% | 5.43% | 5.78% | 52.2% | 64.1% | 99% | -9.80% | -3.26% | -3.68% | 503.3% | 362.2% |
| Class F | 12.66% | 17.04% | 17.55% | 58.8% | 60.3% | 100% | -21.56% | -23.28% | -22.99% | 319.9% | 313.8% |
| Class TGM | 17.96% | 22.00% | 22.94% | 61.7% | 47.5% | 100% | -33.16% | -37.76% | -36.42% | 295.1% | 292.6% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | | | | | | |
|  | **Over ECM-15** | | | | | | **Over VTM-11ecm15** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 | 10.79% | 12.44% | 21.02% | 62.6% | 69.4% | 100% | -19.03% | -14.14% | -23.27% | 499.2% | 461.7% |
| Class A2 | 11.05% | 15.76% | 16.55% | 62.3% | 63.2% | 100% | -22.37% | -23.05% | -28.72% | 478.5% | 524.8% |
| Class B | 8.87% | 16.97% | 16.38% | 57.2% | 63.1% | 99% | -17.71% | -20.06% | -17.24% | 487.2% | 634.6% |
| Class C | 8.42% | 10.47% | 11.35% | 55.5% | 59.4% | 99% | -20.07% | -13.46% | -13.54% | 510.3% | 638.7% |
| Class E |  |  |  |  |  |  |  |  |  |  |  |
| **Overall (Ref)** | 9.57% | 14.09% | 16.00% | 58.8% | 63.3% | 99% | -19.54% | -17.71% | -19.76% | 493.9% | 574.3% |
| Class D | 6.50% | 8.99% | 10.11% | 54.5% | 57.3% | 97% | -22.30% | -15.10% | -15.56% | 497.4% | 673.6% |
| Class F | 14.30% | 17.10% | 17.50% | 63.7% | 65.7% | 99% | -23.41% | -24.51% | -25.08% | 454.5% | 422.3% |
| Class TGM | 19.13% | 24.03% | 24.77% | 67.0% | 68.1% | 99% | -31.46% | -35.60% | -35.04% | 419.7% | 391.3% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | | | | | | | |
|  | **Over ECM-15** | | | | | | **Over VTM-11ecm15** | | | | |
|  | Y | U | V | EncT | DecT | mPeakR | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |  |  |  |  |
| Class B | 9.42% | 13.50% | 13.61% | 49.2% | 53.8% | 99% | -14.27% | -27.11% | -23.06% | 438.0% | 450.7% |
| Class C | 9.25% | 10.66% | 12.52% | 47.2% | 51.1% | 98% | -17.15% | -16.56% | -17.01% | 403.4% | 460.7% |
| Class E | 9.25% | 15.64% | 15.88% | 53.8% | 62.7% | 99% | -14.10% | -14.33% | -12.65% | 446.9% | 387.9% |
| **Overall (Ref)** | 9.32% | 13.09% | 13.82% | 49.6% | 55.0% | 99% | -15.19% | -20.40% | -18.44% | 428.3% | 437.3% |
| Class D | 7.39% | 10.36% | 11.67% | 48.5% | 50.3% | 97% | -19.85% | -17.89% | -17.52% | 418.1% | 549.6% |
| Class F | 14.02% | 16.41% | 16.95% | 60.6% | 62.5% | 98% | -20.58% | -28.62% | -27.85% | 385.7% | 335.5% |
| Class TGM | 17.74% | 25.52% | 27.60% | 69.2% | 61.1% | 100% | -30.10% | -37.91% | -37.17% | 441.2% | 302.3% |

***Summary***

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

1. **Issues**

***Resolved issues***

* #83, ECM-14.1 different intra frame coding results for two successive segments (AHG17 encoding task)

***Open issues***

* #87, ECM-15.0 decoder crashes when "InterLFNST=0"
* #80, Compilation Issue ECM14.0 when disabling GPM reordering
* #78, Encoder crash in RA with --LMChroma=0
* #77, Encoder crash in RA with --DepQuant=0 --NumSignPred=0
* #71, Encode/decode mismatch when using single tree
* #66, [Difference in encoding results of ECM-12.0 in AhG7 group 2 tool off test](https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/issues/66)
* #65, [Different encoding results of ECM12 in AHG7 group1-4 off tests](https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/issues/65)
* #64, [Encode/decode mismatch and decoder crash when inter-CCCM is disable](https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/issues/64)
* #53, Decoding mismatch was observed when AMVR is off

1. **Input contributions**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| [JVET-AK0047](https://jvet-experts.org/doc_end_user/current_document.php?id=15017) | m70721 | 2024-12-09 14:16:13 | 2024-12-09 14:16:38 | 2024-12-12 16:56:00 | AHG7: Complexity and memory footprint assessment of PDP, Neural Network-based Intra Prediction, MTS and LFNST/NSPT, and ALF in ECM-15 | [G. Boisson](mailto:guillaume.boisson@interdigital.com), [T. Dumas](mailto:thierry.dumas@interdigital.com), [F. Galpin](mailto:franck.galpin@interdigital.com), [S. Puri](mailto:saurabh.puri@interdigital.com), [K. Naser](mailto:karam.naser@interdigital.com), [C. Bonnineau (InterDigital)](mailto:charles.bonnineau@interdigital.com) |
| [JVET-AK0048](https://jvet-experts.org/doc_end_user/current_document.php?id=15018) | m70734 | 2024-12-10 22:34:32 | 2024-12-11 19:13:02 | 2024-12-11 19:13:02 | Report of AHG7 conference call | [X. Li](mailto:xlxiangli@google.com) |
| [JVET-AK0049](https://jvet-experts.org/doc_end_user/current_document.php?id=15019) | m70736 | 2024-12-11 12:19:58 | 2024-12-11 12:30:13 | 2024-12-11 17:10:52 | AHG7: Analysis of MTS memory footprint and complexity | [J. Gan (OPPO)](mailto:v-jonathan.gan@oppo.com) |
| [JVET-AK0067](https://jvet-experts.org/doc_end_user/current_document.php?id=15038) | m70793 | 2025-01-07 7:47:49 | 2025-01-07 9:03:27 | 2025-01-07 9:03:27 | AHG7: Tool-off test results of LFNST extension and NSPT | [M. Koo](mailto:moonmo.koo@lge.com), [J. Zhao](mailto:jie.zhao@lge.com), [J. Lim](mailto:jaehyun.lim@lge.com), [S. Kim (LGE)](mailto:seunghwan3.kim@lge.com) |
| [JVET-AK0092](https://jvet-experts.org/doc_end_user/current_document.php?id=15063) | m70819 | 2025-01-07 10:18:35 |  |  | AHG7: Inter LDB tool assessments for each ECM version | [R. Ishimoto](mailto:ishimoto.ryo@mail.sharp), [Z. Fan](mailto:fan.zheming@sharp.co.jp), [T. Chujoh](mailto:chujoh.takeshi@sharp.co.jp), [T. Ikai (Sharp)](mailto:ikai.tomohiro@sharp.co.jp) |
| [JVET-AK0138](https://jvet-experts.org/doc_end_user/current_document.php?id=15109) | m70867 | 2025-01-07 16:29:27 | 2025-01-07 17:01:08 | 2025-01-07 17:01:08 | AHG7: Intra tool assessments between ECM14 and ECM15 | [C.-Y Teng](mailto:zachary.cy.teng@sharp-world.com.tw), [K.-W Liang](mailto:kay.kw.liang@sharp-world.com.tw), [Y.-H Lin](mailto:ekko.yh.lin@sharp-world.com.tw), [Y.-C Yang](mailto:yuchiao.yc.yang@sharp-world.com.tw), [T.Ikai (Sharp)](mailto:ikai.tomohiro@sharp.co.jp) |
| [JVET-AK0215](https://jvet-experts.org/doc_end_user/current_document.php?id=15204) | m70968 | 2025-01-07 23:03:34 |  |  | Non-EE2: MTS with reduced memory | [B. Ray](mailto:bray@qti.qualcomm.com), V. Seregin, M. Karczewicz (Qualcomm) |
| [JVET-AK0228](https://jvet-experts.org/doc_end_user/current_document.php?id=15217) | m70986 | 2025-01-08 4:58:08 |  |  | AHG7: Tool-off test results of ALF fixed filters | [N. Hu](mailto:nanh@qti.qualcomm.com), [V. Seregin](mailto:vseregin@qti.qualcomm.com), [M. Karczewicz (Qualcomm)](mailto:martak@qti.qualcomm.com) |

1. **Recommendations**

* Continue and improve tool assessment
* Resolve identified software issues related to the tool assessment

It was commented that reporting encoding/decoding run time cannot be the only criterion. Criteria such as memory access and bandwidth, number of maximum processing cycles, processing complexity, block decoding dependencies, number of context coded bins, pipeline and parallelization problems should be put more in focus of the AHG, in particular to develop reporting criteria that would allow more precise assessment of tools. Something like that was developed during HEVC and VVC development. Enforce mandates of AHG, conference call in interim period, expecting input contributions to the next meeting. This would later be relevant for a CfP, and some ECM tools may be good examples for this study.

[JVET-AK0008](https://jvet-experts.org/doc_end_user/current_document.php?id=15175) JVET AHG report: Optimization of encoders and receiving systems for machine analysis of coded video content (AHG8) [S. Liu, J. Ström, S. Wang, M. Zhou (AHG chairs)]

1. **Activities**

The AHG 8 used the main JVET reflector, jvet@lists.rwth-aachen.de, for email discussion. There were no email exchanges in the main reflector between the last and this meeting, but active email discussions (50+ emails) among co-chairs, editors and proponents on dense QP result generation and CDTR preparation. There are 7 input contributions related to AHG 8 mandates submitted to this meeting. They are listed in Section 3.

***Software and Common Test Conditions***

AHG 8 related software and documents can be accessed at <https://vcgit.hhi.fraunhofer.de/jvet-ahg-ofm>. This repository contains two projects: common test conditions, reporting templates with anchor results, evaluation scripts and task networks are available in <https://vcgit.hhi.fraunhofer.de/jvet-ahg-ofm/ofm-ctc>, and software implementation examples are hosted in <https://vcgit.hhi.fraunhofer.de/jvet-ahg-ofm/vtm-ofm>.

For this meeting cycle, common test conditions remain unchanged as described in output document JVET-AI2031. Some software implementation examples are available in the repository:

|  |  |
| --- | --- |
| **Branch name** | **Description** |
| JVET-AB0275 | Region-of-Interest-based adaptive QP |
| JVET-AC0086 | Foreground/background separation |
| JVET-AD0122 | Temporal QP offsets |
| JVET-AE0143 | Spatial resampling (Experimental) |
| JVET-AG0212 | Post-processing filter (Removed by proponent) |
| JVET-AH0130 | Spatial resampling |
| JVET-AH0157 | Combined pre- and post-processing |
| JVET-AJ0178 | Lightweight post-processing |
| JVET-AJ0181 | Combined software JVET-AB0275, JVET-AC0086, JVET-AJ0178 |
| JVET-AJ0254 | Pre-analysis based temporal resampling |

Following the discussion in the last (36th) JVET meeting, dense QP experiments were conducted on JVET-AJ0181 to produce results for all QPs in the range from 16 to 58. The results are reported in JVET-AK0122 and crosschecked in JVET-AK0272 and JVET-AK0289.

***Technical Report***

The seventh draft of the technical report (TR) JVET-AJ2030 “Optimization of encoders and receiving systems for machine analysis of coded video content (draft 7)” and was produced, including the following additions compared with draft 6:

* Text on chroma processing (to clause 8) JVET-AJ0168
* Post-processing algorithm (to annex A.3) JVET-AJ0178
* Combinations to new annex (to annex B) JVET-AJ0181, JVET-AJ0232, JVET-AJ0233
* Editorial change on RPR and aggressive downsampling (to clause 7.4) JVET-AJ0232
* Editorial change on useful combinations in the annex (to clause 5.1) JVET-AJ0233
* Non-monotonicity issues and curve fitting (to clause 6) JVET-AJ0235
* Temporal subsampling (to annex A.4) JVET-AJ0254
* Description of MIOU (to clause 6) JVET-AJ0254

CDTR was generated based on JVET-AJ2030 and submitted for ballot (comments).

1. **Input contributions**

There are 7 input contributions related to AHG 8 mandates. They are listed below.

|  |  |  |
| --- | --- | --- |
| **Report** | | |
| [JVET-AK0008](file:////Users/shanl/Documents/contribution/jvet37ak/ahg8/current_document.php%3fid=15175) | JVET AHG report: Optimization of encoders and receiving systems for machine analysis of coded video content (AHG8) | S. Liu, J. Ström, S. Wang, M. Zhou (AHG chairs) |
| **Proposals** | | |
| [JVET-AK0094](file:////Users/shanl/Documents/contribution/jvet37ak/current_document.php%3fid=15065) | AHG8: On combination of adaptive temporal resampling, pre-processing, post-processing and ROI-based adaptive QP algorithm for machine vision | [S. Wang](mailto:shurun.wsr@alibaba-inc.com), [J. Chen](mailto:jiechen.cj@alibaba-inc.com), [Y. Ye](mailto:yan.ye@alibaba-inc.com), [B. Li (Alibaba)](mailto:libinzhe.lbz@alibaba-inc.com), [S. Wang (CityUHK)](mailto:shiqwang@cityu.edu.hk) |
| [JVET-AK0122](file:////Users/shanl/Documents/contribution/jvet37ak/current_document.php%3fid=15093) | AHG8: Dense QP results for joint software | [C. Hollmann](mailto:christopher.hollmann@ericsson.com), [J. Ström (Ericsson)](mailto:jacob.strom@ericsson.com) |
| [JVET-AK0140](file:////Users/shanl/Documents/contribution/jvet37ak/current_document.php%3fid=15111) | AHG9/AHG8: Showcase for Packed regions information SEI | [J. Boyce](mailto:jill.boyce@nokia.com), H. Zhang, [M. M. Hannuksela (Nokia)](mailto:miska.hannuksela@nokia.com) |
| [JVET-AK0141](file:////Users/shanl/Documents/contribution/jvet37ak/current_document.php%3fid=15112) | AHG8: Add Packed regions info SEI to TR on coding video for machine consumption | [J. Boyce](mailto:jill.boyce@nokia.com), [H. Zhang](mailto:honglei.1.zhang@nokia.com), [M. M. Hannuksela (Nokia)](mailto:miska.hannuksela@nokia.com) |
| **Crosschecks** | | |
| [JVET-AK0272](file:////Users/shanl/Documents/contribution/jvet37ak/current_document.php%3fid=15261) | Cross-check of JVET-AK0122 on SFU-HW dataset (AHG8: Dense QP results for joint software) | [S. Wang (Alibaba)](mailto:shurun.wsr@alibaba-inc.com) |
| [JVET-AK0289](file:////Users/shanl/Documents/contribution/jvet37ak/current_document.php%3fid=15278) | Cross-check of JVET-AK0122 on TVD dataset (AHG8: Dense QP results for joint software) | [M. Xu (Tencent)](mailto:motongxu@global.tencent.com) |

1. **Recommendations**

The AHG recommends to:

* Review all input contributions.
* Continue improving draft TR based on CDTR feedback and other inputs, towards TR finalization.
* Continue investigating non-normative technologies and their suitability for machine analysis applications.
* Continue refining test conditions, evaluation and reporting procedures.

A mismatch was observed in the dense QP results, which was resolved according to the newest upload of JVET-AK0122.

Should be presented to WG 4, but no joint meeting necessary at this time.

Align timeline with VSEI v4, which could be DTR in April, final approval (also ITU) in October

[JVET-AK0009](https://jvet-experts.org/doc_end_user/current_document.php?id=15176) JVET AHG report: SEI message studies (AHG9) [S. McCarthy, Y.-K. Wang (co-chairs), J. Boyce, T. Chujoh, S. Deshpande, C. Fogg, M. M. Hannuksela, P. de Lagrange, G. J. Sullivan, H. Tan, A. Tourapis, S. Wenger (vice-chairs)]

**Related contributions**

A total of 65 contributions are identified relating to the mandates of AHG9. Some contributions relate to more than one mandate and/or SEI message. Some contributions also relate to the work of AHG13.

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

***Study the SEI messages in VSEI, VVC, HEVC and AVC (5)***

**JVET-AK2005 Additions and corrections for VVC version 4 (*4*)**

[JVET-AK0053](https://jvet-experts.org/doc_end_user/current_document.php?id=15024) AHG9: NNPF interface text in VVC [Y.-K. Wang, J. Xu, L. Zhang (Bytedance), K. Yang, Y. Li, Y. Xu (SJTU)]

[JVET-AK0080](https://jvet-experts.org/doc_end_user/current_document.php?id=15051) AHG9: Comments on the GFV and GFVE SEI messages [M. M. Hannuksela (Nokia), J. Chen, B. Chen, Y. Ye (Alibaba)]

*Also relates to mandate* *to study JVET-AK2006 (GFV SEI messages*)

[JVET-AK0160](https://jvet-experts.org/doc_end_user/current_document.php?id=15131) AHG9: On VVC interface of object mask information and annotated regions SEI messages [J. Chen, B. Chen, Y. Ye (Alibaba), M. M. Hannuksela (Nokia)]

[JVET-AK0197](https://jvet-experts.org/doc_end_user/current_document.php?id=15184) AHG13: on picture size for the interpretation of the FGC SEI message [P. de Lagrange, E. François (InterDigital)]

**Additions for HEVC *(1)***

[JVET-AK0107](https://jvet-experts.org/doc_end_user/current_document.php?id=15078) AHG9: Modality Information SEI for HEVC [J. Gao, H.-B. Teo, C.-S. Lim, K. Abe, V. Drugeon (Panasonic)]

*Also relates to mandate* *to study JVET-AK2006 (modularity information SEI message*)

***Study JVET-AK2006 Additional SEI messages for VSEI version 4 (32)***

**NNPF SEI messages *(9)***

[JVET-AK0054](https://jvet-experts.org/doc_end_user/current_document.php?id=15025) AHG9: Semantics of the NNPFC and NNPFA SEI messages [Y.-K. Wang, J. Xu, L. Zhang (Bytedance), K. Yang, Y. Li, Y. Xu (SJTU)]

[JVET-AK0070](https://jvet-experts.org/doc_end_user/current_document.php?id=15041) AHG9: On seed parameter for NNPF [T. Chujoh, Z. Fan, R. Ishimoto, T. Ikai (Sharp), L. Jin, H. Watanabe (Waseda Univ.)]

[JVET-AK0072](https://jvet-experts.org/doc_end_user/current_document.php?id=15043) AHG9: Comments on NNPF SEI messages [M. M. Hannuksela, F. Cricri (Nokia)]

[JVET-AK0074](https://jvet-experts.org/doc_end_user/current_document.php?id=15045) AHG9: Handling output variability of an NNPF [M. M. Hannuksela, F. Cricri (Nokia)]

[JVET-AK0147](https://jvet-experts.org/doc_end_user/current_document.php?id=15118) AHG9: On NNPF for Tone Mapping with Colour Volume Information [C.-H. Demarty, E. François, F. Aumont, O. Le Meur (InterDigital)]

[JVET-AK0152](https://jvet-experts.org/doc_end_user/current_document.php?id=15123) AHG9: On Spatial Extrapolation for NNPF [S. Deshpande (Sharp), M. M. Hannuksela (Nokia)]

[JVET-AK0167](https://jvet-experts.org/doc_end_user/current_document.php?id=15138) AHG9: On Enabling the usage of auxiliary layers for NNPF [T. M. Borges, Y. Sanchez, R. Skupin, C. Hellge, T. Schierl (Fraunhofer HHI)]

[JVET-AK0188](https://jvet-experts.org/doc_end_user/current_document.php?id=15159) AHG9: On NNPF filtering control strength value adjustment [J. Xu, Y.-K. Wang (Bytedance)]

**SEI processing order and processing order nesting SEI messages (*5*)**

[JVET-AK0055](https://jvet-experts.org/doc_end_user/current_document.php?id=15026) AHG9: Semantics of the SPO SEI message [Y.-K. Wang, J. Xu, L. Zhang (Bytedance), K. Yang, Y. Li, Y. Xu (SJTU)]

[JVET-AK0112](https://jvet-experts.org/doc_end_user/current_document.php?id=15083) AHG9: On the SPO SEI message [Y. Gao, P. Wu, Y. Bai, S. Xie, M. Jia, W. Niu, C. Huang (ZTE)]

[JVET-AK0156](https://jvet-experts.org/doc_end_user/current_document.php?id=15127) AHG9: On SEI processing order SEI message [C.H. Demarty, E. François, F. Aumont, O. Le Meur (InterDigital)]

[JVET-AK0165](https://jvet-experts.org/doc_end_user/current_document.php?id=15136) AHG9: On SPO sub-chain signaling [T. M. Borges, Y. Sanchez, R. Skupin, C. Hellge, T. Schierl (Fraunhofer HHI)]

[JVET-AK0205](https://jvet-experts.org/doc_end_user/current_document.php?id=15194) AHG9: On SEI messages in spoPropertySeiList [H. Tan, J. Lee, J. Nam, C. Kim, J. Lim, S. Kim (LGE)]

**Encoder optimization information SEI message *(2)***

[JVET-AK0075](https://jvet-experts.org/doc_end_user/current_document.php?id=15046) AHG9: Additional information on object-based optimization for the encoder optimization information SEI message [M. Pedzisz, M. M. Hannuksela, F. Cricri (Nokia)]

[JVET-AK0155](https://jvet-experts.org/doc_end_user/current_document.php?id=15126) AHG9: On Encoder optimization information SEI message [C.H. Demarty, E. François, F. Aumont, O. Le Meur (InterDigital)]

**Source picture timing information SEI message *(1)***

[JVET-AK0168](https://jvet-experts.org/doc_end_user/current_document.php?id=15139) [m70903 [AHG9: On SPTI SEI message: Robustness and source constant framerate [Y. Sanchez, T. M. Borges, R. Skupin, C. Hellge, T. Schierl (Fraunhofer HHI)]

**Modality information SEI message *(1)***

[JVET-AK0107](https://jvet-experts.org/doc_end_user/current_document.php?id=15078) AHG9: Modality Information SEI for HEVC [J. Gao, H.-B. Teo, C.-S. Lim, K. Abe, V. Drugeon (Panasonic)]

*Also relates to mandate* *to study JVET-AK2005*

**Generative face video SEI messages *(8)***

[JVET-AK0080](https://jvet-experts.org/doc_end_user/current_document.php?id=15051) AHG9: Comments on the GFV and GFVE SEI messages [M. M. Hannuksela (Nokia), J. Chen, B. Chen, Y. Ye (Alibaba)]

*Also relates to mandate* *to study JVET-AK2005*

[JVET-AK0124](https://jvet-experts.org/doc_end_user/current_document.php?id=15095) AHG9: On timing information and order of pictures in GFV SEI message [J. Lee, H. Tan, C. Kim, J. Nam, J. Lim, S. Kim (LGE)]

[JVET-AK0127](https://jvet-experts.org/doc_end_user/current_document.php?id=15098) AHG9: On miscellaneous aspects of GFV and DSCI SEI messages [J. Lee, H. Tan, C. Kim, J. Nam, J. Lim, S. Kim (LGE)]

*Also relates to DSC SEI messages*

[JVET-AK0128](https://jvet-experts.org/doc_end_user/current_document.php?id=15099) AHG9: Editorial updates for GFV SEI message [J. Lee, H. Tan, C. Kim, J. Nam, J. Lim, S. Kim (LGE)]

[JVET-AK0154](https://jvet-experts.org/doc_end_user/current_document.php?id=15125) AHG9: Comments on Generative Face Video SEI [A. C. Sidiya, S. Deshpande (Sharp)]

[JVET-AK0164](https://jvet-experts.org/doc_end_user/current_document.php?id=15135) AHG9: Supplementation of value range definition and editorial bugs fixes on the GFVE pupil position SEI messages [F. Ma, A. Trioux, F. Yang (Xidian Univ.), B. Li, F. Xing, Z. Wang (Hisense)]

[JVET-AK0238](https://jvet-experts.org/doc_end_user/current_document.php?id=15227) AHG9: Semantics and syntax fixes for generative face video SEI message [J. Chen, B. Chen, Y. Ye (Alibaba)]

[JVET-AK0239](https://jvet-experts.org/doc_end_user/current_document.php?id=15228) AHG9: On generative face video enhancement SEI message [J. Chen, B. Chen, Y. Ye (Alibaba)]

**Digitally signed content messages *(8)***

[JVET-AK0109](https://jvet-experts.org/doc_end_user/current_document.php?id=15080) AHG9: On signaling and constraint for RefDigest in digitally signed content SEI messages [C. Kim, H. Tan, J. Nam, J. Lee, J. Lim, S. Kim (LGE)]

[JVET-AK0110](https://jvet-experts.org/doc_end_user/current_document.php?id=15081) AHG9: On reference message digest for verification in digitally signed content SEI messages [C. Kim, H. Tan, J. Nam, J. Lee, J. Lim, S. Kim (LGE)]

[JVET-AK0125](https://jvet-experts.org/doc_end_user/current_document.php?id=15096) AHG9: On presence and persistency of digitally signed content SEI messages [J. Lee, H. Tan, C. Kim, J. Nam, J. Lim, S. Kim (LGE)]

[JVET-AK0126](https://jvet-experts.org/doc_end_user/current_document.php?id=15097) AHG9: On signaling the mapping between verification substreams and layer and temporal sublayer for DSCI SEI message [J. Lee, H. Tan, C. Kim, J. Nam, J. Lim, S. Kim (LGE)]

[JVET-AK0127](https://jvet-experts.org/doc_end_user/current_document.php?id=15098) AHG9: On miscellaneous aspects of GFV and DSCI SEI messages [J. Lee, H. Tan, C. Kim, J. Nam, J. Lim, S. Kim (LGE)]

*Also relates to GFV SEI messages*

[JVET-AK0194](https://jvet-experts.org/doc_end_user/current_document.php?id=15165) AHG9: Digitally Signed Content SEI messages for AVC and HEVC [K. Suehring, T. Hinz, Y. Sanchez, J. Pfaff, H. Schwarz, D. Marpe, T. Wiegand (Fraunhofer HHI)]

[JVET-AK0206](https://jvet-experts.org/doc_end_user/current_document.php?id=15195) AHG9: On digitally signed content SEI messages [Sean McCarthy, Christof Fersch, Iraj Sodagar (Dolby)]

[JVET-AK0287](https://jvet-experts.org/doc_end_user/current_document.php?id=15276) AHG9: Multilayer digitally signed content authentication SEI messages [J. Boyce, M. M. Hannuksela (Nokia)]

***Study JVET-AK2032 TuC for VSEI (26)***

**NNPF SEI messages *(1)***

[JVET-AK0073](https://jvet-experts.org/doc_end_user/current_document.php?id=15044) AHG9: Enabling multiple instances of NNPF extrapolated pictures[M. M. Hannuksela, F. Cricri (Nokia)]

**Film grain regions characteristics SEI message *(2)***

[JVET-AK0058](https://jvet-experts.org/doc_end_user/current_document.php?id=15029) [AHG9/AHG13] Comments on Film grain regions characteristics SEI message [S. Xie, P. Wu, Y. Gao, Y. Bai, C. Huang (ZTE)]

[JVET-AK0211](https://jvet-experts.org/doc_end_user/current_document.php?id=15200) AHG9/AHG13: Proposed Film Grain Region SEI message for version 5 of VSEI [S. Wenger, G. Teniou, A.T. Hinds, (Tencent), P. de Lagrange, E. François, D. Doyen, (InterDigital)]

**Constituent rectangle SEI messages (*4)***

[JVET-AK0115](https://jvet-experts.org/doc_end_user/current_document.php?id=15086) AHG9: On supporting 4:4:4 color format in constituent rectangles and enhanced colour format information SEI messages [C. Kim, H. Tan, J. Nam, J. Lee, J. Lim, S. Kim (LGE)]

*Also relates to ECFI SEI message*

[JVET-AK0209](https://jvet-experts.org/doc_end_user/current_document.php?id=15198) AHG9: On constituent rectangles SEI message [H. Tan, C. Kim, J. Nam, J. Lee, J. Lim, S. Kim (LGE)]

[JVET-AK0210](https://jvet-experts.org/doc_end_user/current_document.php?id=15199) AHG9: On constituent rectangle grouping for 4:4:4 color format support in constituent rectangle SEI message [H. Tan, J. Nam, C. Kim, J. Lee, J. Lim, S. Kim (LGE)]

[JVET-AK0278](https://jvet-experts.org/doc_end_user/current_document.php?id=15267) AHG9: Constituent Rectangles / Enhanced Colour Format Information SEI extension to support 4:2:2 source content coded in bitstream with 4:2:0 chroma format [S. Keating, M. Ikeda (Sony)]

*Also relates to ECFI SEI message*

**Quality metrics SEI message *(3)***

[JVET-AK0159](https://jvet-experts.org/doc_end_user/current_document.php?id=15130) AHG9: On miscellaneous aspects of quality metrics SEI message [J. Nam, H. Tan, J. Lee, C. Kim, J. Lim, S. Kim (LGE)]

[JVET-AK0207](https://jvet-experts.org/doc_end_user/current_document.php?id=15196) AHG9: On signalling of average quality in quality metrics SEI message [H. Tan, J. Nam, J. Lee, C. Kim, J. Lim, S. Kim (LGE)]

[JVET-AK0208](https://jvet-experts.org/doc_end_user/current_document.php?id=15197) AHG9: On the inclusion of quality metric SEI message in SEI processing order SEI message [H. Tan, C. Kim, J. Nam, J. Lee, J. Lim, S. Kim (LGE)]

**Lens optical correction SEI messages *(2)***

[JVET-AK0179](https://jvet-experts.org/doc_end_user/current_document.php?id=15150) AHG9: On Lens Optical Correction SEI message [J. Samuelsson-Allendes, S. Deshpande (Sharp)]

[JVET-AK0204](https://jvet-experts.org/doc_end_user/current_document.php?id=15193) AHG9: Proposed Lens Optical Correction SEI message for version 5 of VSEI [S. Wenger, G. Teniou, A. T. Hinds (Tencent)]

**Display overlays SEI message *(6)***

[JVET-AK0099](https://jvet-experts.org/doc_end_user/current_document.php?id=15070) AHG9: On bit depth alignment between target and coded pictures in DOI SEI [T. Biatek, J. Boyce, M. M. Hannuksela (Nokia)]

[JVET-AK0100](https://jvet-experts.org/doc_end_user/current_document.php?id=15071) AHG9: DOI SEI support for target picture with alpha channel [T. Biatek, J. Boyce, M. M. Hannuksela (Nokia)]

[JVET-AK0157](https://jvet-experts.org/doc_end_user/current_document.php?id=15128) AHG9: On the number of DOI SEI message in an AU [J. Nam, H. Tan, J. Lee, C. Kim, J. Lim, S. Kim (LGE)]

[JVET-AK0158](https://jvet-experts.org/doc_end_user/current_document.php?id=15129) AHG9: On reconstruction of target display picture in DOI SEI message [J. Nam, H. Tan, J. Lee, C. Kim, J. Lim, S. Kim (LGE)]

[JVET-AK0190](https://jvet-experts.org/doc_end_user/current_document.php?id=15161) AHG9: On the display overlays information SEI message [J. Xu, Y.-K. Wang (Bytedance)]

[JVET-AK0265](https://jvet-experts.org/doc_end_user/current_document.php?id=15254) AHG9: Display overlay sets for DOI SEI [J. Boyce, T. Biatek, M. M. Hannuksela (Nokia)]

**Bitdepth range information SEI message (1*)***

[JVET-AK0189](https://jvet-experts.org/doc_end_user/current_document.php?id=15160) AHG9: On the bitdepth range information SEI message [J. Xu, Y.-K. Wang (Bytedance)]

**AI usage restrictions SEI message *(2)***

[JVET-AK0114](https://jvet-experts.org/doc_end_user/current_document.php?id=15085) AHG9: Updates and suggestion on AI usage restrictions SEI message [C. Kim, H. Tan, J. Nam, J. Lee, J. Lim, S. Kim (LGE)]

[JVET-AK0281](https://jvet-experts.org/doc_end_user/current_document.php?id=15270) AHG9: On AI usage restrictions SEI message [M. M. Hannuksela, F. Cricri (Nokia)]

**Enhanced colour format information SEI message *(3)***

[JVET-AK0115](https://jvet-experts.org/doc_end_user/current_document.php?id=15086) AHG9: On supporting 4:4:4 color format in constituent rectangles and enhanced colour format information SEI messages [C. Kim, H. Tan, J. Nam, J. Lee, J. Lim, S. Kim (LGE)]

*Also relates to CR SEI message*

[JVET-AK0278](https://jvet-experts.org/doc_end_user/current_document.php?id=15267) AHG9: Constituent Rectangles / Enhanced Colour Format Information SEI extension to support 4:2:2 source content coded in bitstream with 4:2:0 chroma format [S. Keating, M. Ikeda (Sony)]

*Also relates to CR SEI message*

[JVET-AK0286](https://jvet-experts.org/doc_end_user/current_document.php?id=15275) AHG9: Simplified enhanced colour format information SEI [J. Boyce, M. M. Hannuksela (Nokia)]

**Photosensitive content information SEI message *(1)***

[JVET-AK0153](https://jvet-experts.org/doc_end_user/current_document.php?id=15124) AHG9: On Photosensitive Content Information Signaling [S. Deshpande (Sharp)]

**Progression of specifications *(1)***

[JVET-AK0178](https://jvet-experts.org/doc_end_user/current_document.php?id=15149) AHG9: Proposed plan for initiating version 5 of VSEI [A. T. Hinds, S. Wenger (Tencent), P. de Lagrange, E. François (InterDigital)]

***Collect software and showcase information for SEI messages (2)***

[JVET-AK0140](https://jvet-experts.org/doc_end_user/current_document.php?id=15111) AHG9/AHG8: Showcase for Packed regions information SEI [J. Boyce, H. Zhang, M. M. Hannuksela (Nokia)]

[JVET-AK0258](https://jvet-experts.org/doc_end_user/current_document.php?id=15247) AHG9/AHG11: Demo of real-time NNPF inference [J. Funnell (Nokia), M. Santamaria, F. Cricri, M. M. Hannuksela, M. Pedzisz, R. Yang, S. Schwarz]

***Identify potential needs for additional SEI messages, including study of AVC and HEVC SEI messages for use in VVC (4)***

[JVET-AK0136](https://jvet-experts.org/doc_end_user/current_document.php?id=15107) AHG9: Target Colour Volume SEI message [C.H. Demarty, E. François, F. Aumont, O. Le Meur (InterDigital)]

[JVET-AK0142](https://jvet-experts.org/doc_end_user/current_document.php?id=15113) AHG9: Display rectangles SEI [J. Boyce, T. Biatek, M. M. Hannuksela (Nokia)

[JVET-AK0191](https://jvet-experts.org/doc_end_user/current_document.php?id=15162) AHG9: Examples of danmaku applications [J. Xu, Y.-K. Wang, L. Zhang (Bytedance)]

[JVET-AK0203](https://jvet-experts.org/doc_end_user/current_document.php?id=15192) AHG9: On signalling segmentation and object tracking information [E. Thomas, E. Potetsianakis, E. Alexiou, M.-L. Champel (Xiaomi)]

**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. No emails with [AHG9] were exchanged.

**Recommendations**

The AHG recommends to:

* Review all related contributions; and
* Continue SEI messages studies.

Considering extensions of NNPF, it should be assessed which elements are mature enough for v4, which are going to TuC, or even more study needed.

It was commented that it would be premature to request for a version 5 of VSEI, as v4 has not yet reached DAM/DIS stage, and ISO rules would forbid going for another amendment already. As for now, TuC is sufficient to document potential candidates.

[JVET-AK0010](https://jvet-experts.org/doc_end_user/current_document.php?id=15177) JVET AHG report: Encoding algorithm optimization (AHG10) [P. de Lagrange, A. Duenas, R. Sjöberg, A. Tourapis (AHG chairs)]

1. **Activity**

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

Teleconferences were held on December 4th and December 19th, and a few related organizational e-mails were sent to the JVET reflector, as well as a few support requests and answers to reproduce the results from the previous meeting cycle that were used for multilayer viewing tests during the Kemer meeting.

1. **Related contributions**

Two contributions were identified relating to AHG10 and summarized in the following sections.

***General***

**JVET-AK0052 - AHG10: Teleconference on encoder optimization**

This document summarizes discussions during teleconferences on encoder optimization.

The remaining work related to VVC multi-layer coding was discussed, including support for multi-view (suggested tests, test sequences, expected software tools improvements and if specific configurations and encoder optimizations for visual testing would be acceptable.

***Common test conditions***

**JVET-AK0181 - AHG10: Suggested upgrade of CTC for random access for VTM**

This contribution suggests upgrading the VTM settings for random access to better handle a larger span of bitrates:

* Enabling GOP-based RPR as it can be beneficial to encode at a lower resolution for lower bitrates
* Enabling DMVR encoder control to improve subjective quality in some situations also especially at lower bitrates

Minor BD-rate gains (around -0.3%) are reported using the CTC, but more gain (up to -1.6%) when testing higher QPs, and more gain at higher resolutions. Visual quality improvement is reported, especially at low bitrates.

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-AK0011](https://jvet-experts.org/doc_end_user/current_document.php?id=15178) 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)]

**Activities**

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

***Common Test Conditions***

**Anchor Encoding**

Anchor for the NN-based video coding activity made available on Dec. 17 though the Git repository used for the AHG activity:

https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/nnvc-ctc/tree/master/Anchor%20performance?ref\_type=heads

also distributed by AhG14 in JVET-AK0014.

***New training materials***

JVET group received new training materials are response for the call issued during 33rd meeting. In total 78 new video scenes can be potentially added to NNVC training data set.

|  |  |  |
| --- | --- | --- |
| [JVET-AK0255](https://jvet-experts.org/doc_end_user/current_document.php?id=15244) | [AHG11] Response to Call for training materials for neural network-based video coding tool development | [Fan Zhang](mailto:fan.zhang@bristol.ac.uk), [David Bull](mailto:dave.bull@bristol.ac.uk), [Jakub Nawala](mailto:jakub.nawala@bristol.ac.uk), [Yuxuan Jiang](mailto:yuxuan.jiang@bristol.ac.uk), [Xiaoqing Zhu](mailto:xzhu@netflix.com), [Joel Sole](mailto:jsole@netflix.com), [Elena Alshina](mailto:elena.alshina@huawei.com) |

***Dataset access***

It has been reported by multiple proponents trying to retrain models from scratch that the original datasets used to train the current models have changed since their first use. New training materials come with permission to copy them to the JVET hosting. Group should discuss if hosting for training data is possible.

For sequences which are not hosted by JVET changes need to tracked regularly and md5sum updated if needed.

***Interaction with ECM***

Neural network based Intra (not identical, but similar to one in NNVC) was adopted to ECM showing compression performance benefits 0.2% BD-rate in random access configuration (relatively to VTM gain of NN-Intra in NNVC set was more than 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-AK0023](https://jvet-experts.org/doc_end_user/current_document.php?id=15189) | 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) |

***Teleconferences***

The AHG conducted two joint teleconferences with AHG14 and EE1 during the interim period. The teleconferences were held on December 02 and December 20, 2024. In those teleconferences, the following topics were discussed:

* NN-based filter (LOP4) joint training report and LOP4 inference cross-check
* NNVC11.0 software integration status and anchors performance,
* EE1 description finalization

Combination of two proposals on LOP and VLOP filters, ‘Neural network-based in-loop filters using early cropping’ from JVET-AJ054 and ‘Reduced complexity input feature extraction for LOP & VLOP’ from [JVET-AJ0066](https://jvet-experts.org/doc_end_user/current_document.php?id=14653), were jointly trained together, verified though the training cross-check and adoption to NNVC-11.0 was confirmed during joined teleconference of AhG11/14.

As was agreed earlier EE1 proponents design their tests keeping complexity in kMAC/pxl not exceeding competing tool in NNVC and keeping the number of channels as multiple 16. If not possible to meet both requirements at the same time then two sub-tests are conducted (this is the main reason for multiple sub-tests in EE1).

***Performance Evaluation***

The performance and complexity of NN-based tools available in NNVC SW is summarized in the table below. All test data provided by AhG14. Encoding and decoding run time is very dependent on cluster used for simulation. Run time data in this table are all from InterDigital.

In NNVC-11, LOP and VLOP filters performance improved 0.2% (in random access test).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Test vs NNVC (configured as VTM)** | **Random Access cfg** | | | | | **kMAC/pxl** | | | | **Param (Mprm)** | | | |
| **Y** | U | V | Enc | Dec | **Total** | Filter | Intra | SR | **Total** | Filter | Intra | SR |
| NN-Intra & LOP filter (2 tools) | | | | | | | | | | | | | |
| NNVC-11.0 (LOP4) | **-7.6%** | -14.3% | -13.2% | 1.2 | 36 | **21.6** | 16.8 | 4.8 | 0 | **1.5** | 0.21 | 1.3 | 0 |
| NNVC-10.0 (LOP3) | **-7.4%** | -13.6% | -11.6% | 1.2 | 33 | **21.7** | 16.9 | 4.8 | 0 | **1.5** | 0.21 | 1.3 | 0 |
| NNVC-9.1(LOP3) | **-7.3%** | -13.1% | -11.3% | 1.2 | 81 | **24.8** | 16.9 | 7.9 | 0 | **1.7** | 0.21 | 1.5 | 0 |
| NNVC-8.0(LOP2) | **-6.9%** | -13.2% | -12.1% | 1.2 | 73 | **25.0** | 17.1 | 7.9 | 0 | **1.6** | 0.05 | 1.5 | 0 |
| NNVC-7.1(LOP2) | **-6.9%** | -13.2% | -12.1% | 1.3 | 86 | **25.0** | 17.1 | 7.9 | 0 | **1.6** | 0.05 | 1.5 | 0 |
| NNVC-9.1(LOP2CA) | **-8.2%** | -16.5% | -15.5% | 2.5 | 69 | **25.5** | 17.6 | 7.9 | 0 | **1.6** | 0.05 | 1.5 | 0 |
| NN-Intra & HOP filter (2 tools) | | | | | | | | | | | | | |
| NNVC-10.0 (HOP5) | **-14.2%** | -19.5% | -19.9% | 2.6 | 1135 | **471** | 466 | 4.8 | 0 | **2.7** | 1.4 | 1.3 | 0 |
| NNVC-9.1(HOP4) | **-14.1%** | -19.2% | -19.6% | 2.9 | 1447 | **484** | 476 | 7.9 | 0 | **3.0** | 1.44 | 1.5 | 0 |
| NNVC-8.0(HOP3) | **-13.7%** | -13.9% | -14.5% | 2.5 | 1092 | **474** | 466 | 7.9 | 0 | **2.9** | 1.40 | 1.5 | 0 |
| NNVC-7.0(HOP2) | **-13.6%** | -12.5% | -14.2% | 4.1 | 2071 | **485** | 477 | 7.9 | 0 | **3.0** | 1.50 | 1.5 | 0 |
| NN-Intra & VLOP filter (2 tools) | | | | | | | | | | | | | |
| NNVC-11.0 (VLOP3) | **-5.8%** | -6.6% | -5.7% | 1.2 | 7 | **9.9** | 5.10 | 4.8 | 0 | **1.3** | 0.07 | 1.3 | 0 |
| NNVC-10.0 (VLOP2) | **-5.6%** | -7.6% | -6.4% | 1.1 | 15 | **10** | 5.16 | 4.8 | 0 | **1.3** | 0.06 | 1.3 | 0 |
| NNVC-9.1(VLOP) | **-5.3%** | -5.4% | -5.2% | 1.2 | 40 | **13** | 5.12 | 7.9 | 0 | **1.5** | 0.02 | 1.5 | 0 |
| NN-Intra & LOP filter content adaptive (2 tools) | | | | | | | | | | | | | |
| NNVC-10.0 aLOP3 | **-8.2%** | -16.7% | -15.5% | 2.2 | 34 | **21.8** | 16.9 | 4.8 | 0 | **1.5** | 0.21 | 1.3 | 0 |
| NN-Intra & LOP filter & adaptive resolution coding (3 tools) | | | | | | | | | | | | | |
| NNVC-11.0 NNSR | **-8.5%** | -12.2% | -10.9% |  |  | **26.3** | 16.8 | 4.8 | 4.7 | **1.4** | 0.05 | 1.3 | 0.05 |
| NNVC-8.0 RPR | **-7.5%** | -10.9% | -9.7% |  |  | **25.0** | 17.1 | 7.9 | 0 | **1.6** | 0.05 | 1.5 | 0 |
| NNVC-8.0 NNSR | **-7.8%** | -11.8% | -10.5% |  |  | **45.3** | 17.1 | 7.9 | 20.3 | **1.8** | 0.21 | 1.5 | 0.1 |

More details and analysis for tools and tools combination is expected in AhG14 report.

**Input contributions**

There are 29 input contributions related to the AHG mandates. The list of input contributions is provided below.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Reporting (**2**) | | | | | | |
| [JVET-AK0023](https://jvet-experts.org/doc_end_user/current_document.php?id=14700) | 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) | |  | |
| [JVET-AK0050](https://jvet-experts.org/doc_end_user/current_document.php?id=14627) | [AHG11] [AHG14] Teleconference on NNVC | | | | [E. Alshina](mailto:elena.alshina@huawei.com), [F. Galpin](mailto:franck.galpin@interdigital.com) | |
| NNVC training materials (**1**) | | | | | | |
| [JVET-AK0255](https://jvet-experts.org/doc_end_user/current_document.php?id=15244) | [AHG11] Response to Call for training materials for neural network-based video coding tool development | | | | [Fan Zhang](mailto:fan.zhang@bristol.ac.uk), [David Bull](mailto:dave.bull@bristol.ac.uk), [Jakub Nawala](mailto:jakub.nawala@bristol.ac.uk), [Yuxuan Jiang](mailto:yuxuan.jiang@bristol.ac.uk), [Xiaoqing Zhu](mailto:xzhu@netflix.com), [Joel Sole](mailto:jsole@netflix.com), [Elena Alshina](mailto:elena.alshina@huawei.com) | |
| EE1 contributions (**8**) | | | | | | |
| [JVET-AK0077](https://jvet-experts.org/doc_end_user/current_document.php?id=15048) | EE1-2.1: RA/LDB Unified Reference Frame Synthesis for VVC Inter Coding | [Q. Qin](mailto:kippqin@163.com), [C. Jung (Xidian Univ.)](mailto:zhengzk@xidian.edu.cn) | | | | | |
| [JVET-AK0078](https://jvet-experts.org/doc_end_user/current_document.php?id=15049) | EE1-2.2: Transformer-Based Reference Frame Synthesis for VVC Inter Coding | [Q. Qin](mailto:kippqin@163.com), [C. Jung (Xidian Univ.)](mailto:zhengzk@xidian.edu.cn) | | | | | |
| [JVET-AK0093](https://jvet-experts.org/doc_end_user/current_document.php?id=15064) | EE1-1.5: Adaptive skip of LOP filtering based on boundary strength partitions | [H. Kwon](mailto:hspeedkwon@hanyang.ac.kr), [J. Seo](mailto:hyeseo@hanyang.ac.kr), [H. Ko (HYU)](mailto:hyunsuk@hanyang.ac.kr), [D. Kim](mailto:kimddng@etri.re.kr), [S.-C. Lim (ETRI)](mailto:sclim@etri.re.kr) | | | | | |
| [JVET-AK0106](https://jvet-experts.org/doc_end_user/current_document.php?id=15077) | EE1-1.1: Partial Convolution and Over-Parameterization | [J. Chi](mailto:jingwei-chi@std.uestc.edu.cn), [A. Li](mailto:aoli@std.uestc.edu.cn), [C. Zhu](mailto:eczhu@uestc.edu.cn), [L. Luo](mailto:luolei1010@uestc.edu.cn), [H. Guo (UESTC)](mailto:hwguo@uestc.edu.cn), [Y. Huo](mailto:yongkai.huo@transsion.com), [Y. Liu (Transsion)](mailto:yutian.liu@transsion.com) | | | | | |
| [JVET-AK0139](https://jvet-experts.org/doc_end_user/current_document.php?id=15110) | EE1-1.3: Multiscale blocks in LOP4 and VLOP3 filters | [R. Yang](mailto:ruiying.yang@nokia.com), [M. Santamaria](mailto:maria.santamaria_gomez@nokia.com), [F. Cricri](mailto:francesco.cricri@nokia.com), [H. Zhang](mailto:honglei.1.zhang@nokia.com), [J. Lainema](mailto:jani.lainema@nokia.com), [M. M. Hannuksela (Nokia)](mailto:miska.hannuksela@nokia.com), [Y. Li](mailto:yli30@qti.qualcomm.com), [D. Rusanovskyy](mailto:dmytror@qti.qualcomm.com), S. Eadie, M. Karczewicz, J. Wang, L. Kerofsky (Qualcomm), [J. Chi](mailto:jingwei-chi@std.uestc.edu.cn), [A. Li](mailto:aoli@std.uestc.edu.cn), [C. Zhu](mailto:eczhu@uestc.edu.cn), [L. Lei](mailto:luolei1010@uestc.edu.cn), [H. Guo (UESTC)](mailto:hwguo@uestc.edu.cn), [Y. Huo](mailto:yongkai.huo@transsion.com), [Y. Liu (Transsion)](mailto:yutian.liu@transsion.com) | | | | | |
| [JVET-AK0150](https://jvet-experts.org/doc_end_user/current_document.php?id=15121) | EE1-1.2: LOP with improved Attention and residual groups | [Y. Li](mailto:yli30@qti.qualcomm.com), [D. Rusanovskyy](mailto:dmytror@qti.qualcomm.com), [S. Eadie](mailto:seadie@qti.qualcomm.com), [M. Karczewicz](mailto:martak@qti.qualcomm.com), [J. Wang](mailto:jianfwa@qti.qualcomm.com), [L. Kerofsky (Qualcomm)](mailto:kerofsky@qti.qualcomm.com) | | | | | |
| [JVET-AK0195](https://jvet-experts.org/doc_end_user/current_document.php?id=15166) | EE1-1.4: Cross-component enhanced LOP filter | [Y. Li](mailto:yue.li@bytedance.com), J. Li, C. Lin, K. Zhang, L. Zhang (Bytedance) | | | | | |
| [JVET-AK0235](https://jvet-experts.org/doc_end_user/current_document.php?id=15224) | EE1-4.1: A Neural Network Downscaling Filter for RPR | [Y.-Q. Zhu](mailto:yiqingzhu@hust.edu.cn), [W.-X. He](mailto:wxhe@hust.edu.cn), [X. Li](mailto:lx59497@163.com), [J.-D. Y](mailto:ye%20jd@hust.edu.cn), [Q. Liu (HUST)](mailto:q.liu@hust.edu.cn), [Z. Lv (vivo)](mailto:zhuoyi.lv@vivo.com) | | | | | |
| New NNVC (AhG11 or EE1 related) contributions (**8**) | | | | | | |
| [JVET-AK0071](https://jvet-experts.org/doc_end_user/current_document.php?id=15042) | AHG11: Lightweight Multiscale Reference Frame Generation for VVC Inter Coding | | [P. Li](mailto:2415355621@qq.com), [C. Jung](mailto:zhengzk@xidian.edu.cn), [Q. Qin (Xidian Univ.)](mailto:kippqin@163.com) | | |
| [JVET-AK0079](https://jvet-experts.org/doc_end_user/current_document.php?id=15050) | EE1-Related: Combination Test of NNVC Tools and TRFS | | [Q. Qin](mailto:kippqin@163.com), [C. Jung (Xidian Univ.)](mailto:zhengzk@xidian.edu.cn) | | |
| [JVET-AK0119](https://jvet-experts.org/doc_end_user/current_document.php?id=15090) | AHG11: NNLF LOP3 improvement with parallel 1x3/3x1 Backbone | | [T. Shao](mailto:tshao@dolby.com), [P. Yin](mailto:pyin@dolby.com), S. McCarthy (Dolby), [J. N. Shingala](mailto:jay.shingala@ittiam.com), A. Shyam, A. Suneja, S. Badya (Ittiam) | | |
| [JVET-AK0146](https://jvet-experts.org/doc_end_user/current_document.php?id=15117) | [AHG11] A Hybrid Framework Integrating End-to-End Learned Image Codec with Conventional Codec | | [N. Zou](mailto:nannan.zou@nokia.com), [A. Hallapuro](mailto:antti.hallapuro@nokia.com), [F. Cricri](mailto:francesco.cricri@nokia.com), [H. Zhang](mailto:honglei.1.zhang@nokia.com), [M. M. Hannuksela (Nokia)](mailto:miska.hannuksela@nokia.com) | | |
| [JVET-AK0175](https://jvet-experts.org/doc_end_user/current_document.php?id=15146) | AhG11: Dimension-wise decomposed multiplier for content-adaptive loop-filter | | [Z. Xu](mailto:zhuowei.xu@tcl.com), [J. Konieczny (TCL)](mailto:jacek.k@tcl.com) | | |
| [JVET-AK0177](https://jvet-experts.org/doc_end_user/current_document.php?id=15148) | AHG 11: Neural Network Coded Reference Frame for Intra Coding | | [Fabian Brand](mailto:fabian.brand@huawei.de), [Timofey Solovyev](mailto:solovyev.timofey@huawei.com), [Elena Alshina (Huawei)](mailto:elena.alshina@huawei.com) | | |
| [JVET-AK0242](https://jvet-experts.org/doc_end_user/current_document.php?id=15231) | AHG11: Enhancing super-resolution with residual in NNVC | | [T. Yang](mailto:y_tian@hust.edu.cn), [X. Li](mailto:lx59497@gmail.com), [W.-X. He](mailto:wxhe@hust.edu.cn), [Y.-Q. Zhu](mailto:yiqingzhu@hust.edu.cn), [Q. Liu (HUST)](mailto:q.liu@hust.edu.cn), [Z.-Y. Lv (vivo)](mailto:zhuoyi.lv@vivo.com) | | |
| [JVET-AK0258](https://jvet-experts.org/doc_end_user/current_document.php?id=15247) | AHG9/AHG11: Demo of real-time NNPF inference | | J. Funnell (Nokia), M. Santamaria, F. Cricri, M. M. Hannuksela, M. Pedzisz, R. Yang, [S. Schwarz](mailto:sebastian.schwarz@nokia.com) | | |
| Cross-checks (**6** all not yet uploaded at the time report was prepared) | | | | | | |  | |
| [JVET-AK0166](https://jvet-experts.org/doc_end_user/current_document.php?id=15137) | Crosscheck of JVET-AK0139 (EE1-1.3 : multiscale blocks in LOP4 and VLOP3 filters) MISSED | | | [T. Dumas (InterDigital)](mailto:thierry.dumas@interdigital.com) | | |
| [JVET-AK0216](https://jvet-experts.org/doc_end_user/current_document.php?id=15205) | Crosscheck of JVET-AK0195 (EE1-1.4: Cross-component enhanced LOP filter) MISSED | | | [Y. Li (Qualcomm)](mailto:yli30@qti.qualcomm.com) | | |
| [JVET-AK0241](https://jvet-experts.org/doc_end_user/current_document.php?id=15230) | Crosscheck of JVET-AK0093 (EE1-1.5: Adaptive skip of LOP filtering based on boundary strength partitions) | | | [J. Pardo](mailto:johan.esprit.pardo1@huawei.com), [E. Alshina (Huawei)](mailto:elena.alshina@huawei.com) | | |
| [JVET-AK0247](https://jvet-experts.org/doc_end_user/current_document.php?id=15236) | Crosscheck of JVET-AK0235 (EE1-4.1: A Neural Network Downscaling Filter for RPR) | | | [Z. Li](mailto:zylee@stu.pku.edu.cn), J. Zhang(PKU) | | |
| [JVET-AK0263](https://jvet-experts.org/doc_end_user/current_document.php?id=15252) | Crosscheck of JVET-AK0150 (EE1-1.2: LOP with improved Attention and residual groups) | | | [Z.Ameur](mailto:zoubida.ameur@interdigital.com), T.Dumas, [F.Galpin (Interdigital)](mailto:franck.galpin@interdigital.com), | | |
| [JVET-AK0292](https://jvet-experts.org/doc_end_user/current_document.php?id=15281) | Crosscheck of JVET-AK0106 (EE1-1.1 Partial Convolution and Over-Parameterization) | | | [J. Wang (Qualcomm)](mailto:jianfwa@qti.qualcomm.com) | | |

**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 and investigate training stability.
* Consider an up-date of NNVC training set (with re-training of existing NNVC tools)

[JVET-AK0012](https://jvet-experts.org/doc_end_user/current_document.php?id=15179) 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-AJ2024). The combined improvements of the ECM-15.0 over VTM-11.0ecm15.0 anchorfor AI, RA and LB configurations are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | All Intra Main10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -14.31% | -15.16% | -26.55% | 1136.4% | 528.0% |
| Class A2 | -20.82% | -23.47% | -27.99% | 1131.2% | 568.7% |
| Class B | -14.38% | -21.88% | -19.73% | 1052.3% | 601.4% |
| Class C | -14.30% | -11.21% | -12.34% | 1007.8% | 551.3% |
| Class E | -18.63% | -21.83% | -20.07% | 995.7% | 623.0% |
| Overall | -16.13% | -18.65% | -20.66% | 1058.7% | 575.3% |
| Class D | -12.16% | -8.19% | -8.92% | 990.4% | 585.2% |
| Class F | -29.93% | -33.46% | -33.65% | 744.5% | 672.5% |
| Class TGM | -43.12% | -48.77% | -48.02% | 576.4% | 704.0% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Random Access Main 10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -26.85% | -23.26% | -35.35% | 1016.7% | 1018.4% |
| Class A2 | -30.05% | -33.22% | -38.66% | 1001.2% | 1201.5% |
| Class B | -24.46% | -31.49% | -28.80% | 933.8% | 1077.1% |
| Class C | -26.33% | -21.66% | -22.37% | 1003.4% | 1179.8% |
| Class E |  |  |  |  |  |
| Overall | -26.55% | -27.57% | -30.37% | 981.8% | 1115.4% |
| Class D | -27.14% | -22.16% | -23.42% | 949.0% | 1295.2% |
| Class F | -32.70% | -34.86% | -35.64% | 870.5% | 827.1% |
| Class TGM | -42.33% | -47.83% | -47.63% | 735.6% | 656.0% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Low Delay B Main 10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -21.69% | -35.54% | -31.98% | 964.9% | 896.7% |
| Class C | -24.13% | -24.52% | -26.26% | 917.1% | 969.2% |
| Class E | -21.40% | -25.57% | -24.70% | 861.1% | 592.9% |
| Overall | -22.43% | -29.37% | -28.26% | 922.1% | 829.8% |
| Class D | -25.45% | -25.45% | -26.27% | 933.8% | 1110.6% |
| Class F | -30.25% | -38.19% | -37.94% | 824.3% | 729.5% |
| Class TGM | -40.56% | -50.17% | -50.03% | 740.4% | 622.4% |

The rate reduction for natural sequences over VTM in RA configuration for {Y, U, V} increased from ECM-14.0’s {-25.94%, -26.96%, -29.84%} to ECM-15.0’s {-26.55%, -27.57%, -30.37%}. For SCC sequences (class TGM) the rate reduction for RA configuration increased from ECM-14.0’s { -41.56%, -47.16%, -46.90%} to ECM-15.0’s { -42.33%, -47.83%, -47.63%}.

1. **Contributions**

In addition to 39 EE2 contributions, 35 (comparing to 50 last meeting) EE2-related and AHG12-related contributions were received. The EE2-related and AHG12-related contributions can be subdivided as follows:

***Intra (8)***

JVET-AK0089, "Non-EE2: Improvement on MPM", G. Wang, C. Zhou, Z. Lv (vivo)

JVET-AK0090, "Non-EE2: IntraTMP complexity reduction", Z. Fan, T. Chujoh, T. Ikai (Sharp)

JVET-AK0171, "Non-EE2: Picture-level mirroring and rotation", D. Mieloch, M. Lorkiewicz, J. Stankowski (PUT)

JVET-AK0173, "Non-EE2: Adaptive picture-level vertical mirroring", D. Mieloch, M. Lorkiewicz, A. Dziembowski, J. Stankowski, D. Karwowski (PUT)

JVET-AK0202, "AHG12: On cross-component intra prediction", Y.-J. Chang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AK0223, "Non-EE2: CCP merge mode with adjustment", Y. Wang, K. Zhang, W. Yin, Z. Deng, N. hang, L. Zhao, M. Salehifar, L. Zhang (Bytedance)

JVET-AK0234, "Non-EE2: Subblock-based CCCM", F. Pu, N. C. Thuong, P. Yin, S. McCarthy (Dolby)

JVET-AK0262, "EE2-related: On EIP filter shapes", K. Panusopone, M. He, S. Hong, L. Wang, J. Lainema (Nokia)

***Inter (12)***

JVET-AK0096, "Non-EE2: Additional inter merge candidates", N. Zhang, K. Zhang, Z. Deng, L. Zhao, Y. Wang, W. Yin, M. Salehifar, L. Zhang (Bytedance)

JVET-AK0103, "Non-EE2: Template matching based jointly reordering for GPM split modes and partition indexes", C. Ma, X. Xiu, X. Wang (Kwai)

JVET-AK0108, "Non-EE2: CMVP extension for constructed affine merge candidates", C. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba)

JVET-AK0111, "Non-EE2: On interpolation filter for template matching", Z. Dai, R.-L. Liao, J. Chen, X. Li, Y. Ye (Alibaba)

JVET-AK0133, "Non-EE2: Improvements on the BV-based prediction in SGPM", J. Huo, Y. Fei, L. Wang, Y. Ma, F. Yang (Xidian Univ.)

JVET-AK0162, "AHG12: On affine motion compensation", H. Huang, R. Yu, Z. Zhang, Y. Zhang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AK0186, "Non-EE2: Subblock-based spatial MVP ", Z. Deng, K. Zhang, N. Zhang, L. Zhao, Y. Wang, W. Yin, M. Salehifar, L. Zhang (Bytedance)

JVET-AK0192, "Non-EE2: On Sample-Based BDOF", M. Salehifar, Y. He, K. Zhang, L. Zhao, Y. Wang, N. Zhang, Z. Deng, W. Yin, L. Zhang (Bytedance)

JVET-AK0199, "Non-EE2: MV refinement for TMVP", Z. Zhang, J.-L. Lin, Y. Zhang, H. Huang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AK0213, "AHG12: Extended BDOF usage for MV refinement", R. Yu, H. Huang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AK0225, "Non-EE2: Extension on spatial and temporal merge candidates", J.-L. Lin, P.-H. Lin, Z. Zhang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AK0254, "Case study of candidates in AMVP candidate list", Seori Park, Hyunki Jeong, Byeungwoo Jeon (SKKU)

***In Loop Filters (7)***

JVET-AK0113, "AHG12: On CCALF", L. Xu, Y. Yu, H. Yu, N. Song, D. Wang (OPPO)

JVET-AK0117, "Non-EE2: Neural network based loop filtering for ECM", Y. Du, A. Li, J. Liu, C. Zhu, L. Luo, H. Guo (UESTC), Y. Huo, Y. Liu (Transsion)

JVET-AK0132, "AHG12: Boundary strength adjustment for DBV mode", J. Huo, J. Fan, Y. Fei, J.Liu, Y. Ma, F. Yang (Xidian Univ.)

JVET-AK0184, "EE2-related: Additional results for NN-based ILF in ALF", D. Rusanovskyy, Y. Li, M. Karczewicz (Qualcomm)

JVET-AK0200, "Non-EE2: CCSAO with reused CTU control and extended edge classifier", C.-W. Kuo, X. Xiu, X. Wang (Kwai)

JVET-AK0224, "Non-EE2: Cross-Chroma Adaptive Loop-Filter", W. Yin, K. Zhang, H. Liu, Y. Wang, Z. Deng, N. Zhang, L. Zhao, M. Salehifar, L. Zhang (Bytedance)

JVET-AK0237, "Non-EE2: CCALF with dual-chroma inputs", C. Ma, X. Xiu, C.-W. Kuo, X. Wang (Kwai)

***Transform (2)***

JVET-AK0214, "Non-EE2: IntraNN NSPT set", G. Verba, M. Coban, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AK0215, "Non-EE2: MTS with reduced memory ", B. Ray, V. Seregin, M. Karczewicz (Qualcomm)

***Entropy Coding (3)***

JVET-AK0135, "AhG12: CABAC contexts retraining F. Galpin (Interdigital)

JVET-AK0174, "Non-EE2: Optimization of probability estimation in CABAC", D. Karwowski, D. Mieloch, M. Lorkiewicz (PUT)

JVET-AK0279, "Non-EE2: On coefficient level binarization in Transform Skip", [M. Abdoli](mailto:mabdoli@xiaomi.com), R. G. Youvalari, F. Plowman, A. Tissier (Xiaomi)

***Partitioning (1)***

JVET-AK0219, "Non-EE2: Chroma partition prediction in separate tree condition", P.-H. Lin, J.-L. Lin, V. Seregin, M. Karczewicz (Qualcomm)

***CTC and Requirements (2)***

JVET-AK0201, "On additional operation point for the exploration model", T. Solovyev, J. Pardo, J. Sauer, Z. Li, E. Alshina (Huawei)

JVET-AK0232, “On real-time test condition for limited uplink/storage use case and MTT depth suggestions on common test condition”, [T. Ikai (Sharp)](mailto:ikai.tomohiro@sharp.co.jp)

1. Recommendations

The AHG recommends to:

* To review all the related contributions.

[JVET-AK0013](https://jvet-experts.org/doc_end_user/current_document.php?id=15180) 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**

This period was slow and limited to of the second edition of the technical report. The output document was addressed as was discussion on additional topics. A gitlab project was created on git.mpeg.expert to track issues and edits to the technical report [1].

A discussion was held about other issues. The group determined that additional updates and/or fixes should be included in an input document. Examples are the order of precedence and cropping order.

Finally, it is reiterated that the following five items should be in the second edition:

* Updated TR text to fix errors and clumsy wording
* Updated tools for the TR
* Video test sequences specific to film grain to be added to the TR
* Metrics – discussion and/or survey of available techniques or metrics
* Subjective testing of Film Grain

The following are worthy of discussion for the second edition:

* Conformance – this is interesting due to the process tied around an SEI message
* Analysis – preprocessing analysis (parameter estimation)

These are additional topics for consideration and discussion:

* The possible creation of Engineering Guidelines (EG) or Recommended Practice(RP) relating to this topic – This raises the question of what would be useful information for such guiding documents?
* Future parameters and signaling – not a v2 topic

1. **Related contributions**

Five contributions related to AHG13 were identified as of 01/13/2025.

* One contribution was the AHG report:
  + JVET-AK0013 JVET AHG report: Film grain technologies (AHG13)
* Four other contributions were uploaded at the time of the report drafting:
  + JVET-AK0058 AHG9/AHG13: Comments on Film grain regions characteristics SEI message
  + JVET-AK0169 AHG13: film grain synthesis improvement using frequency shaping
  + JVET-AK0197 AHG13: on picture size for the interpretation of the FGC SEI message
  + JVET-AK0211 AHG9/AHG13: Proposed Film Grain Region SEI message for version 5 of VSEI

***Contributions***

There were four contributions uploaded other than the AHG report.

**JVET-AK0058 AHG9/AHG13: Comments on Film grain regions characteristics SEI message**

While the model adaptation method proposed in JVET-AH0166 [1] was adopted into the VSEI TuC doc, i.e. JVET-AJ2032 [2], some model-adaptive modifications to the related syntax and semantics in the Film Grain Regions Characteristics (FGR) SEI message are proposed in this contribution to ensure the standard usage of adaptive film grain technology.

**JVET-AK0169 AHG13: film grain synthesis improvement using frequency shaping**

This contribution reports about the implementation of a soft frequency window instead of a rectangular frequency cutoff, when using film grain synthesis based on the FGC SEI message in frequency-filtering mode. This is said to visually improve the synthetic grain.

**JVET-AK0197 AHG13: on picture size for the interpretation of the FGC SEI message**

This contribution proposes to amend the VVC specification to refer to the SPS instead of the PPS to define picture resolution for the interpretation of the film grain characteristics SEI message, the intent being to describe film grain characteristics at source picture resolution independently of encoder decisions to perform reduced-resolution coding.

It is also proposed to indicate that unspecified cropping and padding can be applied as needed.

**JVET-AK0211 AHG9/AHG13: Proposed Film Grain Region SEI message for version 5 of VSEI**

This contribution proposes to include the Film Grain Region (FGR) SEI message into the working draft of Version 5 of VSEI, should it be established at this meeting. The FGR SEI message is meant to be available for HEVC and VVC.

1. **Recommendations**

The AHG recommends:

* Request opening a second edition for the TR;
* the related input contributions are reviewed;
* any liaisons are reviewed;
* testing of FGS be discussed;
* continued conformance discussion;
* SEI message extensions;
* discussion topics for an EG/RP; and
* continue the study of film grain technologies in JVET.

The text of JVET-AJ2020 is under preparation, could become basis of the ITU output.

[JVET-AK0014](https://jvet-experts.org/doc_end_user/current_document.php?id=15022) JVET AHG report: NNVC software development (AHG14) [F. Galpin (chair), R. Chang, Y. Li, Y. Li, M. Santamaria, J. N. Shingala, Z. Xie (vice-chairs)]

1. **Software development**

***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, enabling deblocking in the RDO, JVET-AH0054 with improved MCTF, JVET-AI0124 for reference picture alignment.

NNVC-11.0 anchor at https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/nnvc-ctc is used for NNVC performance evaluation.

***Software changes***

***NNVC-11.0***

Several commits were merged in the NNVC repository. The following changes were integrated:

|  |  |  |
| --- | --- | --- |
| **Merge request** | **Contribution** | **Description** |
| 280/284/286 | JVET-AJ0066 | inference for reduced inputs |
| 272 | JVET-AJ0056 | wavelet loss NNSR |
| 283 | JVET-AJ0054 | training scripts fix |
| 281 | JVET-AJ0054 | adding models |
| 273 | JVET-AJ0054 | scripts + NNVC code |
| 269 | JVET-AJ0124 | QP per block |
| 271 | JVET-AJ0066 | split block for HOP5 |
| 285/274 | JVET-AJ0125 | SADL update |
| 277 | fix | memory error |
| 279 | fix | memory error |
| 275 | fix | gitlab CI |
| 276 | fix | xLOP backward compatiblity |
| 270 | fix | dual tree data extraction |
| 265 | fix | backward compatible SPS |

In SADL, the following changes were integrated:

|  |  |  |
| --- | --- | --- |
| **Merge request** | **Contribution** | **Description** |
| 126 |  | more info on function calls |
| 125 | JVET-AJ0054 related | conv2D to support pad 0 and improve performance (no skip border anymore) |
| 121/137 |  | speed up for LOP |
| 115 | JVET-AJ0361 | add code to get layers anaysis |
| 118 | JVET-AJ0125 | info in model |
| 112 | JVET-AJ0122 | sparse packed |
| 111 | JVET-AJ0106 | batchnorm experimental |
| 131 | fix | gcc 13 bug |
| 119 | fix | MAC count for sparse packed |
| 117 | fix | allow multiple load of models |
| 116 | fix | correct stack overflow |
| 113 | fix | fix large kernel mismatch |
| 128/136 | fix | correct CI and build |

***Software version***

NNVC-11.0 was tagged December 10th, 2024 (bit accurate with 11rc)

NNVC-11.0rc was tagged December 3rd, 2024

NNVC-10.0 was tagged August 9th, 2024

NNVC-9.1 was tagged May 28th, 2024

NNVC-9.0 was tagged May 13th, 2024

NNVC-8.0 was tagged February 20th, 2024

NNVC-8.0rc3 was tagged February 20th, 2024.

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.

Note: For InterDigital results, the encoding times might not be correct since encoding was done before a bit accurate update of the models. Decoding time are using the last version of NNVC-11.0, using an homogenous cluster with some speed variation.

***Comparison to VTM***

**NNVC-10.0 VTM vs NNVC-11.0 VTM**

VTM configuration of NNVC is not changed between version 10.0 and 11.0.

Note: 1 redundant bit in SPS is removed between the 2 versions (see MR 265).

**NNVC-11.0 VTM vs NNVC-11.0 anchor**

The NNVC-11.0 anchor includes LOP.4 filter and Intra Prediction tools activated.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -8.89% | -13.05% | -14.28% | -9.43% | -14.76% | -16.00% | 119% | 3884% |
| Class A2 | -7.88% | -14.40% | -9.69% | -8.36% | -15.17% | -9.53% | 117% | 3665% |
| Class B | -7.29% | -15.44% | -14.36% | -7.96% | -17.45% | -16.84% | 120% | 3744% |
| Class C | -6.91% | -13.77% | -13.51% | -7.78% | -14.78% | -15.09% | 117% | 3336% |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -7.63% | -14.31% | -13.19% | -8.29% | -15.74% | -14.74% | 118% | 3642% |
| Class D | -7.13% | -13.31% | -12.30% | -6.87% | -14.56% | -13.06% | 113% | 3052% |
| Class F | -4.02% | -8.28% | -7.29% | -5.03% | -10.94% | -10.87% | 122% | 1712% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
|  |  |  |  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 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 | -5.85% | -11.54% | -11.01% | -6.96% | -11.74% | -13.13% | 107% | 3709% |
| Class C | -5.60% | -11.57% | -10.75% | -7.02% | -10.21% | -9.09% | 102% | 3169% |
| Class E | -6.52% | -4.30% | -7.75% | -8.50% | -1.92% | -6.25% | 109% | 3579% |
| **Overall** | -5.94% | -9.74% | -10.11% | -7.36% | -8.78% | -10.07% | 106% | 3488% |
| Class D | -6.33% | -9.37% | -8.05% | -7.09% | -4.97% | -6.54% | 98% | 3003% |
| Class F | -3.33% | -5.41% | -5.64% | -5.52% | -5.27% | -7.43% | 107% | 1819% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
|  |  |  |  |  |  |  |  |  |
|  | **Low delay P Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 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 | -6.32% | -14.44% | -13.28% | -7.35% | -16.25% | -16.87% | 111% | 3619% |
| Class C | -6.00% | -13.66% | -12.66% | -7.21% | -12.92% | -13.31% | 103% | 3033% |
| Class E | -7.36% | -5.63% | -8.77% | -9.47% | -4.59% | -7.92% | 113% | 3656% |
| **Overall** | -6.47% | -11.98% | -11.95% | -7.83% | -12.23% | -13.45% | 109% | 3421% |
| Class D | -6.61% | -11.17% | -9.28% | -7.42% | -7.90% | -8.60% | 99% | 2774% |
| Class F | -3.64% | -6.93% | -6.80% | -5.49% | -8.16% | -11.72% | 108% | 1679% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
|  |  |  |  |  |  |  |  |  |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -9.19% | -15.70% | -16.76% | -9.37% | -18.57% | -18.15% | 171% | 2838% |
| Class A2 | -7.96% | -15.42% | -12.22% | -8.53% | -16.09% | -11.91% | 170% | 2466% |
| Class B | -7.86% | -15.36% | -16.39% | -8.27% | -17.73% | -18.36% | 172% | 2431% |
| Class C | -7.71% | -14.17% | -14.72% | -8.57% | -17.34% | -17.80% | 154% | 1875% |
| Class E | -11.22% | -17.07% | -18.08% | -11.90% | -17.26% | -19.04% | 166% | 2611% |
| **Overall** | -8.63% | -15.45% | -15.67% | -9.17% | -17.43% | -17.24% | 166% | 2389% |
| Class D | -7.61% | -12.34% | -13.39% | -7.97% | -16.35% | -17.00% | 146% | 1679% |
| Class F | -5.39% | -9.85% | -9.70% | -5.81% | -12.75% | -13.41% | 136% | 1955% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |

Note: Results from Interdigital, crosschecked by Qualcomm (AI/RA).

**NNVC-11.0 VTM mode vs NNVC-11.0 HOP.5**

The NNVC-11.0 where LOP.4 filter is replaced by HOP.5 and Intra Prediction tools activated.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -15.60% | -13.69% | -21.60% | -17.26% | -19.06% | -23.79% | 276% | 123461% |
| Class A2 | -15.13% | -20.72% | -26.68% | -14.49% | -20.24% | -23.74% | 257% | 112524% |
| Class B | -13.10% | -24.21% | -17.36% | -11.95% | -22.58% | -16.70% | 272% | 117302% |
| Class C | -13.92% | -17.30% | -16.86% | -12.52% | -14.58% | -14.20% | 215% | 103138% |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -14.22% | -19.57% | -19.94% | -13.67% | -19.28% | -18.86% | 253% | 113563% |
| Class D | -14.83% | -17.41% | -17.34% | -11.03% | -12.96% | -11.40% | 208% | 98862% |
| Class F | -8.99% | -11.35% | -10.48% | -9.24% | -12.20% | -10.80% | 371% | 53532% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
|  |  |  |  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 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 | -9.82% | -12.00% | -6.93% | -9.30% | -8.96% | -4.82% | 291% | #NUM! |
| Class C | -11.65% | -9.46% | -6.96% | -10.48% | -4.78% | -0.10% | 232% | 109681% |
| Class E | -10.33% | -10.68% | -5.57% | -10.01% | -10.57% | -4.81% | 588% | 101925% |
| **Overall** | -10.56% | -10.82% | -6.60% | -9.87% | -7.97% | -3.24% | 321% | #NUM! |
| Class D | -12.84% | -7.60% | -0.50% | -10.53% | 1.16% | 14.54% | 223% | 109822% |
| Class F | -9.00% | -9.70% | -8.81% | -10.13% | -10.56% | -8.72% | 411% | 66305% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
|  |  |  |  |  |  |  |  |  |
|  | **Low delay P Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 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 | -11.07% | -15.39% | -9.62% | -10.31% | -12.22% | -7.20% | 358% | #NUM! |
| Class C | -12.58% | -11.57% | -9.68% | -10.96% | -6.13% | -2.93% | 275% | 105881% |
| Class E | -11.64% | -10.69% | -4.73% | -11.55% | -10.77% | -3.90% | 748% | 111964% |
| **Overall** | -11.72% | -12.94% | -8.42% | -10.84% | -9.83% | -4.95% | 394% | #NUM! |
| Class D | -13.93% | -10.90% | -5.70% | -11.51% | -1.74% | 9.66% | 265% | 101795% |
| Class F | -9.78% | -12.06% | -11.22% | -10.72% | -13.43% | -12.18% | 484% | 65005% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
|  |  |  |  |  |  |  |  |  |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -13.53% | -10.72% | -17.20% | -15.00% | -16.58% | -19.47% | 346% | 103002% |
| Class A2 | -13.56% | -17.53% | -21.27% | -14.30% | -18.54% | -18.26% | 252% | 82417% |
| Class B | -12.13% | -16.96% | -14.32% | -12.03% | -17.06% | -14.79% | 241% | 77815% |
| Class C | -13.07% | -14.27% | -16.16% | -12.79% | -13.89% | -15.46% | 188% | 52191% |
| Class E | -16.66% | -19.95% | -18.84% | -16.38% | -18.82% | -19.54% | 253% | 88218% |
| **Overall** | -13.56% | -15.91% | -17.12% | -13.80% | -16.81% | -17.09% | 246% | 76922% |
| Class D | -12.19% | -12.68% | -15.79% | -11.24% | -12.70% | -14.68% | 172% | 47123% |
| Class F | -9.96% | -12.44% | -10.41% | -9.68% | -12.47% | -10.99% | 168% | 65649% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |

Note: Results from InterDigital, crosschecked by xxx.

**NNVC-11.0 VTM mode vs NNVC-10.0 VLOP**

The NNVC-11.0 where LOP.4 filter is replaced by VLOP3 and Intra Prediction tools activated.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -7.18% | -5.57% | -5.74% | -7.66% | -6.78% | -6.23% | 116% | 1811% |
| Class A2 | -5.92% | -6.40% | -4.16% | -6.51% | -7.01% | -3.96% | 115% | 1643% |
| Class B | -5.45% | -7.56% | -6.50% | -6.25% | -8.21% | -7.34% | 118% | 1689% |
| Class C | -4.95% | -6.49% | -5.83% | -6.01% | -7.79% | -6.45% | 116% | 1504% |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -5.75% | -6.65% | -5.70% | -6.52% | -7.57% | -6.20% | 116% | 1651% |
| Class D | -4.87% | -5.48% | -4.91% | -5.46% | -6.03% | -5.38% | 112% | 1396% |
| Class F | -2.93% | -3.48% | -3.22% | -3.86% | -5.59% | -4.53% | 118% | 773% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
|  |  |  |  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 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 | -3.93% | -6.35% | -3.35% | -5.26% | -8.10% | -7.25% | 106% | 1608% |
| Class C | -3.46% | -5.60% | -4.05% | -5.42% | -9.16% | -5.37% | 105% | 1414% |
| Class E | -4.05% | -1.92% | -1.25% | -6.47% | -4.57% | -2.74% | 107% | 1401% |
| **Overall** | -3.80% | -4.99% | -3.06% | -5.61% | -7.57% | -5.50% | 106% | 1488% |
| Class D | -3.93% | -4.00% | -1.90% | -5.24% | -9.44% | -4.95% | 103% | 1364% |
| Class F | -1.89% | -1.37% | -1.46% | -3.53% | -4.22% | -3.77% | 107% | 803% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
|  |  |  |  |  |  |  |  |  |
|  | **Low delay P Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 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 | -4.11% | -7.42% | -4.07% | -5.08% | -8.88% | -8.13% | 109% | 1489% |
| Class C | -3.75% | -6.75% | -4.65% | -5.32% | -9.92% | -5.75% | 105% | 1314% |
| Class E | -4.60% | -2.57% | 0.10% | -7.08% | -5.17% | -1.87% | 108% | 1448% |
| **Overall** | -4.11% | -5.98% | -3.22% | -5.66% | -8.30% | -5.77% | 108% | 1418% |
| Class D | -4.16% | -5.15% | -2.02% | -5.21% | -9.32% | -6.05% | 104% | 1243% |
| Class F | -1.96% | -1.86% | -2.16% | -3.57% | -5.44% | -6.06% | 108% | 751% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
|  |  |  |  |  |  |  |  |  |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -8.24% | -8.99% | -8.35% | -8.36% | -10.77% | -9.22% | 166% | 1475% |
| Class A2 | -6.80% | -8.35% | -6.85% | -7.16% | -8.55% | -6.21% | 168% | 1333% |
| Class B | -6.73% | -8.80% | -8.72% | -7.11% | -9.70% | -9.50% | 167% | 1339% |
| Class C | -6.25% | -7.86% | -7.59% | -7.16% | -9.37% | -8.65% | 156% | 1109% |
| Class E | -9.34% | -10.06% | -9.98% | -10.08% | -9.77% | -9.18% | 166% | 1400% |
| **Overall** | -7.32% | -8.76% | -8.30% | -7.83% | -9.63% | -8.67% | 165% | 1314% |
| Class D | -6.10% | -6.89% | -6.85% | -6.61% | -8.78% | -8.50% | 144% | 1041% |
| Class F | -4.31% | -5.60% | -5.58% | -4.75% | -7.00% | -6.89% | 134% | 1008% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |

Note: Results from Interdigital, crosschecked by Qualcomm (RA/AI)

**NNVC-11.0 VTM vs NNVC-11.0 NNSR**

Note: only class A sequences are changed, LDx configuration are thus omitted.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -12.07% | -10.17% | -10.56% | -19.18% | -21.21% | -20.88% |  |  |
| Class A2 | -8.86% | -6.49% | -1.95% | -17.40% | -15.13% | -14.57% |  |  |
| Class B | -7.29% | -15.44% | -14.36% | -7.96% | -17.45% | -16.84% |  |  |
| Class C | -6.91% | -13.77% | -13.51% | -7.78% | -14.78% | -15.09% |  |  |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -8.46% | -12.15% | -10.89% | -12.04% | -17.02% | -16.73% |  |  |
| Class D | -7.13% | -13.31% | -12.30% | -6.87% | -14.56% | -13.06% |  |  |
| Class F | -4.02% | -8.28% | -7.29% | -5.03% | -10.94% | -10.87% |  |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -10.98% | -9.96% | -11.73% | -13.82% | -13.69% | -13.78% |  |  |
| Class A2 | -11.00% | -4.54% | -1.25% | -17.09% | -6.78% | -2.52% |  |  |
| Class B | -7.86% | -15.36% | -16.39% | -8.27% | -17.73% | -18.36% |  |  |
| Class C | -7.71% | -14.17% | -14.72% | -8.57% | -17.34% | -17.80% |  |  |
| Class E | -11.22% | -17.07% | -18.08% | -11.90% | -17.26% | -19.04% |  |  |
| **Overall** | -9.43% | -12.68% | -13.00% | -11.34% | -15.07% | -14.95% |  |  |
| Class D | -7.61% | -12.34% | -13.39% | -7.97% | -16.35% | -17.00% |  |  |
| Class F | -5.39% | -9.85% | -9.70% | -5.81% | -12.75% | -13.41% |  |  |

Note: Results from HUTS, crosschecked by xxx.

***Comparison to NNVC-11.0 anchor***

**NNVC-10 anchor vs NNVC-11.0 anchor**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -0.14% | -1.20% | -1.87% | 0.00% | -1.34% | -1.49% | 102% | 114% |
| Class A2 | -0.24% | -1.88% | -1.32% | -0.07% | -1.92% | -1.01% | 102% | 112% |
| Class B | -0.34% | -0.75% | -2.09% | -0.28% | -1.45% | -1.99% | 102% | 111% |
| Class C | -0.33% | 0.00% | -1.67% | -0.17% | -0.07% | -2.14% | 102% | 108% |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -0.28% | -0.86% | -1.78% | -0.15% | -1.15% | -1.73% | 102% | 111% |
| Class D | -0.46% | -0.72% | -1.40% | -0.14% | -1.62% | -1.06% | 101% | 101% |
| Class F | -0.12% | -0.44% | -0.96% | 0.06% | 0.38% | -1.42% | 103% | 113% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
|  |  |  |  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 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 | -0.40% | -1.34% | -2.79% | -0.46% | -1.96% | 0.03% | 106% | 121% |
| Class C | -0.32% | 0.26% | -1.75% | -0.02% | -0.15% | -0.92% | 103% | 110% |
| Class E | -0.60% | 0.71% | -3.87% | -0.48% | 1.63% | -1.07% | 102% | 112% |
| **Overall** | -0.42% | -0.29% | -2.72% | -0.32% | -0.46% | -0.56% | 104% | 115% |
| Class D | -0.48% | 0.33% | -0.41% | -0.15% | 0.04% | 2.64% | 101% | 108% |
| Class F | -0.19% | 1.21% | -3.04% | 0.07% | 1.53% | -0.77% | 104% | 116% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
|  |  |  |  |  |  |  |  |  |
|  | **Low delay P Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 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 | -0.41% | -1.50% | -3.21% | -0.51% | -2.89% | -0.80% | 97% | 108% |
| Class C | -0.37% | 0.34% | -1.61% | -0.01% | 0.02% | -1.67% | 94% | 102% |
| Class E | -0.54% | 0.95% | -4.45% | -0.11% | 1.37% | 0.39% | 96% | 103% |
| **Overall** | -0.43% | -0.28% | -2.99% | -0.24% | -0.86% | -0.79% | 96% | 105% |
| Class D | -0.36% | 0.85% | 0.73% | -0.07% | 1.09% | 4.08% | 94% | 100% |
| Class F | -0.22% | 0.77% | -3.27% | 0.08% | 1.77% | -1.80% | 94% | 100% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
|  |  |  |  |  |  |  |  |  |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | 0.05% | -1.88% | -2.31% | -0.01% | -2.32% | -2.26% | 101% | 113% |
| Class A2 | -0.14% | -1.87% | -1.37% | -0.07% | -2.11% | -1.46% | 100% | 116% |
| Class B | -0.19% | -1.41% | -2.32% | -0.16% | -2.03% | -2.44% | 100% | 112% |
| Class C | -0.27% | -0.86% | -1.98% | -0.20% | -1.58% | -2.60% | 99% | 115% |
| Class E | -0.40% | -1.29% | -2.22% | -0.17% | -1.36% | -2.25% | 98% | 110% |
| **Overall** | -0.19% | -1.42% | -2.07% | -0.13% | -1.88% | -2.25% | 100% | 113% |
| Class D | -0.37% | -0.74% | -1.60% | -0.27% | -1.86% | -2.31% | 97% | 108% |
| Class F | -0.24% | -0.43% | -1.15% | -0.18% | -0.52% | -1.94% | 103% | 113% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |

Note: Results from InterDigital, crosschecked by xxx.

***NNVC-11.0 vs NNVC-11.0 HOP5***

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -7.35% | -1.03% | -8.61% | -8.61% | -5.52% | -9.43% | 232% | 3179% |
| Class A2 | -7.85% | -7.27% | -18.72% | -6.64% | -5.85% | -15.80% | 219% | 3070% |
| Class B | -6.29% | -10.73% | -3.66% | -4.32% | -6.65% | 0.16% | 226% | 3133% |
| Class C | -7.55% | -4.04% | -3.85% | -5.13% | 0.33% | 1.20% | 184% | 3092% |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -7.15% | -6.32% | -7.71% | -5.86% | -4.40% | -4.67% | 214% | 3118% |
| Class D | -8.31% | -4.78% | -5.96% | -4.44% | 1.98% | 1.84% | 184% | 3239% |
| Class F | -5.26% | -3.37% | -3.45% | -4.48% | -1.48% | 0.17% | 303% | 3127% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
|  |  |  |  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 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 | -4.23% | -0.53% | 4.35% | -2.53% | 3.32% | 9.29% | 271% | #NUM! |
| Class C | -6.42% | 2.52% | 4.20% | -3.74% | 6.39% | 10.14% | 228% | 3461% |
| Class E | -4.07% | -6.15% | 2.40% | -1.61% | -8.13% | 1.77% | 542% | 2848% |
| **Overall** | -4.92% | -0.92% | 3.81% | -2.70% | 1.48% | 7.70% | 304% | #NUM! |
| Class D | -6.95% | 1.51% | 8.19% | -3.71% | 6.34% | 22.55% | 227% | 3657% |
| Class F | -5.91% | -4.45% | -3.40% | -4.90% | -5.39% | -1.33% | 385% | 3646% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
|  |  |  |  |  |  |  |  |  |
|  | **Low delay P Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 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 | -5.10% | -1.19% | 3.88% | -3.20% | 4.82% | 11.08% | 324% | #NUM! |
| Class C | -7.02% | 2.51% | 3.22% | -4.07% | 7.91% | 11.49% | 267% | 3491% |
| Class E | -4.62% | -4.93% | 4.43% | -2.23% | -6.10% | 4.33% | 664% | 3062% |
| **Overall** | -5.62% | -0.89% | 3.80% | -3.25% | 3.12% | 9.53% | 363% | #NUM! |
| Class D | -7.84% | 0.30% | 3.90% | -4.42% | 7.44% | 18.18% | 268% | 3670% |
| Class F | -6.43% | -5.48% | -4.78% | -5.56% | -5.55% | -0.50% | 447% | 3872% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
|  |  |  |  |  |  |  |  |  |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -4.76% | 6.00% | -0.40% | -6.14% | 2.71% | -1.60% | 203% | 3629% |
| Class A2 | -6.10% | -1.81% | -10.06% | -6.25% | -2.09% | -7.04% | 148% | 3342% |
| Class B | -4.66% | -1.69% | 2.57% | -4.07% | 0.95% | 4.58% | 140% | 3201% |
| Class C | -5.86% | 0.04% | -1.69% | -4.60% | 4.53% | 2.95% | 122% | 2783% |
| Class E | -6.15% | -3.46% | -0.87% | -4.99% | -1.85% | -0.58% | 152% | 3379% |
| **Overall** | -5.43% | -0.34% | -1.55% | -5.05% | 1.06% | 0.39% | 148% | 3220% |
| Class D | -4.98% | -0.38% | -2.93% | -3.57% | 4.63% | 2.55% | 117% | 2806% |
| Class F | -4.92% | -2.91% | -0.75% | -4.10% | 0.39% | 3.08% | 123% | 3358% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |

***NNVC-11.0 vs NNVC-11.0 VLOP***

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | 1.88% | 9.08% | 10.00% | 1.96% | 10.11% | 11.74% | 97% | 47% |
| Class A2 | 2.14% | 9.96% | 6.25% | 2.00% | 10.40% | 6.23% | 98% | 45% |
| Class B | 2.00% | 9.53% | 9.27% | 1.85% | 11.67% | 11.62% | 98% | 45% |
| Class C | 2.11% | 8.60% | 9.03% | 1.91% | 8.58% | 10.47% | 99% | 45% |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | 2.03% | 9.28% | 8.75% | 1.92% | 10.28% | 10.26% | 98% | 45% |
| Class D | 2.44% | 9.11% | 8.66% | 1.50% | 10.26% | 9.13% | 99% | 46% |
| Class F | 1.17% | 5.49% | 4.47% | 1.24% | 6.42% | 7.53% | 97% | 45% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
|  |  |  |  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 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 | 2.05% | 5.96% | 8.56% | 1.84% | 4.36% | 6.98% | 99% | 43% |
| Class C | 2.28% | 6.82% | 7.66% | 1.75% | 1.28% | 4.50% | 104% | 45% |
| Class E | 2.64% | 2.49% | 7.27% | 2.17% | -2.59% | 4.22% | 99% | 39% |
| **Overall** | 2.27% | 5.38% | 7.94% | 1.89% | 1.59% | 5.46% | 100% | 43% |
| Class D | 2.57% | 5.76% | 7.05% | 2.01% | -3.04% | 2.51% | 105% | 45% |
| Class F | 1.50% | 4.33% | 4.47% | 2.12% | 1.20% | 4.10% | 101% | 44% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
|  |  |  |  |  |  |  |  |  |
|  | **Low delay P Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 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 | 2.36% | 8.27% | 10.58% | 2.46% | 8.96% | 10.61% | 99% | 41% |
| Class C | 2.40% | 8.05% | 9.34% | 2.05% | 3.44% | 8.79% | 102% | 43% |
| Class E | 2.96% | 3.20% | 9.79% | 2.57% | -0.49% | 6.82% | 96% | 40% |
| **Overall** | 2.53% | 6.93% | 9.97% | 2.35% | 4.76% | 9.06% | 99% | 41% |
| Class D | 2.62% | 7.10% | 8.13% | 2.40% | 0.58% | 3.11% | 105% | 45% |
| Class F | 1.76% | 5.52% | 5.01% | 2.06% | 3.28% | 6.59% | 100% | 45% |
| Class H | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
|  |  |  |  |  |  |  |  |  |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | 1.04% | 8.43% | 10.37% | 1.10% | 10.43% | 11.25% | 98% | 52% |
| Class A2 | 1.26% | 8.85% | 6.31% | 1.48% | 9.53% | 6.54% | 99% | 54% |
| Class B | 1.24% | 7.89% | 9.24% | 1.24% | 10.09% | 10.98% | 98% | 55% |
| Class C | 1.61% | 7.61% | 8.62% | 1.53% | 10.19% | 11.56% | 101% | 59% |
| Class E | 2.13% | 8.61% | 10.19% | 2.00% | 9.37% | 12.42% | 100% | 54% |
| **Overall** | 1.44% | 8.20% | 8.96% | 1.45% | 9.96% | 10.65% | 99% | 55% |
| Class D | 1.65% | 6.40% | 7.88% | 1.47% | 9.31% | 10.72% | 99% | 62% |
| Class F | 1.17% | 4.99% | 4.74% | 1.12% | 7.10% | 7.88% | 99% | 52% |

***NNVC-11.0 vs NNVC-11.0+NNSR***

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -3.24% | 3.32% | 4.31% |  |  |  |  |  |
| Class A2 | -0.85% | 8.35% | 7.86% |  |  |  |  |  |
| Class B | 0.00% | 0.00% | 0.00% |  |  |  |  |  |
| Class C | 0.00% | 0.00% | 0.00% |  |  |  |  |  |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -0.82% | 2.33% | 2.43% |  |  |  |  |  |
| Class D | 0.00% | 0.00% | 0.00% |  |  |  |  |  |
| Class F | 0.00% | 0.00% | 0.00% |  |  |  |  |  |
| Class H | #VALUE! | #VALUE! | #VALUE! |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | |  |  |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | |  |  |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU | DecT GPU | PSNR Overlap |
| Class A1 | -1.86% | 6.35% | 5.82% | -4.71% | 5.66% | 5.37% | 131% | #NUM! | #DIV/0! | 97% |
| Class A2 | -3.19% | 11.65% | 11.54% | -9.14% | 9.88% | 9.54% | 95% | #NUM! | #DIV/0! | 99% |
| Class B | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 100% | #DIV/0! | 100% |
| Class C | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 100% | #DIV/0! | 100% |
| Class E | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 100% | #DIV/0! | 100% |
| **Overall** | -0.84% | 3.00% | 2.89% | -2.31% | 2.59% | 2.49% | 104% | #NUM! | #DIV/0! | 99% |
| Class D | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 100% | #DIV/0! | 100% |
| Class F | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 100% | #DIV/0! | 100% |

Results from HUTS, crosschecked by xxx

1. **Contributions**

We have 3 contributions for AhG14 and 1 telco report.

|  |  |  |  |
| --- | --- | --- | --- |
| [JVET-AK0134](https://jvet-experts.org/doc_end_user/current_document.php?id=15105) | AhG14: SADL update | | [F. Galpin (Interdigital)](mailto:franck.galpin@interdigital.com) |
| [JVET-AK0151](https://jvet-experts.org/doc_end_user/current_document.php?id=15122) | AhG14: On the impact of explicit full SIMD implementation for 1x3 and 3x1 DW convolution | [Y. Li](mailto:yli30@qti.qualcomm.com), [D. Rusanovskyy](mailto:dmytror@qti.qualcomm.com), [M. Karczewicz (Qualcomm)](mailto:martak@qti.qualcomm.com), | |
| [JVET-AK0257](https://jvet-experts.org/doc_end_user/current_document.php?id=15246) | AHG14: The extension of SADL library | [N. Fu](mailto:nianxiangfu@whu.edu.cn), [W. Bao](mailto:baoweijie@whu.edu.cn), [Z. Chen (Wuhan Univ.)](mailto:zzchen@whu.edu.cn) | |

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 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 |
| NNVC VTM mode | none | encoder\_xxx\_vtm.cfg |
| NNVC Anchor/EE1 | Intra Pred + LOP.4 | encoder\_xxx\_nnvc.cfg |
| NNVC. HOP | Intra Pred + HOP.5 | encoder\_xxx\_nnvc.cfg + nn-based/HOP5.cfg |
| NNVC. VLOP | Intra Pred + VLOP.3 | encoder\_xxx\_vtm.cfg + nn-based/vlop3.cfg |
| SR | Super-resolution | nn-based/nnsr.cfg + nn-based/nnsr\_classAx.cfg |
| PF | Adaptive post-filters | nn-based/nnpf/nnpf\_xxx.cfg |
| ALOP | Intra pred+adaptive LOP | encoder\_xxx\_vtm.cfg +nn-based/intra.cfg+xxx |

Deprecated options:

|  |  |  |
| --- | --- | --- |
| set0 | Loop filter set #0 | encoder\_xxx\_vtm.cfg + nn-based/NnlfOption\_1.cfg |
| set1 | Loop filter set #1 | encoder\_xxx\_vtm.cfg + nn-based/NnlfOption\_2.cfg |
| set0+rdo | Loop filter set #0 + Rdo | nn-based/NnlfOption\_1.cfg + --EncNnlfOpt=1 |
| set1+rdo+intra+temporal filter | Loop filter set #1 + Temporal filter | nn-based/NnlfOption\_2.cfg  + --NnlfSet1Multiframe=1 |

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-AK0015](https://jvet-experts.org/doc_end_user/current_document.php?id=15181) JVET AHG report: Gaming content compression (AHG15) [S. Puri, J. Sauer (co-chairs), R. Chernyak, A. Duenas, L. Wang (vice chairs)]

**Work in between meetings**

* Anchor generation
* Software repositories (at <https://vcgit.hhi.fraunhofer.de/jvet-ahg-gcc>)
  + Support of reading depth and motion maps (JVET-AJ0300)
  + Update to latest ECM version
* Added repository for scripts for verification of auxiliary information for gaming (JVET-AJ0137)
* Bug reported in ECM (<https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/issues/84>)
  + Encoder crash for RA on BaoleiYard, QPs 22, 27, 32, 37 and 42.

**Test results**

Test results were obtained according to JVET-AJ2027. The list of sequences/classes is copied for convenience below:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Class** | **Sequence name** | **Frame count** | **Low delay frame count** | **Frame rate** | **Bit depth** | **Intra** | **Random access** | **Low-delay** |
| G1 (SDR) | Level1\_SDR | 600 | 300 | 60 | 10 | M | M | M |
| Darktree\_SDR | 600 | 300 | 60 | 10 | M | M | M |
| ARPG2\_SDR | 600 | 300 | 60 | 10 | M | M | M |
| DesertTown3\_SDR | 600 | 300 | 60 | 10 | M | M | M |
| SunTemple3\_SDR | 600 | 300 | 60 | 10 | M | M | M |
| G2 | GTAV | 600 | 300 | 60 | 8 | M | M | M |
| Minecraft | 600 | 300 | 60 | 8 | M | M | M |
| JianlingTemple | 600 | 300 | 60 | 8 | M | M | M |
| BaoleiYard | 300 | - | 60 | 8 | M | M | - |
| G3 (HDR) | Level1\_HDR | 600 | 300 | 60 | 10 | O | O | O |
| Darktree\_HDR | 600 | 300 | 60 | 10 | O | O | O |
| ARPG2\_HDR | 600 | 300 | 60 | 10 | O | O | O |
| DesertTown3\_HDR | 600 | 300 | 60 | 10 | O | O | O |
| SunTemple3\_HDR | 600 | 300 | 60 | 10 | O | O | O |

For class G3 the HDR versions of Level1 and Darktree are still missing.

***VTM-11.0ecm15.0 vs ECM-15.0***

**SDR**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | | |
|  | **Over Anchor** | | | | | |
|  | Y | U | V | EncT | DecT | VmPeak |
| Class G1 | -11.33% | -23.56% | -18.28% | 895.65% | 348.38% | #REF! |
| Class G2 | -18.36% | -22.51% | -26.64% | 1240.13% | 691.59% | #REF! |
| **Overall** | -14.45% | -23.09% | -21.99% | 1035.02% | 472.50% | #REF! |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access Main 10** | | | | | |
|  | **Over Anchor** | | | | | |
|  | Y | U | V | EncT | DecT | VmPeak |
| Class G1 | -27.08% | -42.28% | -34.12% | 671.18% | 1193.40% | #REF! |
| Class G2\* | -23.98% | -35.97% | -37.25% | 869.45% | 1515.22% | #REF! |
| **Overall** | -25.92% | -39.91% | -35.29% | 739.59% | 1305.18% | #REF! |

\* ECM crashed for BaoleiYard. Results are averaged without BaoleiYard

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Low delay B Main 10** | | | | | |
|  | **Over Anchor** | | | | | |
|  | Y | U | V | EncT | DecT | VmPeak |
| Class G1 | -20.63% | -52.33% | -41.62% | 686.55% | 848.99% | #REF! |
| Class G2\* | -18.83% | -55.86% | -55.24% | 1035.63% | 1397.60% | #REF! |
| **Overall** | -20.11% | -53.34% | -45.51% | 772.11% | 978.93% | #REF! |

\* JianlingTemple did not finish in time for ECM. It has been excluded from averages.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Low delay P Main 10** | | | | | |
|  | **Over Anchor** | | | | | |
|  | Y | U | V | EncT | DecT | VmPeak |
| Class G1 | -18.94% | -58.18% | -49.65% | 592.71% | 856.06% | #REF! |
| Class G2 | -15.62% | -54.08% | -54.41% | 923.03% | 1205.41% | #REF! |
| **Overall** | -17.70% | -56.64% | -51.43% | 699.82% | 973.28% | #REF! |

**HDR**

There are still no HDR versions for Level1 and Darktree. Due to a configuration error AI was run without using HDR config, hence results for AI are omitted.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access** | | | | | | | | | |
|  | **Over VTM11.0ecm15.0** | | | | | | | | | |
|  |  |  | **wPSNR** |  |  | **PSNR** |  |  |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class G3 | -40.39% | -26.60% | -27.79% | -57.87% | -56.17% | -27.99% | -58.41% | -56.40% | 774% | 1435% |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Low Delay B** | | | | | | | | | |
|  | **Over VTM11.0ecm15.0** | | | | | | | | | |
|  |  |  | **wPSNR** |  |  | **PSNR** |  |  |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class G3 | -36.84% | -20.85% | -20.09% | -62.36% | -57.80% | -20.36% | -63.00% | -57.71% | 914% | 1602% |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Low Delay P** | | | | | | | | | |
|  | **Over VTM11.0ecm15.0** | | | | | | | | | |
|  |  |  | **wPSNR** |  |  | **PSNR** |  |  |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class G3 | -36.67% | -20.56% | -19.78% | -62.15% | -56.97% | -20.02% | -62.71% | -56.94% | 888% | 1158% |

**Input contributions**

* AHG15: Analysis of ECM 12-bit internal bit depth in gaming content (JVET-AK0082)
* AHG15: Analysis of high-level control of intra prediction methods in gaming content   
  (JVET-AK0083)
* AHG15: On the compression of depthmaps from auxiliary data (JVET-AK0170)

**Recommendations**

The AHG recommends to:

* Review input documents on gaming content compression
* Identify if/how auxiliary information can be used for coding of gaming content
* Further work on unification of auxiliary data format

[JVET-AK0016](https://jvet-experts.org/doc_end_user/current_document.php?id=15182) JVET AHG report: Generative face video compression (AHG16) [Y. Ye (chair), H.-B. Teo, Z. Lyu, S. McCarthy, S. Wang (vice chairs)]

1. **AHG activities**

Regarding the mandate on developing and maintaining the GFVC software, the AHG16 GFVC software tool and accompanying usage instructions and exemplar configurations for experimentation are maintained in the GIT repository at <https://vcgit.hhi.fraunhofer.de/jvet-ahg-gfvc>. During this AHG period, the AHG16 software was updated by replacing the existing CFTE, FOMM, FV2V, and DAC models with respective multi-resolution models as provided in JVET-AJ0052. Implementation of the GFV SEI message following the syntax defined in the VSEI version 4 draft 4 in JVET-AI2006 has been merged into VTM-23.6.

Regarding testing of higher resolution content, GFVC test conditions and evaluation procedures as defined in JVET-AJ2035 have been updated to include additionally 33 test sequences at 512x512 resolution.

Regarding coordination with AHG9 to develop the GFV SEI message, the GFV and GFVE SEI messages have been integrated into VSEI version 4 draft 4 JVET-AJ2006. At this meeting, 8 input contributions addressing various high-level syntax aspects of the GFV and GFVE SEI messages have been received.

1. **Related contributions**

The following two input contributions to this meeting are related to the activities of AHG16:

1. [JVET-AK0068](file:///C:\Users\Yan.Ye\Desktop\2025-01-Geneva\AHG%20reports\current_document.php%3fid=15039), AHG16: GFVC Extension of the VVC Standard [L. Liu, C. Jung (Xidian Univ.)]
2. [JVET-AK0069](file:///C:\Users\Yan.Ye\Desktop\2025-01-Geneva\AHG%20reports\current_document.php%3fid=15040), AHG16: QP-Adaptive GFVC [W. Kang, L. Liu, C. Jung (Xidian Univ.)]

The following input contributions are related to the high-level syntax aspects of the generative face video (GFV) and/or generative face video enhancement (GFVE) SEI messages:

1. [JVET-AK0080](file:///C:\Users\Yan.Ye\Desktop\2025-01-Geneva\AHG%20reports\current_document.php%3fid=15051), AHG9: Comments on the GFV and GFVE SEI messages [M. M. Hannuksela (Nokia), J. Chen, B. Chen, Y. Ye (Alibaba)]
2. [JVET-AK0124](file:///C:\Users\Yan.Ye\Desktop\2025-01-Geneva\AHG%20reports\current_document.php%3fid=15095), AHG9: On timing information and order of pictures in GFV SEI message [J. Lee, H. Tan, C. Kim, J. Nam, J. Lim, S. Kim (LGE)]
3. [JVET-AK0127](file:///C:\Users\Yan.Ye\Desktop\2025-01-Geneva\AHG%20reports\current_document.php%3fid=15098), AHG9: AHG9: On miscellaneous aspects of GFV and DSCI SEI messages [J. Lee, H. Tan, C. Kim, J. Nam, J. Lim, S. Kim (LGE)]
4. [JVET-AK0128](file:///C:\Users\Yan.Ye\Desktop\2025-01-Geneva\AHG%20reports\current_document.php%3fid=15099), AHG9: Editorial updates for GFV SEI message [J. Lee, H. Tan, C. Kim, J. Nam, J. Lim, S. Kim (LGE)]
5. [JVET-AK0154](file:///C:\Users\Yan.Ye\Desktop\2025-01-Geneva\AHG%20reports\current_document.php%3fid=15125), AHG: Comments on Generative Face Video SEI [A. C. Sidiya, S. Deshpande (Sharp)]
6. [JVET-AK0164](file:///C:\Users\Yan.Ye\Desktop\2025-01-Geneva\AHG%20reports\current_document.php%3fid=15135), AHG9: Supplementation of value range definition and editorial bugs fixes on the GFVE pupil position SEI messages [F. Ma, A. Trioux, F. Yang (Xidian Univ.), B. Li, F. Xing, Z. Wang (Hisense)]
7. [JVET-AK0238](file:///C:\Users\Yan.Ye\Desktop\2025-01-Geneva\AHG%20reports\current_document.php%3fid=15227), AHG9: Semantics and syntax fixes for generative face video SEI message [J. Chen, B. Chen, Y. Ye (Alibaba)]
8. [JVET-AK0239](file:///C:\Users\Yan.Ye\Desktop\2025-01-Geneva\AHG%20reports\current_document.php%3fid=15228), AHG9: On generative face video enhancement SEI message [J. Chen, B. Chen, Y. Ye (Alibaba)]
9. **Recommendations**

The AHG recommends to:

* Review related contributions, including those on the high-level syntax aspects of the GFV and GFVE SEI messages;
* To continue AHG16 to study GFVC-related topics.

[JVET-AK0017](https://jvet-experts.org/doc_end_user/current_document.php?id=15183) JVET AHG report: Testing of video coding technology beyond CTC (AHG17) [M. Wien (chair), Y. Ye, V. Baroncini, E. Alshina (vice chairs)]

**Activities**

Two AHG calls were held on 2024-11-20 and 2024-12-18, respectively. The reports of these meetings are available in JVET-AK0042.

***Identification of test sequences and preparation of bitstreams***

The set of sequences to be explored was determined including sequences discussed during breakout activities at the last JVET meeting and new content. The sequences were assigned to volunteers for the encoding task and the corresponding RD data and bitstreams were collected.

The discussion and the results of this activity are reported in JVET-AK0042.

The set of sequences consists of nine classes including SDR\_RA\_UHD/4K, SDR\_RA\_HD, SDR\_LD\_HD, HDR\_RA\_UHD, HDR\_RA\_8Kcrop, Gaming\_LD\_HD. And UGC\_RA, making a total of 24 UHD/4K sequences and 27 sequences at HD resolution. For distribution of the encoding tasks, the intended QP range was split into interleaved subset for the VTM and the ECM, respectively. The sets contain one overlapping QP each for crosschecking purposes. The overall encoding effort comprised a total of 1152 rate points with 504 point at UHD/4K and 648 points at HD resolution (384 ECM bitstreams and 768 VTM bitstreams).

|  |  |  |
| --- | --- | --- |
| **Resolution** | **QP set Name** | **Values** |
| HD | VTM QP set 1 | QP=20:4:48 |
| HD | VTM QP set 2 | QP=22:4:50, 36 |
| HD | ECM QP set 1 | QP=22:8:46 |
| HD | ECM QP set 2 | QP=26:8:50, 38 |
| UHD | VTM QP set 1 | QP=26:4:50 |
| UHD | VTM QP set 2 | QP=28:4:52, 38 |
| UHD | ECM QP set 1 | QP=28:8:52 |
| UHD | ECM QP set 2 | QP=32:8:48, 36 |

Before the meeting, 363 out of 384 ECM bitstreams and 732 out of 768 were uploaded (about 95% for each of the test models.

Input contributions JVET-AK0041, JVET-AK0043, JVET-AK0044, and JVET-AK0045 have been covered in the AHG phase.

**Related contributions**

|  |  |  |
| --- | --- | --- |
| [JVET-AK0041](https://jvet-experts.org/doc_end_user/current_document.php?id=15011) | AHG17: Materials for assisting sequence selection for Non-CTC testing | [J. Pardo](mailto:johan.esprit.pardo@huawei.com), [J. Sauer (Huawei)](mailto:johannes.sauer@huawei.com) |
| [JVET-AK0042](https://jvet-experts.org/doc_end_user/current_document.php?id=15012) | AHG17: AhG meeting notes | [M. Wien](mailto:wien@lfb.rwth-aachen.de) |
| [JVET-AK0043](https://jvet-experts.org/doc_end_user/current_document.php?id=15013) | AHG17: Additional sequences for testing of video coding technology beyond CTC | [R.-L. Liao](mailto:ruling.lrl@alibaba-inc.com), [Y. Ye (Alibaba)](mailto:yan.ye@alibaba-inc.com) |
| [JVET-AK0044](https://jvet-experts.org/doc_end_user/current_document.php?id=15014) | AHG17: Cropped HDR sequences for non-CTC testing | [A. Filippov](mailto:alexey.filippov@tcl.com), [J. Konieczny](mailto:jacek.k@tcl.com), [Z. Zhi](mailto:zhou.zhi@tcl.com), [V. Rufitskiy](mailto:vasily.rufitskiy@tcl.com), [H. Qin](mailto:hongdong.qin@tcl.com), [T. Dong](mailto:tianyu.dong@tcl.com), [X. Tang (TCL)](mailto:xiaoyong.tang@tcl.com), [S. Shen](mailto:s.shen@nercdtv.org), [Y. Qin](mailto:yf.qin@nercdtv.org), [Y. Wang](mailto:y.wang@nercdtv.org), [X. Wang](mailto:xs.wang@nercdtv.org), [Y. Guan (NERC-DTV)](mailto:yf.guan@nercdtv.org) |
| [JVET-AK0045](https://jvet-experts.org/doc_end_user/current_document.php?id=15015) | AHG17: Test sequences from YouTube | [X. Li](mailto:xlxiangli@google.com), Y. Wang, N. Birkbeck, B. Adsumilli (Google) |
| [JVET-AK0066](https://jvet-experts.org/doc_end_user/current_document.php?id=15037) | AHG4/AHG17: 8K HDR sequences as a response to the call for new HDR test materials for future video coding | [Y. Li (CMG)](mailto:liyan@cctv.com), [Q. Zhang](mailto:zhangqian@abp2003.cn), [J. Ning (ABP)](mailto:ningjinhui@abp2003.cn) |
| [JVET-AK0196](https://jvet-experts.org/doc_end_user/current_document.php?id=15167) | AHG17: Additional sequences for testing of video coding technology beyond CTC | [J. Pardo](mailto:johan.esprit.pardo1@huawei.com), Y. Sun, T. Solovyev, E. Alshina (Huawei) |
| [JVET-AK0218](https://jvet-experts.org/doc_end_user/current_document.php?id=15207) | AHG17: UGC test sequences from Huawei | [Y. Zhao](mailto:yin.zhao@huawei.com), Y. Sun, P. Liu, Y. Lu, C. Wang, [E. Alshina (Huawei)](mailto:elena.alshina@huawei.com) |
| [JVET-AK0221](https://jvet-experts.org/doc_end_user/current_document.php?id=15210) | [AHG17] JVET CTC and non-CTC sequences performance analysis | [E. Alshina](mailto:elena.alshinaQhuawei.com), J. Pardo, J. Sauer, T. Solovyev (Huawei) |
| [JVET-AK0246](https://jvet-experts.org/doc_end_user/current_document.php?id=15235) | Crosscheck of JVET-AK0196 (AHG17: Additional sequences for testing of video coding technology beyond CTC) | [Z. Xiang (Tencent)](mailto:alanzxiang@global.tencent.com) |
| [JVET-AK0282](https://jvet-experts.org/doc_end_user/current_document.php?id=15271) | AHG4/AHG17: Additional HDR sequences for non-CTC testing | [A. Filippov](mailto:alexey.filippov@tcl.com), [J. Konieczny](mailto:jacek.k@tcl.com), [Z. Zhi](mailto:zhou.zhi@tcl.com), [V. Rufitskiy](mailto:vasily.rufitskiy@tcl.com), [H. Qin](mailto:hongdong.qin@tcl.com), [T. Dong](mailto:tianyu.dong@tcl.com), [X. Tang (TCL)](mailto:xiaoyong.tang@tcl.com), [S. Shen](mailto:s.shen@nercdtv.org), [Y. Qin](mailto:yf.qin@nercdtv.org), [Y. Wang](mailto:y.wang@nercdtv.org), [X. Wang](mailto:xs.wang@nercdtv.org), [Y. Guan (NERC-DTV)](mailto:yf.guan@nercdtv.org) |
|  |  |  |

**Recommendations**

The AHG recommends:

* To review the remaining input contributions related to AHG17 during the meeting.
* To review the provided bitstreams and the corresponding rate-distortion results in a BoG activity in order to assert correctness of the completed encoding tasks.
* To perform informal viewing sessions together with SC 29/AG 5 for checking the characteristics of the provided test sequences, with a focus on new or previously rarely used material, and the potential range of QPs suitable for conducting formal subjective tests.
* To prepare a proposed list of candidate sequences for each category for further consideration for testing.

At maximum roughly 60 sequences in total (51 from the coordinated activity, 6 from 0196 and 0218, 4 additional from an update of 0045)

* As first step, parties who have encoded the sequences shall share the proposed lowest and highest rate points for ECM in an update of the spread sheet by Wednesday morning, decoding those points
* First viewing with a smaller group of people for pre-selection (Wed. afternoon), also to investigate appropriateness of “voting methodology” about the correctness of the suggested rate points and suitability of the sequences.
* BoG to decide on reduction of the set (Thu.)
* First round of “real” test on Friday morning (after larger number of displays would be available)

An update of the status achieved was given by MW Thu. 16 Jan. 1400.

* All bitstreams decoded
* Displays arrived, to be set up
* ACR to be used for the assessment
* Initial viewing with smaller group, set of sequences to be viewed reduced from 59 to 51
* First viewing session Friday 0900, 24 volunteers needed
* Discussion on results on next steps Sat. late afternoon
* Lack of content showing conversational scenes

# Project development (37)

## AHG1: Development, deployment and advertisement of standards (1)

Contributions in this area were discussed at 0950–1025 on Sunday 19 Jan. 2025 (chaired by JRO).

[JVET-AK0178](https://jvet-experts.org/doc_end_user/current_document.php?id=15149) AHG9: Proposed plan for initiating version 5 of VSEI [A. T. Hinds, S. Wenger (Tencent), P. de Lagrange, E. François (InterDigital)]

(add abstract)

According to proponents, due to the ISO/IEC rule about not starting an amendment when another one is at committee stage (see discussion under opening plenary / AHG9 report), this contribution became somewhat obsolete. It was suggested instead to consider potentially promoting stable SEI messages from TuC into v4. Further to be discussed Monday morning.

See further notes under plenary section.

## AHG2: Text development and errata reporting (0)

This section is kept as a template for future use.

## 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 (0)

This section is kept as a template for future use.

## AHG4: Test and training material (0)

This section is kept as a template for future use.

Proposals on new test materials are also included in 4.17 and 5.1.4

## Codec performance under non-CTC conditions (5)

Contributions in this area were discussed at 1025–1345 on Sunday 19 Jan. 2025 (chaired by JRO).

[JVET-AK0057](https://jvet-experts.org/doc_end_user/current_document.php?id=15028) [AHG3]?: NAL unit loss software [S. Wenger, A. Hinds, G. Teniou (Tencent)]

Included herein and proposed for error resilience tests—if such were agreed by JVET—is a packet loss software package based on JCTVC-H0072. JCTVC-H0072 appears to be the last error resilience software in use by the J\* groups, at least based on the contributors’ investigation. The “loss” software parses Annex-B bitstreams (such as those created by our reference software), removes individual NAL units as indicated in an error pattern file, and writes the remaining NAL units. Vis-à-vis JCTVC-H0072, the software has been modified so to support 3 and 4 octet start codes, and the NAL unit types specified in VVC and related application logic. The “loss” software has been tested on a small ECM bitstream and it is asserted that no problems were detected, though the software can benefit from further tests. Also included is the converter from ITU-T Rec. G.1050-generated error pattern format to the error pattern format used herein. This software is untested. Finally, included as the error patterns that were used during some of the H.264 and H.265 error resilience tests.

It was asked what happens to the output of the tool when fed into ECM decoder? It crashes. Same with VTM, HM, JM (the latter in some earlier version had an error concealment).

It was commented that having such a software is useful. It was explained that “mode 4” only loses SEI NAL units (as in the old software).

[JVET-AK0193](https://jvet-experts.org/doc_end_user/current_document.php?id=15164) Low latency Wi-Fi transmission simulation and suggestions on codec requirements [S. Ikonin, X. Ma, I. Gribushin, M. Sychev, V. Khamidullin, R. Shabaev, E. Alshina (Huawei)]

This contribution illustrates video transmission performance simulated using wireless Wi-Fi network simulator under low latency constraint. It is observed that average and worst case reconstructed quality, and frame freeze ratio strongly depend on the latency requirements specified in the test. With the transmission latency restricted to 16 ms, corresponding to one frame interval at 60 fps, the video freeze ratio is more than 20%, and average reconstruction quality is degraded by 5.0 dB (from 39.1 to 34.1 dB). Due to low latency restriction and wireless channel fluctuation decoder cannot be always guaranteed to receive all the packets necessary for frame reconstruction in time. It is suggested to consider packet loss resilient reconstruction as a necessary functionality to support low latency requirement.

Simulator uses a pattern for losses (using NS-3, similar concept as JVET-AK0057), but additionally allows generating losses based on latency. Encoding/decoding delay is currently not considered.

Error patterns of various network types could be used.

A video demo is included in the contribution.

At decoder, a lost slice is discarded. Encoder-side error concealment is used, changing reference picture of subsequent pictures based on feedback that a slice was lost. This would require the feedback channel to be reliable. Loss of I picture is not considered currently. Other approaches of error concealment might be useful to be studied.

It is commented by various experts that the topic of JVET-AK0057 and JVET-AK0193 is highly relevant.

It was requested installing a new AHG (S. Ikonin to take) on ultra-low delay study.

Decision(SW): Create an AHG branch in VTM for the software of JVET-AK0057 and JVET-AK0193.

[JVET-AK0201](https://jvet-experts.org/doc_end_user/current_document.php?id=15188) On additional operation point for the exploration model [T. Solovyev, J. Pardo, J. Sauer, Z. Li, E. Alshina (Huawei)]

This document describes the benefits of presence in the exploration model additional operation point, corresponding to the real-time encoding scenario. Possible parameters of this operation point such as the ratio between encoding run time and real time are discussed. Potential strategies of implementation of such a new operation point are discussed.

Unclear whether the goal is optimization of ECM or VTM

It was asked why this information was not brought to AHG10

Implicit MTS extensions of ECM might be an example candidate for extending VVEnc faster preset for real-time encoding with marginal runtime increase and some benefit in compression

[JVET-AK0232](https://jvet-experts.org/doc_end_user/current_document.php?id=15221) On real-time test condition for limited uplink/storage use case and MTT depth suggestions on common test condition [T. Ikai (Sharp)]

The uploading of user-generated videos from smartphones presents a significant opportunity for entertainment. However, the uplink speed in wireless communication is limited. Concurrently, cloud storage is widely utilized by individuals; nevertheless, it becomes quickly overwhelmed when dealing with videos, forcing people to either abandon video storage (opting instead for still images) or incur significantly higher costs. To address the uplink and storage challenges, a higher compression efficiency in real-time encoding on smartphones, with limited complexity—such as 3x—is essential. Conversely, the current common test conditions are established using explosive partitioning patterns, encoding difficult tool set, and numerous distortion optimizations, without considering real-time encoding.

A preliminary experiment has shown that

* Reducing MTT depth provides a favorable balance of coding efficiency in terms of encoder complexity, and a reduced MTT depth with some tools disabled would constitute a good real-time test condition.
* The anticipated coding performance with new coding tools may yield 10% improvement at 3x EncT and 15% improvement at 5x EncT in real-time test conditions.
* In real-time conditions, the anchor should always include the vtm fast mode anchor as the second anchor in addition to the previous test model (e.g. ECM anchor) to check complexity - coding efficiency balance is on the way in terms of the former codec standard.

It is proposed to develop a new test condition as a real-time test condition.

It is additionally suggested to reduce MTT depth in the common condition as well for CfP and standard development to achieve a good coding efficiency balance and to facilitate the new real time test condition.

ECM is not suitable for developing real-time encoder at this moment, and .

At this moment, it may not be useful for VTM to investigate real-time encoding settings, as this work has already been done by VVEnc, for example.

There could be potential adding VVEnc as a close to RT encoding anchor in a CfP, but new proposals could also demonstrate reduced complexity benefit against VVC.

Lower encoding complexity is not necessarily low-delay or random access.

More discussion on requirements for next generation standard necessary, also to decide whether a lower-complexity scenario would be relevant for a CfP.

Difficult to assess if something is capable for real time encoding.

It was commented that it would be premature to define real-time conditions before a CfP, before it is known which potential tools enabling fast encoding should be investigated. For evaluation within the CfP, something like VVEnc could be used as a comparison point.

Real-time encoding is not only achieved by tools, but also overall implementation optimization of system (as done in VVEnc).

It was suggested to further study the proposed MTT configuration and potential other improvements in the current VTM (also providing crosscheck). This study can well be performed in the context of AHG10, and software maintained by AHG3, with the potential of adding an alternative config setting for faster encoding by the next meeting.

[JVET-AK0236](https://jvet-experts.org/doc_end_user/current_document.php?id=15225) Use cases and requirement of next generation video coding standard for broadcasting services [S. Iwamura, S. Nemoto, A. Ichigaya (NHK)]

(joint meeting)

[JVET-AK0338](https://jvet-experts.org/doc_end_user/current_document.php?id=15327) Gap analysis and suggested updates to next-generation video coding requirements draft [J. Samuelsson-Allendes, S. Deshpande, T. Ikai (Sharp)] [late]

(joint meeting)

## AHG5: Conformance test development (0)

This section is kept as a template for future use.

## AHG7: ECM tool assessment (6)

Contributions in this area were discussed at 1130–XXXX on Friday 17 Jan. 2025 (chaired by JRO).

[JVET-AK0047](https://jvet-experts.org/doc_end_user/current_document.php?id=15017) AHG7: Complexity and memory footprint assessment of PDP, Neural Network-based Intra Prediction, MTS and LFNST/NSPT, and ALF in ECM-15 [G. Boisson, T. Dumas, F. Galpin, S. Puri, K. Naser, C. Bonnineau (InterDigital)]

Was reviewed in conf call JVET-AK0048

[JVET-AK0048](https://jvet-experts.org/doc_end_user/current_document.php?id=15018) Report of AHG7 conference call [X. Li]

**Tool analysis**

1. PDP

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Total size of the lookup tables (KB) | Worst-case memory reading (byte/sample) | Worst-case multiplication per sample | Stage number of the worst case |
| JVET-AJ0372 | 7672.232 | 264 | 132 |  |
| JVET-AK0047 | 16bits per element  PDP conventional: 3869  PDP-MIP 3804 | - | - |  |
| Conclusion | 16bits per element  PDP conventional: 3869  PDP-MIP: 3804 | 264 | 132 | - |

1. NNIP

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Total size of the lookup tables (KB) | Worst-case memory reading (byte/sample) | Worst-case multiplication per sample | Worst Stage |
| JVET-AJ0372 | 2285 | 12102 | 3138 | - |
| JVET-AK0047 | 2439 kB using 16 bits for each element (current implementation, with sparsity considered)  2378 kB using 16 bits for a value of non-zero coefficient, 11 bits for the index of non-zero coefficient, and 10 bits for the number of non-zero coefficients | 9875 (4x4) for the ECM-15 | To add number with matrix multiplication only  3990 MAC/sample (current implementation with sparsity)  3138 MAC/sample for an algorithm fully exploiting the sparsity) | 3 for 4x4  4 for 16x16 |
| Conclusion | 2439 kB using 16 bits for each element (current implementation)  2378 kB using 16 bits for a value of non-zero coefficient, 11 bits for the index of non-zero coefficient, and 10 bits for the number of non-zero coefficients | 9875 (4x4) for the ECM-15 | To add number with matrix multiplication only  3990 MAC/sample (current implementation with sparsity)  3138 MAC/sample for an algorithm fully exploiting the sparsity) | 3 for 4x4  4 for 16x16 |

1. Enhanced MTS

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Total size of the lookup tables (KB) | Worst-case memory reading (byte/sample) | Worst-case multiplication per sample | Stage number of the worst case |
| JVET-AJ0372 | 1541.255  (Note that some parts of the table are not actually used) | 2 | To be calculated | 2 |
| JVET-AK0047 | 439.275 | 32.768 kB for 128x128=2 coefficient/sample  (only used transforms in CTC are counted) | 16-bit  (M x N) x (M + N) / (M x N)=M+N  256 (for 128x128) | 2 |
| JVET-AK0049 | Some transforms are in the software implementation but not used in CTC  Calculation is based on 16-bit  Assume only the coefficients for the supported transforms in Table 2 are stored for matrix multiplication and assume IDTR coefficients must be stored. The storage requirements are 21,840 x 4 + 336 x 2 = 88,032 coefficients, or 176,064 bytes  Assuming the IDTR transform can be performed without storing any coefficients. The storage requirements are 21,840 x 3 + 336 x 2 = 66,192 coefficients, or 132,384 bytes. | 2 coefficients = 4 bytes | Same as above | 2 |
| Qualcomm | If max MTS size is limited to 64, the total decoder memory (16-bit inverse transform) is 111.6 KB |  |  |  |
| Conclusion | No conclusion yet, to be further checked by InterDigital, OPPO and Qualcomm | 4 bytes | 256 | 2 |

1. LFNST extension

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Total size of the lookup tables (KB) | Worst-case memory reading (byte/sample) | Worst-case multiplication per sample | Stage number of the worst case |
| JVET-AJ0372 | 564.577 | 16 | 16 |  |
| JVET-AK0047 | 322.56  LFNST16 + LFNST8 + LFNST4:  322.56 + 215.04 + 26.88 kB = 564.48 kB | 8-bit  3.072 kB (for LFNST16) = 12  Comment: LFNST 8/4 are also used in CTC. But the number is not the worst case | LFNST 16x16: 32x96 = 3072 MACs=12 |  |
| Conclusion | 564.48 kB (to be double checked by LG) | 12 | 12 | 1 (DCT2 is not considered) |

1. NSPT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Total size of the lookup tables (KB) | Worst-case memory reading (byte/sample) | Worst-case multiplication per sample | Stage number of the worst case |
| JVET-AJ0372 | 3601.92 | 40 | 40 |  |
| JVET-AK0047 | 5921.28  CU level worst case 24x256=6.144 kB | NSPT 8X16/16X8: 5120  8-bit | 40 |  |
| Conclusion | 5921.28  CU level worst case 24x256=6.144 kB | NSPT 8X16/16X8: 5120  8-bit | 40 | 1 |

1. Fixed loop filter in ALF

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Total size of the lookup tables (KB) | Worst-case memory reading (byte/sample) | Worst-case multiplication per sample | Stage number of the worst case |
| JVET-AJ0372 | 1949.696  (Note that some tables only need 2 bits/element, the total size could be reduced by 40%) | 73 | 146 |  |
| JVET-AK0047 | Decoder: 1,950 kB (could be reduced to 935 kB) | CTU: 342 kB (could be reduced to 162 kB) / (128x128)=20.8 (could be reduced to 9.9 | AI: 41+64+12=117 MAC/sample  RA/LD: 41+64+12+21=138 MAC/sample  Chroma: 41  Total: AI 137.5  RA/LD: 158.5  Gaussian filter is considered |  |
| Conclusion | Decoder: 1,950 kB (could be reduced to 935 kB) | CTU: 342 kB (could be reduced to 162 kB) / (128x128)=20.8 (could be reduced to 9.9  (to be double check by Qualcomm, InterDigital and Google) | RA/LD: 146.5 (Gaussian filter is not considered) | Not discussed |

**Recommendations**

It is agreed to create an AHG7 ECM branch and merge the tool control into that branch. The MRs of the controls of LFNST extension, NSPT, and fixed filters in ALF have been merged into the branch. The MR of the controls of MTS extension are under review. This MR is expected to be merged soon.

It is agreed to focus the analysis on ECM-15 CTC.

Note that all MTS related memory size calculations are based on 16-bit and LFNST related memory calculations are based on 8-bit.

The tool analysis is summarized in the table below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Tool | Cfg control | Total size of the lookup tables (KB) | Worst-case memory reading (byte/sample) | Worst-case multiplication per sample | Stage number of the worst case |
| PDP | PDP | 16bits per element  PDP conventional: 3869  PDP-MIP: 3804 | 264 | 132 | - |
| NN based Intra prediction | IntraNN | 2439 kB using 16 bits for each element (current implementation, with sparsity considered)  2378 kB using 16 bits for a value of non-zero coefficient, 11 bits for the index of non-zero coefficient, and 10 bits for the number of non-zero coefficients | 9875 (4x4 in ECM-15 | 53kMAC/sample for 4x4 (Hypothetical dense matrices)  3990 MAC/sample (current implementation with sparsity)  3138 MAC/sample for an algorithm fully exploiting the sparsity) | 3 for 4x4  4 for 16x16 |
| Enhanced MTS | MaxMTSSize (cannot achieve the goal of completely disabling Enhanced MTS). | 307.112 KB (Including DCT2) | 4 bytes | 256 | 2 |
| LFNST extension | LFNSTExt | 564.48 kB | 12 | 12 | 1 (DCT2 is not considered) |
| NSPT | NSPT | 5921.28  CU level worst case 24x256=6.144 kB | 8-bit 8X16/16X8: 5120 | 40 | 1 |
| Fixed filter in ALF | AlfFixedFilter (Gaussian filter is not disabled. Filter 0, filter 1 and the fixed filter of residuals are controlled. The shapes of other filters are not impacted) | Decoder: 1,950 kB (could be reduced to 935 kB) | CTU: 342 kB (could be reduced to 162 kB) / (128x128)=20.8 (could be reduced to 9.9) | RA/LD: 146.5 (Gaussian filter is not considered) | 3 |

From discussion in JVET: Number of necessary sequential operations would more precisely expressing the latency caused by a tool, rather than number of stages, as the number of operations per stage may be different for each tool.

The criticality of the latency of a tool highly depends on its position in the processing pipeline, and dependencies, e.g. in intra prediction for small blocks.

It is to be noted that such analysis is not only necessary for group 5 as done in the table, but in the AHG study some representative tools that are deemed to be critical should be selected as “study objects” which criteria need to be analysed.

Worst case of reference picture memory access is another criterion that was used in the past, but never studied for ECM tools.

Lots of these methodologies can be found in old JCT-VC and JVET documents, and in CfPs for VVC/HEVC.

Decision(SW): The tool control from AHG7 branch shall be moved to the main branch.

[JVET-AK0049](https://jvet-experts.org/doc_end_user/current_document.php?id=15019) AHG7: Analysis of MTS memory footprint and complexity [J. Gan (OPPO)]

Was reviewed in conf call JVET-AK0048.

[JVET-AK0067](https://jvet-experts.org/doc_end_user/current_document.php?id=15038) AHG7: Tool-off test results of LFNST extension and NSPT [M. Koo, J. Zhao, J. Lim, S. Kim (LGE)]

In the 36th JVET meeting, tool-off tests for a group of ECM tools with large memory access were requested, including LFNST extension and NSPT. In the last AHG7 BoG meeting, it was pointed out that SW support for disabling each of LFNST extension and NSPT with configuration options is required, which will allow measurement of coding benefit of each relative to VVC. To facilitate this, the merge request for the SW support (i.e. !807: Support for tool-off test control of LFNST extension and NSPT) has been issued the ECM15\_AHG7 branch. In this contribution, the BD-rate and complexity changes of the following three tool-off cases over ECM-15.0 are reported below {for Y, U, V, EncT, DecT}:

1. LFNST extension off: VVC LFNST is applied instead of LFNST extension.
   * AI { 0.31%, 0.18%, 0.23%, 100%, 101% }
   * RA { 0.32%, 0.13%, 0.34%, 100%, 101% }
   * LDB { 0.24%, -0.42%, -0.12%, 100%, 100% }
   * LDP { 0.21%, 0.64%, 0.63%, 100%, 100% }
2. NSPT off: LFNST extension are applied to all block sizes.
   * AI { 0.73%, 0.68%, 0.73%, 95%, 100% }
   * RA { 0.49%, 0.38%, 0.45%, 99%, 101% }
   * LDB { 0.39%, -0.55%, -0.02%, 99%, 100% }
   * LDP { 0.42%, 0.29%, 0.40%, 99%, 99% }
3. LFNST extension off + NSPT off: VVC LFNST are applied to all block sizes.
   * AI { 1.71%, 1.51%, 1.68%, 89%, 101% }
   * RA { 1.34%, 0.94%, 1.05%, 96%, 101% }
   * LDB { 0.99%, -0.14%, -0.02%, 98%, 100% }
   * LDP { 1.06%, 0.39%, 0.04%, 98%, 99% }

Results indicate that the on/off implementation is working properly for the purposes of AHG7.

Decision(SW): Adopt JVET-AK0067

[JVET-AK0092](https://jvet-experts.org/doc_end_user/current_document.php?id=15063) AHG7: Inter LDB tool assessments for each ECM version [R. Ishimoto, Z. Fan, T. Chujoh, T. Ikai (Sharp)]

ECM keeps getting updates and the latest one is version 15. This report analyzes the trends in coding efficiency gain, decoding time, and encoding time of major inter tools for each version of ECM, highlighting which areas have improved in the evolution of ECM. This report is focused more on individual tools rather than group experiments. This experiment was conducted by turning off one target tool and seeing the difference to CTC anchor, a.k.a. off tests. This result is the average of the LDB for classes B, C and E on the CTC.

v2 added the trend of the RA's luma gain across each ECM version.

v3 added the trend of encoding time vs luma gain, as well as decoding time vs luma gain from ECM1 to ECM15 with VTM11-ecm150 as the anchor and scatter plots of decoding time and luma gain, as well as encoding time and luma gain for ECM 1.0.

Interesting analysis. It was commented that the runtime analysis made on slide 2 for the entire ECM (and in the end of the word document for ECM1 vs. ECM15) would be interesting. Proponents are asked to provide Excel sheets with tool results for all ECM versions (if available).

[JVET-AK0138](https://jvet-experts.org/doc_end_user/current_document.php?id=15109) AHG7: Intra tool assessments between ECM14 and ECM15 [C.-Y. Teng, K.-W. Liang, Y.-H. Lin, Y.-C. Yang, T. Ikai (Sharp)]

This contribution provides a quantitative analysis of the luma intra prediction modes in ECM-14.0 and ECM-15.0. The analysis encompasses two perspectives: intra tool selection rate and BD-rate. The selection rate statistics were derived on a block-based basis, and the BD-rate results were obtained through tool-off testing. This contribution aims to provide insights into the impact of the newly adopted neural network-based intra prediction mode on other intra prediction tools.

It was commented that due to introduction of NN Intra in ECM15, it is not surprising that other intra tools suffer. MIP relatively loses most gain.

It was suggested that a similar analysis for intra tools targeting chroma gain would be interesting.

[JVET-AK0228](https://jvet-experts.org/doc_end_user/current_document.php?id=15217) AHG7: Tool-off test results of ALF fixed filters [N. Hu, V. Seregin, M. Karczewicz (Qualcomm)] [late]

In ECM-15.0, for luma component, two Laplacian and variance classifier-based fixed filters in adaptive loop filter (ALF) are applied to input to deblocking filters (DBF) and input to ALF. In addition, one fixed filter is also applied to residual data in random access and low delay configurations. The outputs of fixed filters are then used to derive luma ALF output. For chroma component, one Laplacian and variance classifier-based fixed filter is applied to input to DBF and input to ALF. The outputs of this fixed filter are then used to derive chroma ALF output. In this contribution, on top of ECM-15.0 with common test condition, simulation results of disabling fixed filters are reported as follows.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | BD-rate Y | BD-rate U | BD-rate V | EncT | DecT | VmPeakE | VmPeakD |
| All intra | 1.46% | 1.09% | 1.30% | 96.8% | 88.5% | 90.4% | 100.2% |
| Random access | 2.08% | 1.28% | 1.77% | 97.3% | 91.4% | 92.3% | 100.2% |
| Low delay B | 2.65% | 1.68% | 1.55% | 94.6% | 89.1% | 87.9% | 100.0% |

Results indicate that the on/off implementation is working properly for the purposes of AHG7.

Decision(SW): Adopt JVET-AK0228

[JVET-AK0319](https://jvet-experts.org/doc_end_user/current_document.php?id=15308) Proposing an Ad Hoc Group on Enhanced Complexity Analysis [[A. Stein](mailto:alan.stein@v-nova.com), S. Ferrara (V-Nova)] [late]

JVET has long recognized the importance of evaluating both coding gain and complexity during the development and tool selection process for video codecs. While significant progress has been made in standardizing and improving coding gain evaluation methods, such as BD-Rate and HDR metrics, complexity assessment remains a relatively underdeveloped area. Currently, complexity is often measured and compared using single-threaded command-line program execution time of reference software, which may not accurately reflect real-world usage scenarios.

This contribution reviews some prior work and proposes a survey and study activity to explore and develop more comprehensive and effective complexity analysis metrics and methods for existing and future MPEG and JVET codecs. Examples are reported of more detailed complexity analysis published in papers, e.g. count of numbers of operations necessary for various building blocks of AVC.

Various criteria are proposed that could be used. It was commented that very similar criteria had been used in the past during HEVC and VVC development, and also that VTM has macros that support such analysis. Without doubt, such detailed analysis has not been done in ECM and NNVC, but for a new standard it is necessary, also for a CfP

It had already been discussed earlier during the meeting that defining and investigating methods useful for more detailed complexity analysis is needed, and that AHG7 should investigate methods used in the past. Input on new criteria which could be used for current hardware/software architectures should also be solicited.

It was further commented that one problem always has been that to some extent complexity is highly dependent on implementation architecture, e.g. in terms of cache size, multi-threading implementation, etc.

## AHG8: Optimization of encoders and receiving systems for machine analysis of coded video content (3+1)

Contributions in this area were discussed at 1805–1930 on Saturday 18 Jan. 2025 (chaired by JRO).

[JVET-AK0094](https://jvet-experts.org/doc_end_user/current_document.php?id=15065) AHG8: On combination of adaptive temporal resampling, pre-processing, post-processing and ROI-based adaptive QP algorithm for machine vision [S. Wang, J. Chen, Y. Ye, B. Li (Alibaba), S. Wang (CityUHK)]

This contribution reports the performance of jointly enabling adaptive temporal resampling, pre-processing, post-processing and ROI-based adaptive QP algorithm for machine vision. According to experimental results, it is reported that when these four technologies are all enabled, 33.77% (RA), 44.33% (LD) and 80.91% (AI) averaged BD-rate savings are achieved for object detection on SFU-HW dataset, and 62.89% (RA), 59.96% (LD) and 92.33% (AI) averaged BD-rate savings are achieved for object tracking on TVD dataset. It is reported that such gains are higher than enabling pre- and post-processing and ROI-based adaptive QP algorithm together or adaptive temporal resampling algorithm only for machine vision.

Reference to JVET-AD0047 should be changed to JVET-AC0086.

The BD rate gain is consistently and significantly better than for the combination of three tools (from JVET-AJ0181) for both tasks of object detection and tracking.

Decision(SW): Include in the software git

It is agreed to extend the table in the TR by the combination of all 4 tools.

[JVET-AK0302](https://jvet-experts.org/doc_end_user/current_document.php?id=15291) Crosscheck of JVET-AK0094 (AHG8: On combination of adaptive temporal resampling, pre-processing, post-processing and ROI-based adaptive QP algorithm for machine vision) [H. Zhang (Nokia)] [late]

[JVET-AK0122](https://jvet-experts.org/doc_end_user/current_document.php?id=15093) AHG8: Dense QP results for joint software [C. Hollmann, J. Ström (Ericsson)]

In this contribution results for dense QP simulations for the joint software in the study of optimizations for encoders and receiving systems for machine analysis of coded video content are reported. It is proposed to use the results from set 2 for the comparison with other efforts in this area.

In version 2 of this contribution a minor bug that affected the results was corrected; new Excel documents were attached and the results in section 2 updated.

Set 1 is JVET-AJ0181

Set 2 is JVET-AJ0181 plus JVET-AD0122 (temporal layer dependent QP adaptation which gives additional gain for RA and LB, according to CTC JVET-AI02031.

Dense QP results are provided for both sets to be used by WG 4 with their current CTC as negotiated in the last meeting.

Decision(SW): Include software for Set 2 in the git

It is agreed to extend the table in the TR by the combination of all 4 tools from Set 2.

It was suggested to generate until the next meeting a “Set 3” which would be Set 1 plus JVET-AJ0254, i.e. generating dense QP for JVET-AK0094, and another “Set 4”, which would be Set 1 plus JVET-AJ0254, expecting even better gain for RA and LB. If the latter would show benefit, the table in the TR could be further extended by a combination of all 5 tools.

An appropriate place for storing bitstreams would be on the content server of the mpeg.expert site (clarify with Marius Preda)

[JVET-AK0272](https://jvet-experts.org/doc_end_user/current_document.php?id=15261) Cross-check of JVET-AK0122 on SFU-HW dataset (AHG8: Dense QP results for joint software) [S. Wang (Alibaba)] [late]

[JVET-AK0289](https://jvet-experts.org/doc_end_user/current_document.php?id=15278) Cross-check of JVET-AK0122 on TVD dataset (AHG8: Dense QP results for joint software) [M. Xu (Tencent)] [late]

[JVET-AK0140](https://jvet-experts.org/doc_end_user/current_document.php?id=15111) AHG9/AHG8: Showcase for Packed regions information SEI [J. Boyce, H. Zhang, M. M. Hannuksela (Nokia)]

This contribution provides a showcase for the packed regions information (PRI) SEI message. The showcase demonstrates bitrate reduction ranging from 8.68% to 89.09% and pixel rate from 38% to 90% of the source pixel rate for the VCM test conditions using the PRI SEI message.

It is also proposed to include the Packed regions information SEI message in VSEI v4.

The results and demo shown are showing benefit, though probably the rate gains likely would have overlap with foreground/background preprocessing.

[JVET-AK0141](https://jvet-experts.org/doc_end_user/current_document.php?id=15112) AHG8: Add Packed regions info SEI to TR on coding video for machine consumption [J. Boyce, H. Zhang, M. M. Hannuksela (Nokia)]

It is proposed to include the Packed regions information SEI message in the technical report “Optimization of encoders and receiving systems for machine analysis of coded video content.” The Packed regions information SEI message is included in the Technologies under Consideration (TuC) for future versions of VSEI and is proposed in JVET-AK0140 to be included in VSEI v4.

It is asserted that for machine analysis tasks where only parts of a scene are interesting this can be useful, also including options of subsampling regions individually, and identifying where the regions came from.

As the SEI message was later asserted to be mature enough in terms of design/syntax/semantics to be included in JVET-AK2006 (see HLS discussion conducted Mon. 20 Jan.), it should also be included in the TR.

Decision: Add a text describing potential usage of the PRI SEI in the TR.

Text was presented (as included in the contribution) Tuesday 21 Jan. 1240. The text is asserted to be appropriate.

It was asked if there would be relevant applications in the context of the TR where this could be combined with the annotated regions SEI (which is in principle possible). This is left for further study.

## AHG10: Encoding algorithm optimization (2)

Contributions in this area were discussed at 0910–0950 on Saturday 18 Jan. 2025 (chaired by JRO).

[JVET-AK0052](https://jvet-experts.org/doc_end_user/current_document.php?id=15023) AHG10: Teleconference on encoder optimization for multi-layer coding [P. de Lagrange]

No need for presentation

[JVET-AK0181](https://jvet-experts.org/doc_end_user/current_document.php?id=15152) AHG10: Suggested upgrade of CTC for random access for VTM [K. Andersson, P. Wennersten, J. Ström, W. Ahmad, V. Shchukin, D. Liu, J. Enhorn (Ericsson)]

This contribution suggests upgrading the VTM settings for random access to better handle a larger span of bitrates. It is well known that it can be beneficial to encode at a lower resolution for lower bitrates. It has also been shown that the DMVR encoder control can improve subjective quality in some situations also especially at lower bitrates. It is therefore suggested to enable GOP based RPR (encode in lower resolution when suitable) and the DMVR encoder control in the default settings.

Impact on BD rate for CTC:

RA Y/U/V: -0,29%/0,41%/0,14%

Impact on BD rate for QP 32, 37, 42 and 47:

RA Y/U/V: -1,61%/0,71%/-0,83%

It is also asserted that the suggested settings can improve visual quality especially when encoding at lower bitrates.

In version 2 results were added for disabling DMVR encoder control which is default in VTM when GOP based RPR is enabled. MSSSIM results are also included.

Impact on BD rate for CTC:

RA Y/U/V: -0,32%/0,37%/0,09%

Impact on BD rate for QP 32, 37, 42 and 47:

RA Y/U/V: -1,96%/0,28%/-1,24%

Under CTC, practically all benefit comes from A classes.

In NNVC, a similar approach is used in experiments for superresolution.

In the ECM software, RPR is included as option, but it was reported that currently it is broken, likely due to some recent tools integration.

It would also be relevant for investigating/resolving the DMVR bug.

Some concern was raised about changing CTC by enabling the GOP based RPR optimization for all cases, but it was agreed that it would be good to raise its visibility, as it may be relevant for certain experiments, and also for subjective investigations.

It was suggested to add a section for “alternative configurations” in the CTC file.

Decision(CTC): Include the existing configuration for GOP based RPR in JVET-AK2010, but also including a minor change for the GOP based DMVR encoder control. As another alternative configuration, the “high performance” configuration should be added. AHG3 should also report results on these alternative configurations.

## AHG13: Film grain synthesis (3+2)

Contributions in this area were discussed at 0950–1025 on Saturday 18 Jan. 2025 (chaired by JRO).

[JVET-AK0058](https://jvet-experts.org/doc_end_user/current_document.php?id=15029) [AHG9/AHG13] Comments on Film grain regions characteristics SEI message [S. Xie, P. Wu, Y. Gao, Y. Bai, C. Huang (ZTE)]

See section 6.2.4

[JVET-AK0169](https://jvet-experts.org/doc_end_user/current_document.php?id=15140) AHG13: Film grain synthesis improvement using frequency shaping [P. de Lagrange (InterDigital)] [late]

This contribution reports about the implementation of a soft frequency window instead of a rectangular frequency cutoff, when using film grain synthesis based on the FGC SEI message in frequency-filtering mode. This is reported to visually improve the synthetic grain.

A graph of a function

Description automatically generated A graph of a function

Description automatically generated

**example of soft frequency window (natural and squared)**

![A black and white square

Description automatically generated](data:image/png;base64,iVBORw0KGgoAAAANSUhEUgAAAEAAAABAEAAAAADfkvJBAAAAPElEQVRo3u3OMQ0AMAwEsTQYAifwC6cgCuJXez/pzn0V2cn6zvKcAQMGDBgwYMCAAQMGDBgwYAAAAADgA4k6AkDxgkEAAAAAAElFTkSuQmCC) A black and white image

Description automatically generated

![A black and grey square

Description automatically generated](data:image/png;base64,iVBORw0KGgoAAAANSUhEUgAAAEAAAABAEAAAAADfkvJBAAAAOUlEQVRo3u3OsQ0AMAzDsDif9rKe2h7hldwFKPdN5aTrt8t7BgwYMGDAgAEDBgwYMGDAgAEAAACAD4S6AkDI30cPAAAAAElFTkSuQmCC) A light shining on a black background

Description automatically generated

**example of 2D soft frequency window (compared to rectangular window):  
natural (top) and squared (bottom)**

Only the generation of grain patterns is changed, such that typical hardware implementations could still be used.

The rectangular window approach is described as an example in the SEI message, but not mandatory to use it.

Agreed to include in JVET-AK2020.

It is noted that software used in recent investigations on film grain are in a git repository of AHG13. It should be investigated to potentially attach relevant parts of it to the TR version 2, and/or add them to VTM (new/updated mandate to AHG).

[JVET-AK0197](https://jvet-experts.org/doc_end_user/current_document.php?id=15184) AHG13: on picture size for the interpretation of the FGC SEI message [P. de Lagrange, E. François (InterDigital)]

This contribution proposes to amend the VVC specification to refer to SPS instead of PPS to define picture resolution for the interpretation of the film grain characteristics SEI message, the intent being to describe film grain characteristics at source picture resolution independently of encoder decisions to perform reduced-resolution coding.

It is also proposed to indicate that unspecified cropping and padding can be applied as needed.

Upsampling the video with film grain synthesized at low resolution seems not appropriate. Therefore this modification clarifying how the SEI message is to be used together with RPR appears reasonable.

This was further discussed in plenary on Sunday. It is agreed that the problem exists, as the picture width/height can vary in a sequence at picture level due to encoder optimization, where just an off-the-shelf encoder might be operated which is not specifically knowing about film grain. It was however commented that the problem requires further study and the proposed solution might potentially violate some of the principles applied so far in interfacing VVC and VSEI. It was furthermore commented that the impact on existing hardware implementations should be studied. A simple solution would be adding a note in VSEI that it is not recommended to use FGC when RPR is used in VVC.

In a follow-up discussion on Sunday morning, it was commented by GJS that when originally designing the text in VVC on film grain, it was not expected that it was used with RPR. The concept of assuming that the max resolution signalled in SPS would be the one to be used, is not following the usual assumption that the SPS would express the original picture size, and that the max parameters are really indicating what happens in the bitstream. These are rather foreseen to indicate the needs of max buffer size. Furthermore, the proposal might break existing implementations of the SEI message.

A potential solution might be to express the case “when the pic width/height changes within a CVS”, where it would be up to the implementer of the SEI message which of the values to be interpreted as PicWidth/PicHeight in the SEI message.

It was also mentioned that at least two sequence-level flags exist (e.g., sps\_ref\_pic\_resampling\_enabled\_flag) that signal potential usage of RPR, which could be used for this purpose.

Further offline work was requested to be conducted by GJS, PdL, YKW, WH, JS.

It was reported that an idea of resolving the issue in a way that is in better harmony with the intended usage of SPS was discussed, but then it was found that additional issues might exist when FGC is operating in SPO and resolution is changed. Potentially a new SEI message about intended display size could resolve this issue at least for systems that are able to interpret it.

It was commented that the text considered would improve the situation in the non-SPO case, such that the interpretation of the pic width/height could be handled by a formulation “UpsamplingInferredFlag may be set as follows”, and also restricting this to the case where FGC is not part of SPO, or is the first step would at least improve the situation. It was requested to produce a new input document.

[JVET-AK0339](https://jvet-experts.org/doc_end_user/current_document.php?id=15328) AHG9: Interface text and semantics of the FGC SEI message (Y.-K.Wang (Bytedance), ??] [late]

TBP

[JVET-AK0211](https://jvet-experts.org/doc_end_user/current_document.php?id=15200) AHG9/AHG13: Proposed Film Grain Region SEI message for version 5 of VSEI [S. Wenger, G. Teniou, A. T. Hinds (Tencent)]

See section 6.2.4

## Implementation studies (2)

Contributions in this area were discussed at 1025–1100 on Saturday 18 Jan. 2025 (chaired by JRO).

[JVET-AK0258](https://jvet-experts.org/doc_end_user/current_document.php?id=15247) AHG9: Demo of real-time NNPF inference for VVC decoding on consumer laptop [J. Funnell, M. Santamaria, F. Cricri, M. M. Hannuksela, M. Pedzisz, R. Yang, S. Schwarz (Nokia)] [late]

In this demo, Nokia showcases the integration of a neural network post-processing filter (NNPF) for visual quality enhancement with Versatile Video Coding (VVC) through the Versatile Supplemental Enhancement Information (VSEI) standard. The demo verifies the possibility to run NNPF inference (complexity block-wise: 34 kMAC/pixel, number of parameters 109068) for VVC-coded 1080P video in real-time on current consumer hardware. Nokia plans to release the related software in the near future.

Demo was shown on projector. The residual generated by the filter was also demonstrated, which showed an effect that edge sharpening happens.

It was asked if it would also be beneficial to use the same filter in the loop? Not investigated.

It was commented that it is interesting to see that an NNPF with double complexity of LOP can be run in real time on NPU for 1080P25 (1080P30 might be current limit according to presenter). This would mean that 1080P60 might roughly be possible with LOP. VVC uses the hardware of the Intel processor in the PC.

[JVET-AK0274](https://jvet-experts.org/doc_end_user/current_document.php?id=15263) Development of VVC live encoder with content layering capability [S. Nemoto, S. Iwamura, A. Ichigaya (NHK)] [late]

This contribution reports development of VVC multilayer live encoder with content layering capability. It is known that the complexity of multilayer encoding is large and real-time processing is challenging especially for high resolution. In the developed system, encoding process of each layer was separated into two encoders to achieve real-time multilayer encoding. Furthermore, considering the developed encoder is designed specifically for content layering, RD optimization can be significantly simplified for non-overlaid region in the enhancement layer (EL) encoding since input image of the EL in those regions are identical to the corresponding regions in the base layer. In the experiment, it was confirmed that 4K/4K real-time multilayer encoding can be performed using workstation with the CPU (2x AMD EPYC Genoa 9654 96C/192T 2.4GHz) with a frame rate of 60Hz.

It was asked what the compression performance compared to VTM is? Was not investigated.

It was asked if it has compression benefit compared to simulcast, performing the overlay by mechanism at system layer? Was not investigated.

## Profile/tier/level specification (0)

This section is kept as a template for future use.

## Gaming content compression (AHG15) (4)

Contributions in this area were discussed at 1120–1225 on Saturday 18. Jan. 2025 (chaired by JRO).

[JVET-AK0081](https://jvet-experts.org/doc_end_user/current_document.php?id=15052) AHG6: wPSNR log output for gaming content in ECM [X. Liang, K. Choi (Kyung Hee Univ.)]

This contribution addresses the issue of ECM not producing a wPSNR summary during the G3 sequence encoding process. It provides the modifications implemented to enable ECM to produce a wPSNR summary for the encoding results of the HDR sequence in Class G3.

It was recommended to submit software fix and a merge request.

Decision(SF/BF): Agreed that this is necessary, not only for G3, but also for the investigation of HDR materials in AHG17.

[JVET-AK0082](https://jvet-experts.org/doc_end_user/current_document.php?id=15053) AHG15: Analysis of ECM 12-bit internal bit depth in gaming content [X. Liang, K. Choi (Kyung Hee Univ.), S. Lee (Atins)]

This contribution examines the ECM 12-bit internal bit depth for game content. In ECM-15.0, a 12-bit internal bit depth has been implemented for all tools within the ECM software. In this contribution, a 12-bit internal bit depth was applied to analyze the behavior of game content. The impact on BD-rate performance when applying a 12-bit internal bit depth under the ECM CTC for game applications is reported as follows:  
  
AI: { -0.17% Y, -0.18% U, 0.06% V, 99.1% EncT, 104.0% DecT } for SDR Class (G1, G2)  
AI: { -0.51% Y, -4.72% U, -5.65% V, 93.5% EncT, 94.3% DecT } for HDR Class (G3)

(run times not reliable, 12 bit is probably slower)

It is noted that for other classes the gain of ECM was around 0.25-0.3%, and around 0.5% for RA.

For VTM, very similar gain.

Interesting information to know numbers at least for AI in the cases of classes G. RA not yet available due to long run time.

[JVET-AK0083](https://jvet-experts.org/doc_end_user/current_document.php?id=15054) AHG15: Analysis of high-level control of intra prediction methods in gaming content [X. Liang, K. Choi (Kyung Hee Univ.), S. Lee (Atins)]

This contribution investigates the high-level control of intra prediction methods for game content. In ECM-15.0, the approach of high-level control of intra prediction methods has been adopted in the ECM software, demonstrating efficient improvements in coding performance for screen content. In this contribution, the method was tested to analyze its behavior with game content. The impact on BD-rate performance when applying the high-level control of intra prediction methods under the ECM CTC for game applications is reported as follows:

Test1 (whole tools off):  
AI: { 0.11% Y, 0.04% U, 0.04% V, 98.2% EncT, 89.0% DecT } for SDR Class (G1, G2)  
AI: { 0.20% Y, 0.12% U, 0.30% V, 99.0% EncT, 102.7% DecT } for HDR Class (G3)  
  
Test2 (adaptive tools off):   
AI: { 0.05% Y, 0.03% U, 0.03% V, 99.5% EncT, 100.6% DecT } for SDR Class (G1, G2)  
AI: { 0.00% Y, 0.00% U, 0.00% V, 92.8% EncT, 95.6% DecT } for HDR Class (G3)

The complete or adaptive disabling of three tools (reference sample filtering, PDPC, intra prediction fusion) was originally introduced to avoid degrading of graphics-generated content. In CTC of other classes, tools are never turned off for camera content, and adaptively turned off for classes F and TGM.

For gaming content CTC also uses the three tools always enabled. The results indicate that PSNR-wise, this may be correct, however examples are shown where visual quality is better with turning tools.

Further study would be needed to investigate this. As gaming content is sometimes similar to regular video, if at all, an adaptive disabling would be appropriate, or potentially more specific encoder optimization would be necessary.

[JVET-AK0170](https://jvet-experts.org/doc_end_user/current_document.php?id=15141) AHG15: On the compression of depth maps from auxiliary data [J. Sauer, T. Solovyev, E. Alshina (Huawei)]

Sequences in classes G1/G3 of the “Common Test Conditions (CTC) for gaming applications” (JVET-AJ2027) are provided along with auxiliary data, which consists of auxiliary meta data (e.g. camera parameters) and auxiliary video data (e.g. a depth map or a motion map). This contribution is a prelimiary study of compressing the auxiliary depth maps.

Depth maps are converted into an integer representation, which is then compressed as aux picture. PSNR is calculated using the depth maps for global motion compensation between subsequent frames. This cannot capture object motion, or newly appearing content, which is indicated by the fact that this PSNR sometimes saturates at relatively low values.

However, as saturation occurs at around 14-16 bit of depth map integer precision (without compression), this may indicate that FP precision might not be necessary.

Before further investigating compression, it should be clarified by which precision depth information is needed at the receiver end by gaming applications.

As a similar example, in applications of MV-HEVC and 3D-HEVC, depth maps were expected to be needed for view synthesis, for which a certain precision of the depth map was required to achieve a certain quality, and then both video and depth were compressed together.

Further study in AHG on these aspects.

## Generative face video (AHG16) (2)

Contributions in this area were discussed at 1225–1320 on Saturday 18 Jan. 2025 (chaired by JRO).

Also refer to section 6.1.4

[JVET-AK0068](https://jvet-experts.org/doc_end_user/current_document.php?id=15039) AHG16: GFVC Extension of the VVC Standard [L. Liu, C. Jung (Xidian Univ.)]

This contribution proposes GFVC extension of the VVC standard, named VVC GFVC. VVC GFVC employs a generic video coding standard VVC for face video compression, which does not need any additional SEI message in transmission. To save bit and maintain visual quality, VVC GFVC adopts intraframe guided face generation that uses the intraframe compressed at low QP as guidance and the inter frames compressed at high QP as target. VVC GFVC utilizes multi-step asymmetric Cycle GAN to ensure training stability and efficiency in face generation. Experimental results show that VVC GFVC achieves average BD-rate savings of 37.42%, 40.79%, -3.19% and -3.55% in DISTS, LPIPS, PSNR, and SSIM over the VVC anchor (VTM 22.2) on Class A and Class B test sequences of GFVC CTC, respectively.

Intra picture is coded by QP22, inter frame by QP52. The generative model is only post processing, no parameters are encoded, no normative change to VVC necessary. Potentially the EOI SEI could be used to signal that something like that has been done, not recommending the direct decoder output for viewing. It was asked how the performance would be if other QPs or other resolution would be used.

It would be desirable for a real-world application to allow variation of rates. It was suggested to investigate if QP could be used as input to the model.

Also support for other resolutions would be useful.

Model has “very high complexity” – more information about number of parameters, kMAC/pixel would be useful.

It was suggested as a first step studying the benefit when modifying inter QP to other values, and also looking into visual performance with some of the more difficult sequences.

[JVET-AK0069](https://jvet-experts.org/doc_end_user/current_document.php?id=15040) AHG16: QP-Adaptive GFVC [W. Kang, L. Liu, C. Jung (Xidian Univ.)]

This contribution proposes a QP-adaptive GFVC framework that incorporates varying feature sizes in supplemental enhancement information (SEI) message according to the given QP value. At low QP (e.g., QP22 and QP32), the proposed framework adds optional setting to CFTE anchor by providing additional features of 8x8 size. The optional setting remarkably enhances face generation results based on multi-level Multi-Deconv Transformer Attention (MDTA) and Spatially Adaptive DEnormalization (SPADE) by capturing the difference between keyframe and inter frames. The proposed framework achieves average BD-rate reductions of 60.11%, 58.58%, 6.06%, and 28.06% in DISTS, LPIPS, PSNR, and SSIM over the VVC anchor (VTM 22.2) for test sequences of Class A and Class B. Moreover, the proposed framework provides a GFVC solution to be adaptable to the network bandwidth.

The syntax of the current SEI message could be used without changes.

It was asked how much gain comes from the other network architecture, and how much from QP adaption.

It was recommended to study encoding the data within the SEI message, and report rates comparing against other schemes when also the SEI message overhead is counted in the bitrate.

If it the proposed method does not require normative changes and provides improvement, it could be included as another model in the GFVC software.

Also information about complexity should be provided.

## Testing for future standardization (AHG17) (10)

Contributions in this area were discussed at 1600–1745 on Saturday 18 Jan. 2025, and at 1510–XXXX on Sunday 19 Jan. 2025 (joint with AG 5, chaired by JRO and MW).

Some aspects discussed under 4.7 could also be relevant here.

[JVET-AK0041](https://jvet-experts.org/doc_end_user/current_document.php?id=15011) AHG17: Materials for assisting sequence selection for non-CTC testing [J. Pardo, J. Sauer (Huawei)]

Was reviewed in AHG telco

[JVET-AK0042](https://jvet-experts.org/doc_end_user/current_document.php?id=15012) AHG17: AhG meeting notes [M. Wien]

No need for presentation, included in AHG report

[JVET-AK0043](https://jvet-experts.org/doc_end_user/current_document.php?id=15013) AHG17: Additional sequences for testing of video coding technology beyond CTC [R.-L. Liao, Y. Ye (Alibaba)]

Was reviewed in AHG telco

[JVET-AK0044](https://jvet-experts.org/doc_end_user/current_document.php?id=15014) AHG17: Cropped HDR sequences for non-CTC testing [A. Filippov, J. Konieczny, Z. Zhi, V. Rufitskiy, H. Qin, T. Dong, X. Tang (TCL), S. Shen, Y. Qin, Y. Wang, X. Wang, Y. Guan (NERC-DTV)]

Was reviewed in AHG telco

[JVET-AK0045](https://jvet-experts.org/doc_end_user/current_document.php?id=15015) AHG17: Test sequences from YouTube [X. Li, Y. Wang, N. Birkbeck, B. Adsumilli (Google)]

Was reviewed in AHG telco

[JVET-AK0066](https://jvet-experts.org/doc_end_user/current_document.php?id=15037) AHG4/AHG17: 8K HDR sequences as a response to the call for new HDR test materials for future video coding [Y. Li (CMG), Q. Zhang, J. Ning (ABP), Y. Zhao (Huawei)]

As a response to the call for new HDR test materials for future video coding [1], four 8K HDR test sequences are provided in this proposal by China Media Group (CMG) and Academy of Broadcasting Planning (ABP), National Radio and Television Administration (NRTA).

Presented by Yin Zhao (Huawei) on behalf of the owners. It was asked to put his name into the list of authors.

SG16 should be updated SG21 in the copyright statement.

Sequences were already included in viewing session.

[JVET-AK0180](https://jvet-experts.org/doc_end_user/current_document.php?id=15151) On User Generated Content and non-pristine source videos [J. Samuelsson-Allendes (Sharp)]

This document discusses the topics of User Generated Content (UGC) and Non-Pristine Source (NPS) videos in the context of selecting test material for next-generation video coding standard development. Three different scenarios are described for NPS videos, and it is analysed how conventional methods for subjective and objective quality evaluation applies to these scenarios. Based on the analysis, a few proposals are made for next steps regarding UGC and NPS videos.

The second version of the document adds an example screenshot from one of the UGC test sequences being evaluated at the 37th JVET meeting.

Non-Pristine Source (NPS) is a term that may be used to describe source videos that include visual impairments or degradations that causes them to give a reduced visual experience compared to the “optimal” visual quality. The reason for the impairments can generally be attributed to one or both of the following:

* Non-ideal capturing conditions and capturing equipment. This includes out-of-focus blur, shaking camera, camera noise, rolling shutter skew, etc.
* Pre-compression. This includes for example blockiness, blurriness, over-smoothness, loss of details, motion artefacts, etc.

For the second case, it might happen that a second encoding provides better quality, therefore measuring PSNR versus the pre-compressed “original” is not appropriate.

For scenarios 1 and 2, it would be most appropriate to retain the “documentary style” of UGC, potentially with preprocessing.

As in scenario 3 (upload social platform + transcoding) preprocessing may be used before re-encoding. Something similar might be necessary, or industry might be asked to provide uploaded and preprocessed sequences before re-encoding them, or before/after preprocessing – needs more investigation.

A first step of transcoding already happens on the smartphone, where the content stored after capturing is further compressed for the upload.

“Non pristine” conditions such as low lighting/noisy, shaky, blurred etc. would also be important for industry needs. It would be useful to define a separate class for testing that, potentially with narrower QP range. Content to be selected should be realistic as coming out from camera, and difficult to encode. Also this needs more investigation.

For the current moment, priority should be given to select sequences appropriate for visual testing, and not too simple to encode. This would mostly be “pristine” sequences.

[JVET-AK0196](https://jvet-experts.org/doc_end_user/current_document.php?id=15167) AHG17: Additional sequences for testing of video coding technology beyond CTC [J. Pardo (Huawei)]

This contribution proposes additional candidates for AHG17 testing.

Test results for JVET-AK0066, cropped versions of.

* ChineseDancing
* WaterfallForest
* Bookstore
* WomenFootball

These cropped versions had already been included in the viewing.

All sequences could be useful, where compression artifacts are best visible in WaterfallForest and WomenFootball, least in Bookstore.

The sequences should be made available in the content server

[JVET-AK0246](https://jvet-experts.org/doc_end_user/current_document.php?id=15235) Crosscheck of JVET-AK0196 (AHG17: Additional sequences for testing of video coding technology beyond CTC) [Z. Xiang (Tencent)] [late]

[JVET-AK0218](https://jvet-experts.org/doc_end_user/current_document.php?id=15207) AHG17: UGC test sequences from Huawei [Y. Zhao, Y. Sun, P. Liu, Y. Lu, C. Wang, E. Alshina (Huawei)]

Two 1080p 60fps uncompressed UGC sequences are provided in this contribution. They are recorded by Huawei Mate 70 smartphone.

New sequences not included in test

Sequences are 10 seconds long. ECM/VTM encodings at CTC are also available for viewing.

RA Simulation results with ECM (ECM-15.0) and VTM (vtm-11.0ecm15.0):

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Reference | | | | Test | | | | BD-rate | | |
|  |  | kbps | Y psnr | U psnr | V psnr | kbps | Y psnr | U psnr | V psnr | Y | U | V |
| Blackswan | 22 | 61233.27 | 38.34 | 43.64 | 45.93 | 58423.98 | 38.54 | 43.93 | 46.27 | -9.46% | -33.73% | -30.91% |
|  | 27 | 24771.86 | 34.45 | 41.53 | 44.51 | 23750.20 | 34.65 | 42.07 | 44.97 |  |  |  |
|  | 32 | 9302.71 | 31.24 | 39.97 | 43.11 | 9015.99 | 31.45 | 40.74 | 43.62 |  |  |  |
|  | 37 | 3285.28 | 28.62 | 38.65 | 41.57 | 3155.01 | 28.80 | 39.30 | 42.11 |  |  |  |
| Camellia | 22 | 22911.49 | 40.89 | 44.54 | 46.00 | 20060.45 | 41.13 | 44.58 | 46.04 | -24.76% | -20.18% | -20.14% |
|  | 27 | 8812.49 | 37.89 | 42.43 | 44.02 | 7442.16 | 38.26 | 42.57 | 44.14 |  |  |  |
|  | 32 | 4081.49 | 35.32 | 40.27 | 42.02 | 3417.12 | 35.77 | 40.48 | 42.21 |  |  |  |
|  | 37 | 2054.40 | 32.69 | 38.26 | 39.96 | 1754.25 | 33.19 | 38.42 | 40.22 |  |  |  |

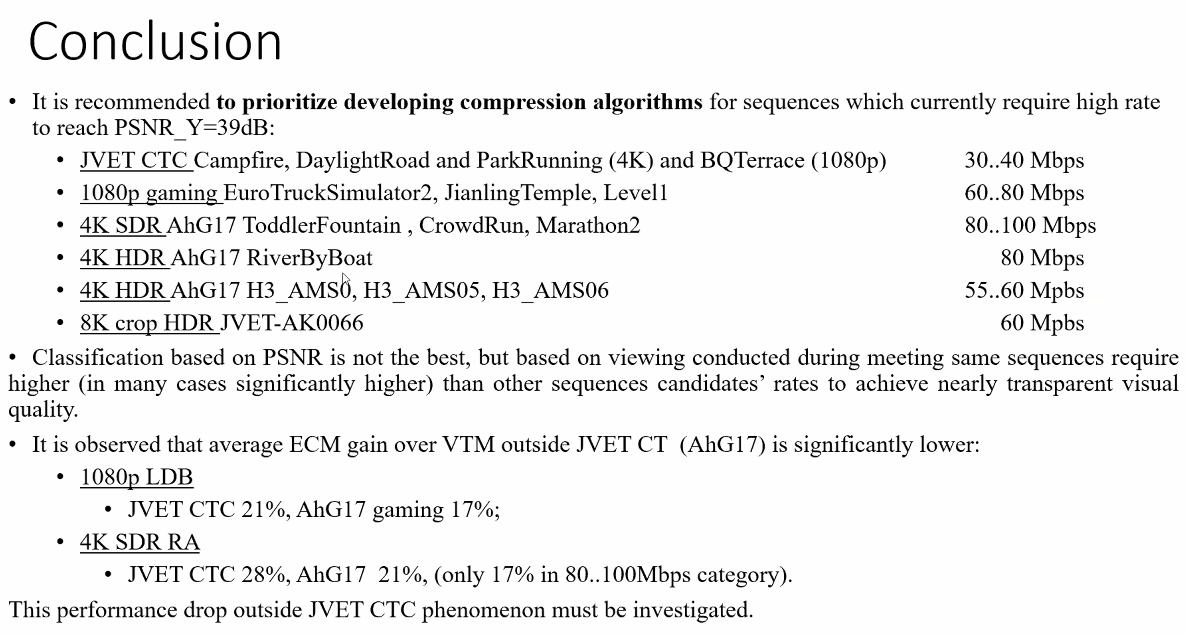
It was reported that more uncompressed sequences captured by smartphone could be made available. This would be welcome, but it was commented that availability of uncompressed or slightly compressed sequences from other manufacturers would also be useful, potentially also 10 bit HDR – sequences in JVET-AK0218 are 8 bit SDR.

[JVET-AK0221](https://jvet-experts.org/doc_end_user/current_document.php?id=15210) [AHG17] JVET CTC and non-CTC sequences performance analysis [E. Alshina, J. Pardo, J. Sauer, T. Solovyev (Huawei)]

This contribution analyzes performance of VTM and ECM for sequences from JVET common test conditions and outside of those, including sequences studied in AhG17 and some sequences from new proposals. Sequences are classified based on the difficulty of encoding. The most challenging sequences are recommended to be in the focus of JVET, studying the needs of the evidence for next generation video coding.

The point of assuming transparent quality at 39 dB PSNR may be too high for more complex sequences. Comparison with the visual test results shows that MOS 8-9 is often achieved by much lower rates.

Investigation is made also based on typical bit rates used for encoding certain types of material in certain application domains. It was commented that for 4K HDR the assumption of up to 85 Mbit/s goes too high – typically not more than double rate would be used than for SDR.



The approach of selecting target rates for testing based on typical application scenario requirements is deemed to be useful.

It was commented that in practice rate control would be used to reach certain target rates. With constant QP, many bits might be spent in vain by coding certain areas with better quality.

It was commented that other metrics than PSNR might be used.

The test should not only focus on content that is difficult to code. However, content that is more in the “normal” range would be relevant as well. Of high importance is the suitability for subjective viewing as well.

[JVET-AK0282](https://jvet-experts.org/doc_end_user/current_document.php?id=15271) AHG4/AHG17: Additional HDR sequences for non-CTC testing [A. Filippov, J. Konieczny, Z. Zhi, V. Rufitskiy, H. Qin, T. Dong, X. Tang (TCL), S. Shen, Y. Qin, Y. Wang, X. Wang, Y. Guan (NERC-DTV)] [late]

This contribution provides a response of TCL and NERC-DTV to the call for new HDR materials for future video coding development issued at the 35th JVET meeting in Sapporo (Japan) and proposes 2 more HDR sequences. The parameters of the proposed sequences are as follows:

* resolution: 8192x4320;
* frame-rate: 60 fps;
* bit-depth: 16;
* transfer function: PQ.

New sequences not included in test

Sequences are generated in 16 bit and would be provided as 10 bit YUV (PQ). Both versions are available from an ftp server.

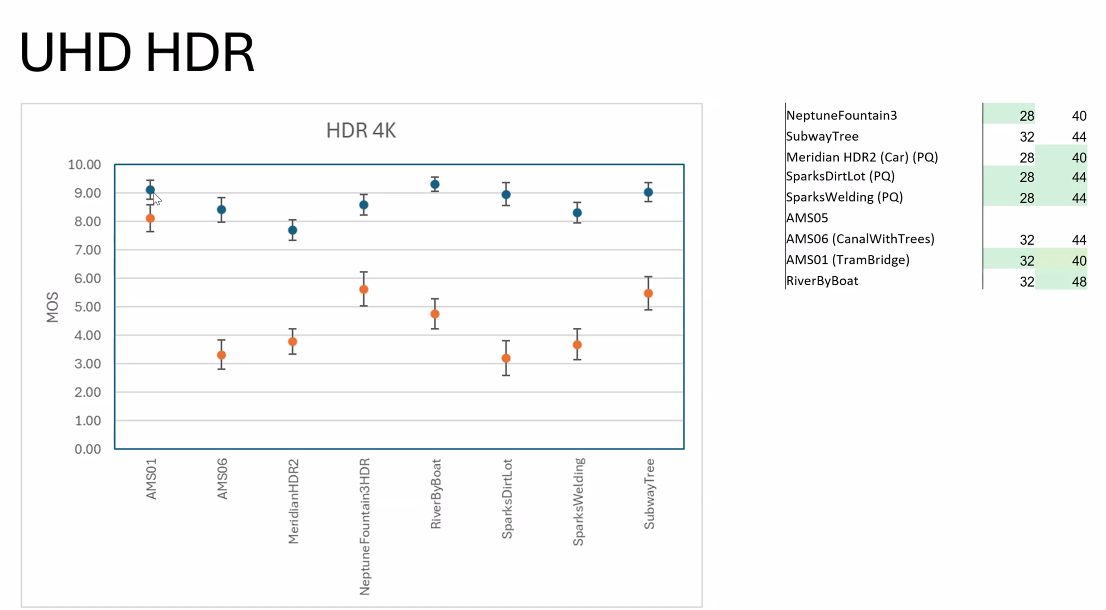
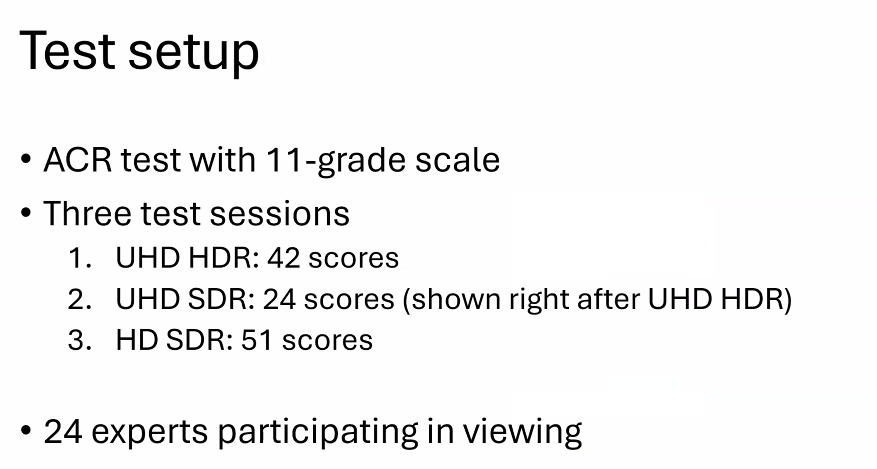
Coded results not available yet.

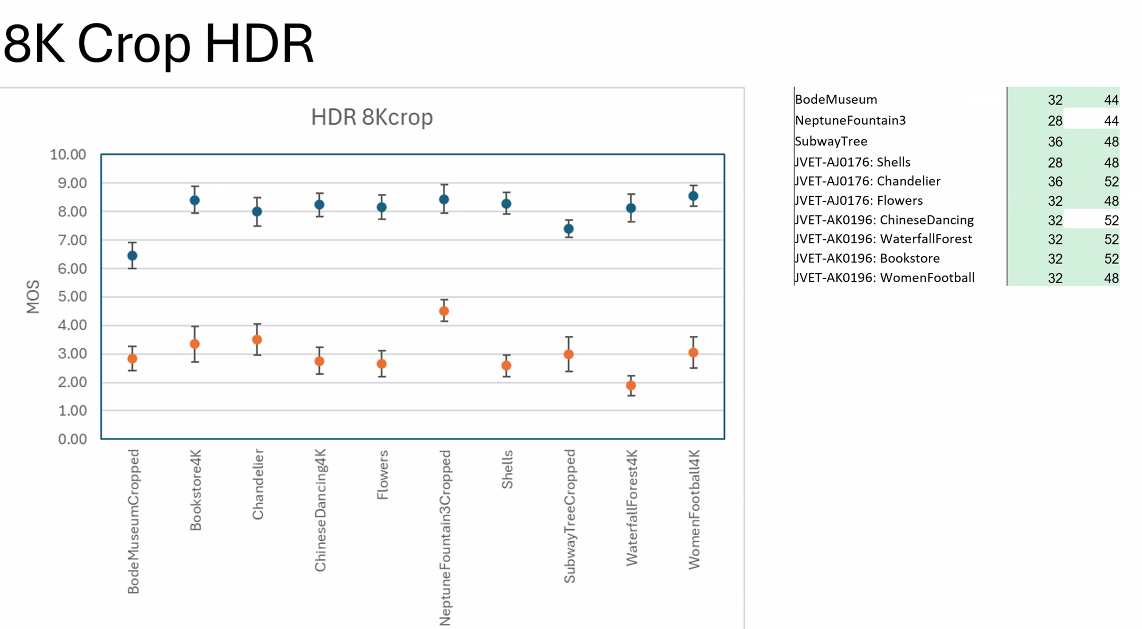
One of the sequences (melt metal) is dark and has complex motion of particles.

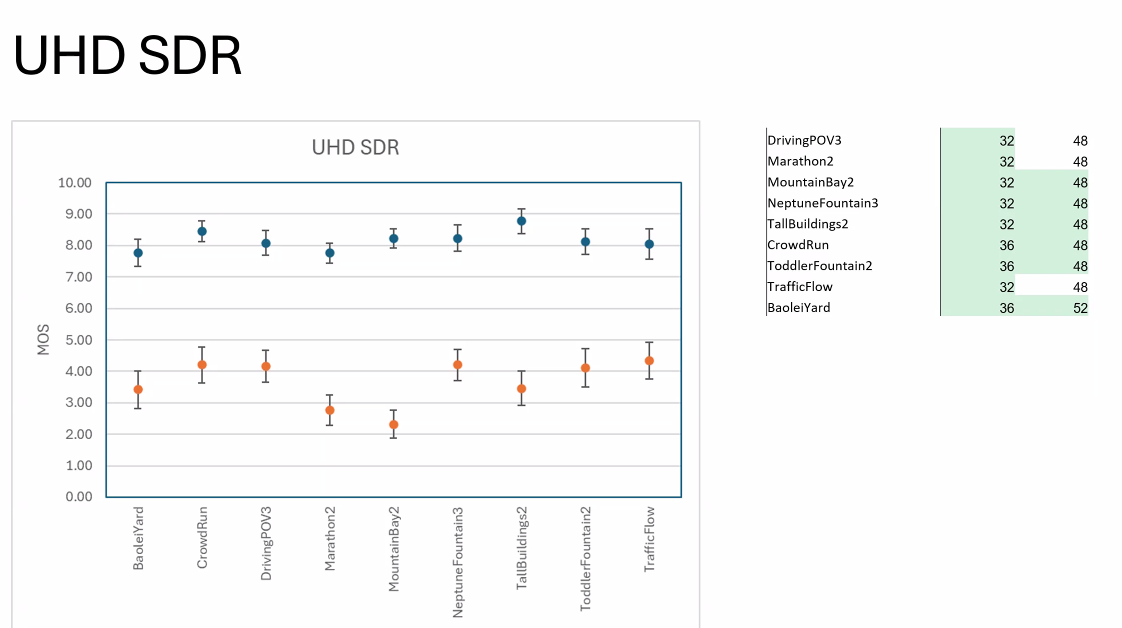
Viewing session to be planned for JVET-AK0218 and JVET-AK0282. Monday evening after the joint meeting might be a good time (to be announced).

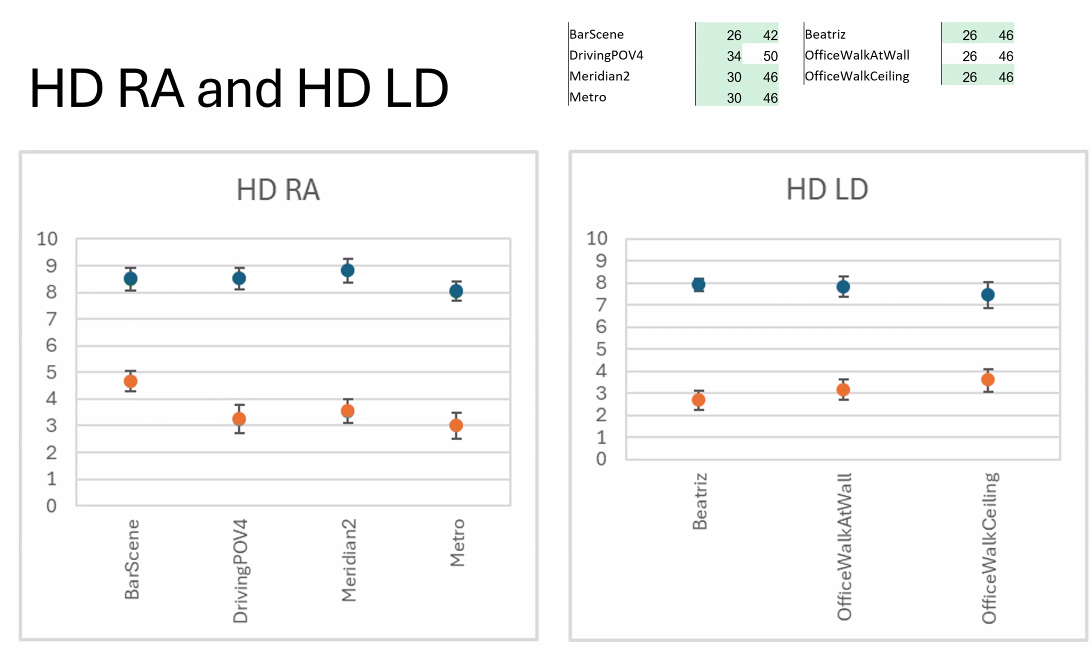
[JVET-AK0335](https://jvet-experts.org/doc_end_user/current_document.php?id=15324) AHG17: Results of expert viewing of non-CTC sequences [M. Wien (RWTH)]

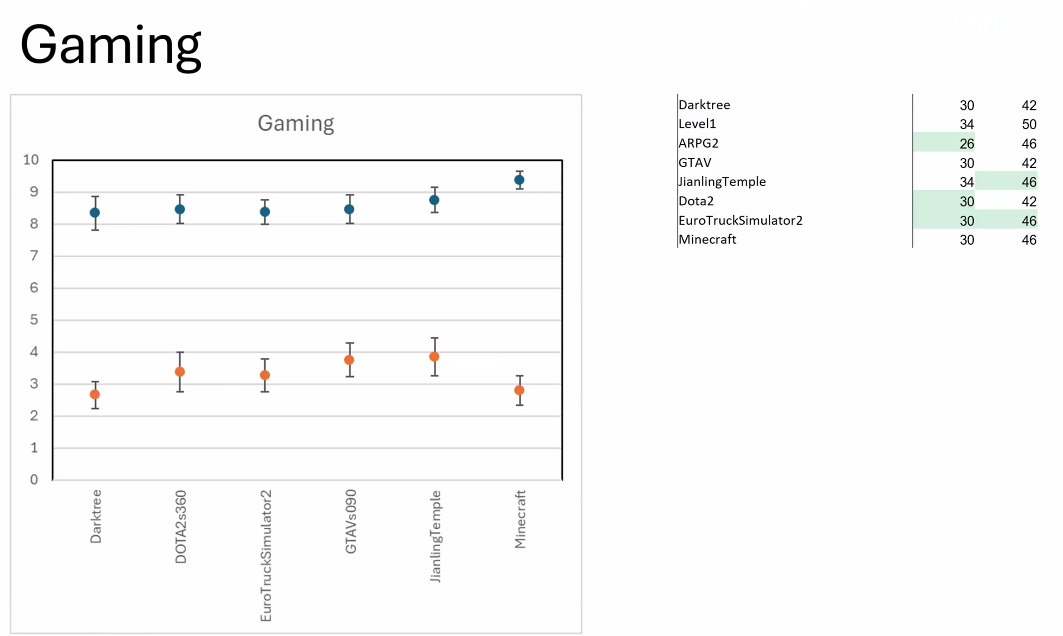
Preliminary version presented Saturday 18 Jan.



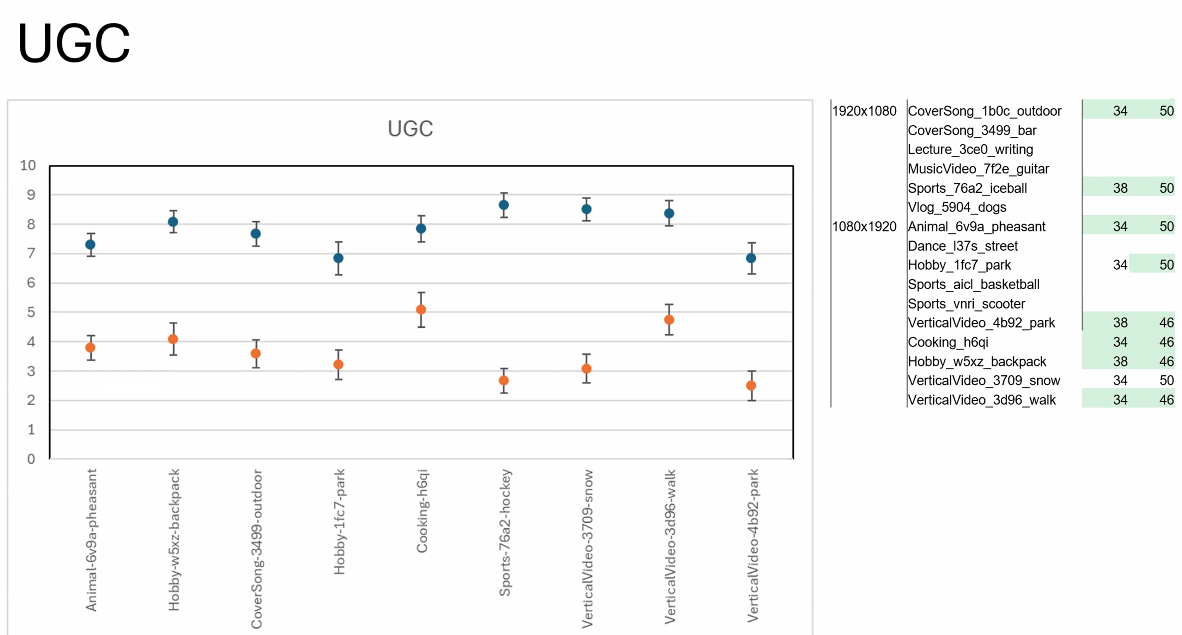








ARPG and Level 1 were accidentally missed/forgotten.



There are also comments on specific sequences which will be added. Suitability for viewing and for recognizing artifacts is important selection criterion.

It was commented that in previous calls with 4 rate points, typically factors of approx. 4 between lowest and highest rate, with sufficient difference in quality.

Definition of lowest/highest rate point should depend on whether main target is about significantly improving quality at same rate regardless of complexity, or also reaching no/small improvement at lower complexity than VVC.

Follow-up discussion Sunday afternoon:

Criteria that could be used to identify sequences that are unsuitable:

* MOS between lowest and highest rate points too close
* Rate between lowest and highest rate point not sufficiently different.
* Sequences with low amount of motion (e.g., the selected crop from Bode Museum 8K), which would appear as almost a still picture, but is slightly differently coded at each GOP.

For cases where the confidence interval of MOS at highest rate point is completely below 8, it should be checked whether the QP of highest rate point needs to be decreased.

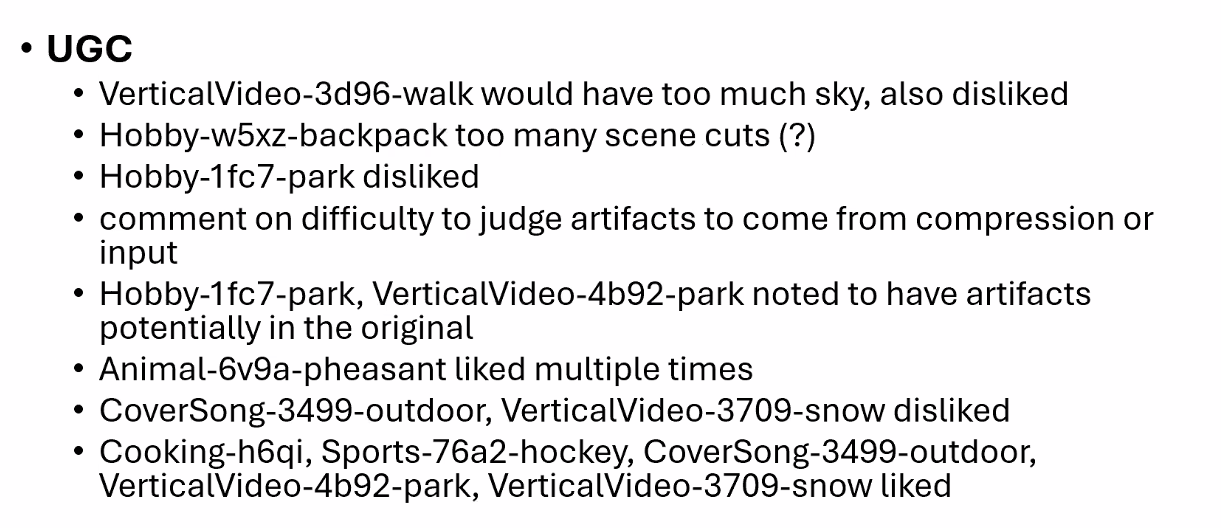
For lowest rate point, ECM around 3 is asserted to be OK, as it could be expected that VTM would still be above MOS 1.

As in the past, the merit of proposals could mainly be assessed by comparing the quality benefit, either against anchors, or one proposal against another. For this, 4 rate points are enough. If a 5th rate point would be added for the VTM anchor at the high end, giving approximately same quality as ECM at highest rate, to a certain extent the rate saving of a proposal could be assessed by comparing rate points 1-4 of a proposal against 2-5 of VTM. It should be avoided that quality of proposals saturates already at the highest two rate points.

For sequences where MOS of lowest rate point is too high, it needs to be identified if it QP can be increased, or if the problem is that compression artifacts are difficult to see, or if quality breaks down suddenly (“sharp knee of RD curve”)







Mathias will provide a new version with the comments above

Experts who were involved in the pre-selection screening are asked to analyse the results along the criteria listed above and comments, to have an initial categorization of sequences as “rather not useful”, “definitely useful”, and “potentially useful”.

Follow-up discussion Tuesday afternoon.

As a result from the joint meeting on Monday, JVET was tasked to work towards preparing a Call for evidence, with a first output document calling this a Call for Evidence to be issued as output from the next meeting.

The development of this Call is JVET business, but the Call in terms of defining which evidence is sought should as much as possible consider the requirements of the parent bodies (which are not finalized yet, and not unified yet between the parent bodies, a joint AHG was set up). The CfE can be developed in parallel, whereas by the time of CfP a stable version of requirements doc is needed.

What was discussed and edited in the joint meeting, was an input document m70750, which was an updated version of WG 2 N 411 from the Kemer meeting.

From the current meeting, a new version of the document JVET-AK2026 should be produced, updating the set of sequences and defining an initial range of rates per sequence (adding notes where further updates would be necessary, and identifying in which categories the number of sequences is too low).

Work to be done until an AHG telco (beginning of February)

* defining an initial range of rates and rate points per sequence (Mathias, Kenneth, Elena, Philippe, Yan), and provide an input on that, or include into a first version of JVET-AK2026.
* Proposals of additional sequences with at least VTM encoding, and suggested range of rates
* Encoding results for additional conversational with at least VTM encoding, and suggested range of rates

Updated version of JVET-AK2026 would be result of that telco.

Work to be done until the AHG meeting in March:

* For ECM, produce encodings for a denser distance of QPs (QP distance of 2, still with ECM 14, as some existing bitstreams can be reused)
* For VTM, produce encodings for a denser distance of QPs (using same VTM as before, QP distance of 1)
* For the new sequences from JVET-AK0218 and JVET-AK0282 (definition of cropping areas necessary), and others suggested until the AHG meeting – encoding needs to be performed from scratch with same distance of QPs (2 for ECM, 1 for VTM)
* Encoding VVenc in “medium” mode should be done additionally, with the perspective of potentially including in the CfE a category of testing algorithms with fast encoding (details of turning off rate control, perceptual optimization, multi-threading, or testing both and comparing both as if two proposals would be compared)

Encoding and decoding times should also be reported.

The lowest and highest rate points defined should not be exactly those produced by ECM, but somewhat reasonably rounded values, e.g. quantized with a precision of 200 or 500 kbit/s (likely also depends on the actual rate range of a given sequence).

Initial rate points might be defined on a logarithmic scale between lowest and highest selected rate points. At this stage, it is not necessary to perform rate matching, using the next available value would be good enough to assess if the quality of ECM and VTM is suitable at the defined rate points, and if sufficient quality difference can be observed between one rate point and the next.

Work to be done at the AHG meeting in March

* Perform testing of the sequences with all rate points such that the information needed can be extracted from the results, and refine set of sequences, rate points, etc.
* Identify an additional higher rate point for VTM where it would have approximately same visual quality as ECM at the highest rate point.
* Prepare a proposed draft text for CfE as input to the 38th JVET meeting.

A proposal for classifying suitability of sequences was presented as v3. For the HD categories (both RA and LD) still more sequences would be needed. It was suggested that additional conversational-type sequences for LD could be drawn from the set proposed in JVET-Y0071 which seem not yet available on the content server. This content should also be included in the investigation, and potentially more for RA could be found in the set of sequences provided by Netflix.

It was agreed to remove all sequences classified in category 0.

# Low-level tool technology proposals (XX)

## AHG11/AHG14: Neural network-based video coding (23)

### Summary and BoG reports

Contributions in this area were discussed at 1600–1815 on Tuesday 14 Jan. 2025 (chaired by JRO).

[JVET-AK0023](https://jvet-experts.org/doc_end_user/current_document.php?id=15189) 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 36th and 37th JVET meetings to evaluate **Neural Network-based Video Coding** (NNVC) technologies, analyze their performance, evaluate their complexity aspects, and clarify training procedure. Two teleconferences have been conducted during this meeting circle.

NNVC has three major operation points with NN-intra always enabled and NN-based in-loop filters with different level of complexity: High Operational Point (HOP, 471kMAC/pixel), Low Operational Point (LOP, 21.6kMAC/pixel) and Very Low Operational Point (VLOP, 9.9kMAC/pixel). Compared with VVC, NNVC HOP5 provides {14.2%, 19.5%, 19.9%}, NNVC LOP4 provides {7.6%, 14.3%, 13.2%} and NNVC VLOP3 provides {5.8%, 6.6%, 5.7%} bitrate saving for {Y, Cb, Cr} channels in the RA configuration. Observed run time increase relatively to VVC is ×2.6, ×1.2, ×1.2 (encoding) and ×1136, ×36, ×17 (decoding) for HOP, LOP and VLOP configurations respectively. EE1 tests were targeting improving the performance complexity trade-off of all three operation points of neural-network based filter and super-resolution.

**Introduction**

It was decided that the code base for the test should be NNVC-11.0. The anchor for all tests is default configuration of NNVC-11.0 (NN-Intra as adopted from [JVET-AI0130](https://jvet-experts.org/doc_end_user/current_document.php?id=14372)and LOP-4 filter are enabled). Anchor performance results were provided by AhG 14 and have been used by all proponents.

For tests competing with technologies in NNVC-11.0 it was agreed to configure the proposed solution targeting close to existing NNVC tool complexity, but not exceeding it**:**

* kMAC/pxl of EE1 test ≤ kMAC/pxl NNVC (*must*),
* Number of Parameters EE1 test ≤ Number of Parameters NNVC (*if possible*).

The minor violation of this requirements is allowed for the final model, if the reason is keeping the number of channels a multiple of 16 (or at least 8), since it is very helpful for neural network algorithms implementation on wide range of platforms. Reported results should follow the constraints for evaluation purpose but the final model might violate the constraints.

NN architecture provided in this test description should not be changed, beside minor adjustment for parameters (such as channels number) in order to meet recommendation above.

**Exact parameters settings were announced by proponents by 2nd AhG11/14 teleconference on December 20.**

Table 1 provides performance and complexity number of several tools’ combinations in NNVC relatively to VVC anchor.

*Table 1 Performance of NNVC tools combinations relatively to VVC*

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Test | Random Access | | | | | All Intra | | | | | Total kMAC/pxl | Total Param (Mprm) | |
| Y | U | V | Enc | Dec | Y | U | V | Enc | Dec |  |
| NNIntra+LOP | **-7.6%** | -14.3% | -13.2% | 1.2 | 36 | **-8.6%** | -15.4% | -15.7% | 1.7 | 24 | 21.6 | 1.5 |
| NNIntra+HOP | **-14.2%** | -19.6% | -20% | 2.5 | 1135 | **-13.7%** | -15.9% | -17.1% | 2.5 | 769 | 471 | 2.7 |
| NNIntra+VLOP | **-5.8%** | -6.6% | -5.7% | 1.2 | 17 | **-7.3%** | -8.8% | -8.3% | 1.7 | 13 | 9.9 | 1.4 |
| NNIntra+LOP+SR | **-8.5%** | -12.1% | -10.9% |  |  | **-9.4%** | -12.7% | -13.0% |  |  | 26.4 | 1.5 |

**List of tests**

This round of EE1 tests includes:

* EE1-1: LOP and VLOP in-loop filter
  + EE1-1.1 – **Partial Convolution and Over-Parameterization** 
    - Report: [JVET-AK0106](https://jvet-experts.org/doc_end_user/current_document.php?id=15077) (UESTC), (Transsion), cross-check: [JVET-AK0292](https://jvet-experts.org/doc_end_user/current_document.php?id=15281) (Qualcomm)
  + EE1-1.2 – **LOP with improved Attention and residual group**
    - Report: [JVET-AK0150](https://jvet-experts.org/doc_end_user/current_document.php?id=15121) (Qualcomm), cross-check [JVET-AK0263](https://jvet-experts.org/doc_end_user/current_document.php?id=15252) (Interdigital), JVET-AK0300 (Bytedance)
  + EE1-1.3 - **Multiscale blocks in LOP4 and VLOP3 filters**
    - Report: [JVET-AK0139](https://jvet-experts.org/doc_end_user/current_document.php?id=15110) (UESTC), (Transsion), (Qualcomm), (Nokia), cross-check: [JVET-AK0166](https://jvet-experts.org/doc_end_user/current_document.php?id=15137) (Interdigital)
  + EE1-1.4 – **Cross-component enhanced LOP filter**
    - Report: [JVET-AK0195](https://jvet-experts.org/doc_end_user/current_document.php?id=15166) (Bytedance) cross-check: [JVET-AK0](https://jvet-experts.org/doc_end_user/current_document.php?id=15166)216 (Qualcomm), [JVET-AK](https://jvet-experts.org/doc_end_user/current_document.php?id=15166)xxxx (Ittiam)
  + EE1-1.5 - **Adaptive skip of LOP filtering based on boundary strength partitions**
    - Report: [JVET-AK0093](https://jvet-experts.org/doc_end_user/current_document.php?id=15064) (HYU, ETRI) cross-check: [JVET-AK0241](https://jvet-experts.org/doc_end_user/current_document.php?id=15230) (Huawei)
  + EE1-1.6: **Reduced complexity input feature extraction** – withdrawn
* EE1-2: NN-inter prediction
  + EE1-2.1 **RA/LDB Unified Reference Frame Synthesis for VVC Inter Coding**
    - Report: [JVET-AK0077](https://jvet-experts.org/doc_end_user/current_document.php?id=15048) (Xidian Uni) cross-check: ??? (Nokia)
  + EE1-2.2 –  **Transformer-Based Reference Frame Synthesis for VVC Inter Coding**
    - Report: [JVET-AK0078](https://jvet-experts.org/doc_end_user/current_document.php?id=15049)  cross-check: ???
* EE1-4: NN-based super resolution
  + EE1-4.1 – **A Neural Network Downscaling Filter for RPR**
    - Report: [JVET-AK0235](https://jvet-experts.org/doc_end_user/current_document.php?id=15224) (VIVO), cross-check: [JVET-AK0247](https://jvet-experts.org/doc_end_user/current_document.php?id=15236) (PKU)

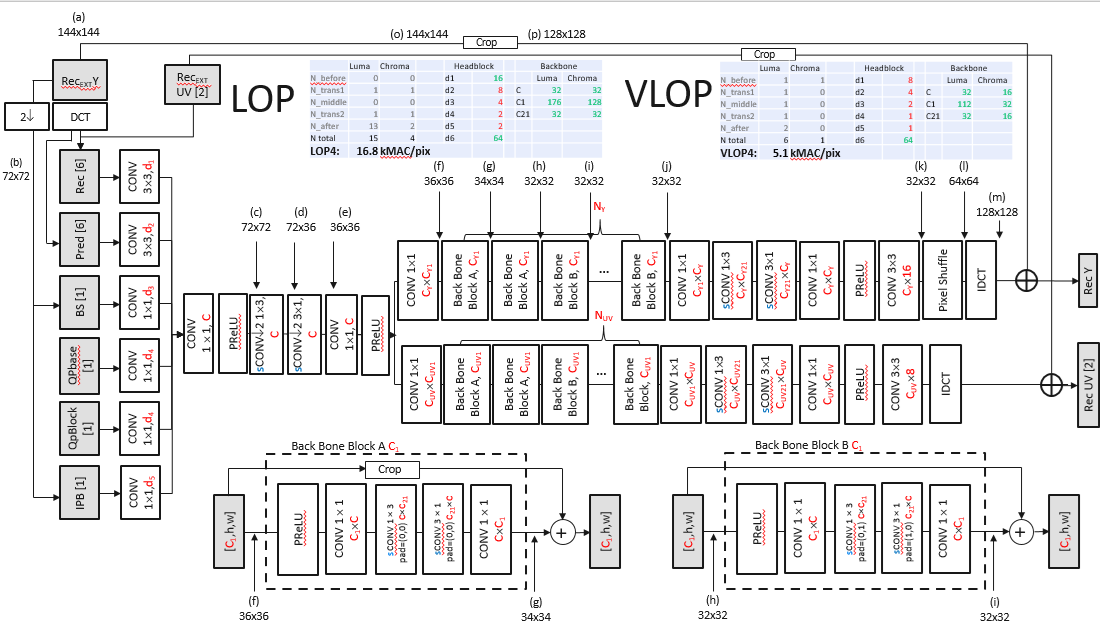
**Test results summary**

Details of each test can be found in attached presentation.

***LOP and VLOP filter modifications***

In NNVC-11 LOP4 and VLOP3 filters architectures are unified. Just number of channels and back bone blocks are different to match two different level of complexity.

Figure 1. LOP4 and VLOP3 filters in NNVC-11



Proponents of tests EE1-1.1-3 provide their solutions for both LOP and VLOP keeping filters unified:

* EE1-1.1 – Partial Convolution and Over-Parameterization, (UESTC), (Transsion)
  + Complexity reduction proposal. Splits input tensor in channel dimension into two parts, one part goes though convolution, another part skips convolution, so decoding run time reduced 23-27%.
* EE1-1.2 – LOP with improved Attention and residual groups (Qualcomm)
  + Performance improvement proposal. For each pair of Back Bone Blocks skip connection, PReLu and attention block are added. Attention block includes: depth-wise convolution blocks, MaxPool operation, HardSigmoid and bi-linear interpolation. For operations not yet in SADL proponent provides SIMD implementation (subject for confirmation by AhG14). Peak memory usage increases ×3 (block level processing). BD-rate reduction 0.3…0.6%.
* EE1-1.3 – Multiscale blocks in LOP4 and VLOP3 filters (UESTC), (Transsion), (Qualcomm), (Nokia)
  + Additionally to EE1-1.1 and EE1-1.2 multi-scale blocks are added (includes EE1-1.1 & EE1-1.2). No memory usage increment compared to EE1-1.2. Froma partially available results peromance benefits are not clear.

Tests EE1-1.4 and EE1-1.5 provide solution and test results only for LOP.

* EE1-1.4 – Cross-component enhanced LOP filter (Bytedance)
  + After the most of NN-based filter operation Luma tensor is added as auxiliary input to the Chroma processing (only ‘tail’ operations of NN-filter use extra input). Approximately 1% Chroma BD-rate gain.
* EE1-1.5 - Adaptive skip of LOP filtering based on boundary strength partitions (HYU, ETRI)
  + NN-filtering operations are by-passed based on number of block boundaries and BS. Decoding run time reduction 35% (LDB) and 23% (RA), almost no drop and gain in Chroma. Non-normative solution is slightly worse.

*Table 2* ***EE1-1: LOP in-loop filter***

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Random Access | | | | | All Intra | | | | | kMAC/pxl | | | Num Param, M | | |
| Y | U | V | EncT | DecT | Y | U | V | EncT | DecT | Total | Filter | Intra | Total | Filter | Intra |
| NNVC-LOP4 | 0.0% | 0.0% | 0.0% |  |  | 0.0% | 0.0% | 0.0% |  |  | 21.6 | 16.8 | 4.8 | 1.5 | 0.248 | 1.3 |
| EE1-1.1.1 |  |  |  |  |  | 0.0% | 0.5% | 0.4% | 101% | 77% | 21.4 | 16.6 | 4.8 | 1.5 | 0.246 | 1.3 |
| EE1-1.1.2 |  |  |  |  |  | 0.0% | -0.1% | -0.3% | 99% | 73% | 21.5 | 16.7 | 4.8 | 1.5 | 0.205 | 1.3 |
| EE1-1.2 | -0.6% | 0.2% | -0.3% |  |  | -0.6% | 0.1% | 0.0% | 99% | 100% | 21.4 | 16.6 | 4.8 | 1.5 | 0.247 | 1.3 |
| EE1-1.3.2 |  |  |  |  |  |  |  |  |  |  | 21.6 | 16.8 | 4.8 | 1.5 | 0.267 | 1.3 |
| EE1-1.4 | 0.0% | -1.5% | -1.3% | 100% | 98% | 0.0% | -0.7% | -0.5% | 100% | 99% | 21.6 | 16.8 | 4.8 | 1.5 | 0.247 | 1.3 |
| EE1-1.5.1.1 | 0.1% | -0.2% | -0.1% | 100% | 77% | 0.0% | -1.5% | -1.0% | 101% | 65% | 21.6 | 16.8 | 4.8 | 1.5 | 0.248 | 1.3 |
| EE1-1.5.2 | 0.2% | -0.1% | 0.0% | 100% | 76% | 0.2% | -1.2% | -0.5% | 101% | 63% | 21.6 | 16.8 | 4.8 | 1.5 | 0.248 | 1.3 |

*Table 3* ***EE1-1: VLOP in-loop filter***

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Random Access | | | | | All Intra | | | | | kMAC/pxl | | | Num Param, M | | |
| Y | U | V | EncT | DecT | Y | U | V | EncT | DecT | Total | Filter | Intra | Total | Filter | Intra |
| NNVC-VLOP3 | 0.0% | 0.0% | 0.0% |  |  | 0.0% | 0.0% | 0.0% |  |  | 9.9 | 5.1 | 4.8 | 1.4 | 0.067 | 1.3 |
| EE1-1.1 |  |  |  |  |  | 0.2% | 0.7% | 0.8% | 100% | 84% | 9.9 | 5.1 | 4.8 | 1.4 | 0.070 | 1.3 |
| EE1-1.2 | -0.3% | 0.6% | 0.8% |  |  | -0.4% | 0.5% | 0.4% | 100% | 104% | 9.8 | 5.0 | 4.8 | 1.4 | 0.069 | 1.3 |
| EE1-1.3.2 | 0.1% | 0.9% | 1.0% |  |  | -0.1% | 0.6% | 0.9% | 97% | 112% |  | 5.1 |  |  | 0.071 |  |

1-1.1.x could be seen beneficial in AI due to complexity reduction, but preliminary RA results (incomplete) indicate significant losses. Applies to both LOP and VLOP.

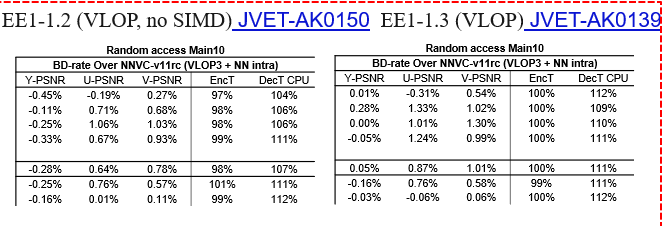
1-1.2 shows interesting gain, but training crosscheck JVET-AK0263 not successful so far (larger deviation than usually seen (in particular different luma/chroma balance). However, different luma/chroma weights were used. Results indicate even loss in performance in case of VLOP). Reason not clear at this moment; might be caused by using different hyperparameters (was later confirmed that these were the same, and that also quantization can be excluded causing it). It was also commented that perhaps only retraining the last iteration might not be sufficient as there are more significant changes relative to the original LOP design. Another training crosscheck JVET-AK0300 is less deviating for LOP with results available so far. However, that second crosscheck was never finalized. Considering the fact that overall gain in luma+chroma is almost identical in cross-check and the deviation in balance can be explained by using different weights in cross-check, the training can be considered successful in the LOP case.

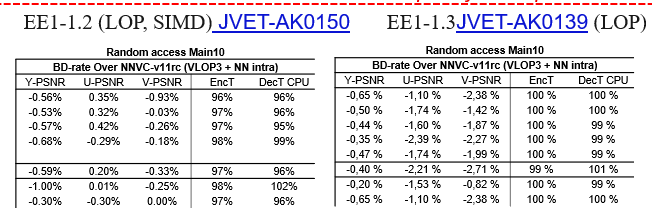
More precise description of the architecture was included in EE summary report v3.

Decision: Adopt JVET-AK0150, case of LOP.

Further study necessary for VLOP case to identify issues with deviation of training crosscheck.

1-1.3 has more complete results in the powerpoint presentation than in the table (results below are from subtest 1.3.2):





For VLOP, worse performance than 1-1.2 (where the latter does not have successful training crosscheck yet). For LOP, less gain in luma, but significantly larger in chroma due to the concept it is understandable that it provides another luma/chroma balance. Training of floating point model was conducted and gave better results than with proponents.

In the original upload, only subtest 1-1.3.2 was reported, in which reparameterization is only applied to luma. Furthermore, the results above are for floating point model, and a new version reports different results with quantized model, with better results for luma, but worse for chroma.

Furthermore, results on 1-1.3.3 are reported, which are better for VLOP and LOP at least for AI and RA. VLOP applies reparameterization also on chroma.

It was asked if the number of channels is multiple of 16 for both luma and chroma.

The following information was requested by proponents of 1-1.2 and 1-1.3 (was provided in EE summary report v3, where EE1-1.3 is a combination from 1.1, 1.2 and multi-scale blocks):

* Exact architecture description for LOP and VLOP, number of channels for each element of the pipeline (was provided, also confirming that number of channels is multiple of 16
* Hyper parameters used for training (by proponents and cross checker)
* Training time compared to current LOP and VLOP (is increased from 2.4h per epoch to 3.4h per epoch, likely caused by overparameterization of 1.1, but 1.3 uses that).
* Possible reasons for deviations of training cross-check (only for 1.2, see notes above)

Training crosscheck JVET-AK0166 was successful for 1.3.2 combination for LOP, not for VLOP (which is understandable as it was also not successful for 1.2).

Concern was raised about the significant increase of training runtime which is caused by overparameterization. The combination of 1.2 and multi-scale blocks from 1.3 and partial convolution from 1.1 (which had also been integerized in 1.3), but without the overparameterization was not tested so far, but might provide similar benefit without increasing training time, but decreasing decoding time. Further investigate this in EE.

To be tested:

LOP5+multi-scale

LOP5+partial convolution

LOP5+partial convolution+multi-scale

The overparameterization of 1.1 is not of interest as long it cannot be conducted without significant training time increase.

(move the following paragraph potentially to discussion of training with new material)

Retraining with new dataset for both LOP4 and LOP5, LOP5 to be confirmed in telco

EE1-1.4 (only relevant for LOP) shows some benefit for chroma, and minor complexity reduction. In the investigated architecture, a latency problem between luma and chroma was resolved. Not applicable to VLOP due to different cropping strategy (earlier cropping only possible in luma).

Number of channels in chroma is increased to 144, but operations are saved due to 1x1 convolution in later stage.

It was reported on Monday 20 Jan. that the training was finished, and inference with the retrained model was matching with small margin (even slightly), also for RA where classes A1 and A2 were missing (but by tendency matching from the QP points that were done.

The training cross-check for 1.4 can be concluded to be successful. It also might be expected to have complementary benefit together with 1.2, which targets improving luma, and the major change is in the chroma branch. However, 1.4 also has an element that modifies the kernel size of convolution from 3x3 to 1x1 in one layer, which should not have relevant impact on the performance.

Decision: Conditional adoption of JVET-0195 EE1-1.4, to be confirmed in telco, provided that the combined model is retrained and crosschecked and performance is as expected. This should still be conducted with the old training data set

In a follow-up discussion on Tue. 21 Jan. proponents 1.2 and 1.4 should reported that the complexity of the combination would still be lower than LOP4 (16.6 kMAC/pix vs. 16.8 of LOP4) A model had been sent to the SW coordinator for checking this in SADL.

If conditional adoption would be confirmed, LOP5=LOP4+EE1.2+EE1.4

Otherwise, LOP5=LOP4+EE1-1.2

Proponents of 1.3 requested also to be included in this investigation. In a follow-up discussion on Tue. 21 Jan., E. Alshina provided an analysis of the architectural changes necessary for that by showing a first version of EE1 description. The necessary changes appeared significantly larger than for the combination 1.2/1.4. Moreover, in that case, 1.2/1.3\*, 1.2/1.3\* and 1.2/1.3\*/1.4 would need to be trained and tested. 1.3\* is only the multi-scale aspect of 1.3, which had been studied and training cross-checked standalone in a previous EE round, but based on LOP3. Further investigate in EE, likely to give benefit also on top of LOP5, as it modifies a different part of chroma branch.

Further study on the aspect of using this in VLOP is recommended.

It was commented that any of the proposals so far would cause larger deviation between LOP and VLOP, where it would however be questionable if a harmonization between the two is really relevant in terms of implementation.

EE1-1.5 is not changing the model, but skips its usage by blockwise decision, which reduces decoding time. Most effective for screen content class F. Worst case complexity would not be reduced, and logic would need to be added if it was normative.

It is interesting to see that NN filter is not very beneficial in quite some cases. It was commented that potentially using it during training would even provide benefit in performance. It was however also commented that decoding runtime is not a serious bottleneck in the context of the exploration.

Decision(SW/non-CTC): Adopt EE1-1.5.2 encoder-only option.

***HOP filter modifications***

There are no tests for HOP in this round of EE1.

***NN-Inter***

Two versions of NN-based reference frame generation are studied:

* EE1-2.1 URFS Unified Reference Frame Synthesis (649 kMAC/pxl, 1.6M param)
* EE1-2.2 TRFS Unified Reference Frame Synthesis (261 kMAC/pxl, 2.0 M param)

Training performed on Vimeo 90K (15K video scenes) triplet and BVI-DVC (200 video scenes). Training on Vimeo 90K triplet resulted in 4..5% higher BD-rate.

Variant with transformers (based on partially available data) has 1.5% lower compression performance (likely training set is not big enough for 2M parameter model).

Performance of both methods in presence of NNVC tools is 12% lower than in absence NNVC tools (test on top of VTM).

No decoder run time reported due to decoder crash. It was commented that this may be due to a modification of GOP structure since NNVC4. Encoder runtimes also appear unreasonable in comparison to NNVC anchor.

Too early for taking action. Training crosscheck would be desirable.

More results were shown on Tue. 21 Jan., but not fully complete. The tendency that results are significantly worse than Vimeo90K training is still observed. To be further investigated in EE with extended training data set.

***NN-based super resolution (NNSR)***

Existing in NNVC-11.0 NNSR operated as a combination of adaptive resolution coding (scaling factors 1.5 and 2 both supported) and NN-based post-filter for up-sampling.

In tests EE1-4.1:

* EE1-4.1.1 scaling ratio is only 2, NN-based down-sampling has been added.
* EE1-4.1.2 scaling ratio is only 2, NN-based down-sampling and up-sampling filters are trained together,
* EE1-4.1.3 scaling ratio 1.5 added

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Test | Random Access | | | | | All Intra | | | | |
| Y | U | V | EncT | DecT | Y | U | V | EncT | DecT |
| EE1-4.1.1 | -0.01% | -0.10% | -0.09% | ?? | ?? | -0.04% | -0.13% | -0.11% | ?? | ?? |
| EE1-4.1.2 | 0.00% | -0.11% | -0.08% | ?? | ?? | -0.05% | -0.14% | -0.08% | ?? | ?? |
| EE1-4.1.3 | 0.76% | 12.45% | 8.52% | ?? | ?? | 0.03% | 10.15% | 7.28% | ?? | ?? |

Minor gain, mostly in Chroma. EE1-4.1.3 (scaling ratio 1.5) has a big loss likely due to a bug.

Further study recommended.

[JVET-AK0050](https://jvet-experts.org/doc_end_user/current_document.php?id=15020) [AHG11] [AHG14] Teleconference on NNVC [E. Alshina, F. Galpin]

### EE1 contributions: Neural network-based video coding (9)

Contributions in this area were discussed in the context of the EE summary report JVET-AK0023.

[JVET-AK0077](https://jvet-experts.org/doc_end_user/current_document.php?id=15048) EE1-2.1: RA/LDB Unified Reference Frame Synthesis for VVC Inter Coding [Q. Qin, C. Jung (Xidian Univ.)] [late]

Initial version rejected as “placeholder”

[JVET-AK0315](https://jvet-experts.org/doc_end_user/current_document.php?id=15304) Crosscheck of JVET-AK0077 (EE1-2.1: RA/LDB Unified Reference Frame Synthesis for VVC Inter Coding) [L. Murn (Nokia)] [late]

[JVET-AK0078](https://jvet-experts.org/doc_end_user/current_document.php?id=15049) EE1-2.2: Transformer-Based Reference Frame Synthesis for VVC Inter Coding [Q. Qin, C. Jung (Xidian Univ.)] [late]

Initial version rejected as “placeholder”

[JVET-AK0316](https://jvet-experts.org/doc_end_user/current_document.php?id=15305) Crosscheck of JVET-AK0078 (EE1-2.2: Transformer-Based Reference Frame Synthesis for VVC Inter Coding) [L. Murn (Nokia)] [late]

[JVET-AK0093](https://jvet-experts.org/doc_end_user/current_document.php?id=15064) EE1-1.5: Adaptive skip of LOP filtering based on boundary strength partitions [H. Kwon, J. Seo, H. Ko (HYU), D. Kim, S.-C. Lim (ETRI)]

[JVET-AK0241](https://jvet-experts.org/doc_end_user/current_document.php?id=15230) Crosscheck of JVET-AK0093 (EE1-1.5: Adaptive skip of LOP filtering based on boundary strength partitions) [J. Pardo, E. Alshina (Huawei)] [late]

[JVET-AK0106](https://jvet-experts.org/doc_end_user/current_document.php?id=15077) EE1-1.1: Partial Convolution and Over-Parameterization [J. Chi, A. Li, C. Zhu, L. Luo, H. Guo (UESTC), Y. Huo, Y. Liu (Transsion)]

[JVET-AK0292](https://jvet-experts.org/doc_end_user/current_document.php?id=15281) Crosscheck of JVET-AJ0106 (EE1-1.1 Partial Convolution and Over-Parameterization) [J. Wang (Qualcomm)] [late]

[JVET-AK0139](https://jvet-experts.org/doc_end_user/current_document.php?id=15110) EE1-1.3: Multiscale blocks in LOP4 and VLOP3 filters [R. Yang, M. Santamaria, F. Cricri, H. Zhang, J. Lainema, M. M. Hannuksela (Nokia), Y. Li, D. Rusanovskyy, S. Eadie, M. Karczewicz, J. Wang, L. Kerofsky (Qualcomm), J. Chi, A. Li, C. Zhu, L. Lei, H. Guo (UESTC), Y. Huo, Y. Liu (Transsion)]

[JVET-AK0166](https://jvet-experts.org/doc_end_user/current_document.php?id=15137) Crosscheck of JVET-AK0139 (EE1-1.3: multiscale blocks in LOP4 and VLOP3 filters) [T. Dumas (InterDigital)] [late]

[JVET-AK0150](https://jvet-experts.org/doc_end_user/current_document.php?id=15121) EE1-1.2: LOP with improved Attention and residual groups [Y. Li, D. Rusanovskyy, S. Eadie, M. Karczewicz, J. Wang, L. Kerofsky (Qualcomm)]

[JVET-AK0263](https://jvet-experts.org/doc_end_user/current_document.php?id=15252) Crosscheck of JVET-AK0150 (EE1-1.2: LOP with improved Attention and residual groups) [Z. Ameur, T. Dumas, F. Galpin (InterDigital)] [late]

[JVET-AK0300](https://jvet-experts.org/doc_end_user/current_document.php?id=15289) Crosscheck of JVET-AK0150 (EE1-1.2: LOP with improved attention and residual groups) [Y. Li (Bytedance)] [late] [miss]

[JVET-AK0195](https://jvet-experts.org/doc_end_user/current_document.php?id=15166) EE1-1.4: Cross-component enhanced LOP filter [Y. Li, J. Li, C. Lin, K. Zhang, L. Zhang (Bytedance)]

[JVET-AK0216](https://jvet-experts.org/doc_end_user/current_document.php?id=15205) Crosscheck of JVET-AK0195 (EE1-1.4: Cross-component enhanced LOP filter) [Y. Li (Qualcomm)] [late] [miss]

[JVET-AK0301](https://jvet-experts.org/doc_end_user/current_document.php?id=15290) Crosscheck of JVET-AK0195 (EE1-1.4: Cross-component enhanced LOP filter) [J. N. Shingala, A. Shyam (Ittiam)] [late]

[JVET-AK0235](https://jvet-experts.org/doc_end_user/current_document.php?id=15224) EE1-4.1: A Neural Network Downscaling Filter for RPR [Y.-Q. Zhu, W.-X. He, X. Li, J.-D. Y, Q. Liu (HUST), Z. Lv (vivo)] [late]

[JVET-AK0247](https://jvet-experts.org/doc_end_user/current_document.php?id=15236) Crosscheck of JVET-AK0235 (EE1-4.1: A Neural Network Downscaling Filter for RPR) [Z. Li, J. Zhang (PKU)] [late]

[JVET-AK0212](https://jvet-experts.org/doc_end_user/current_document.php?id=15201) EE2-2.1: OBMC modifications [R. Yu, H. Huang, C.-C. Chen, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-AK0256](https://jvet-experts.org/doc_end_user/current_document.php?id=15245) Crosscheck of JVET-AK0212 (EE2-2.1: OBMC modifications) [Y. Kidani, K. Kawamura (KDDI)] [late]

[JVET-AK0290](https://jvet-experts.org/doc_end_user/current_document.php?id=15279) Crosscheck of Test 2.1a and b in JVET-AK0212 (EE2-2.1: OBMC modifications) [Z. Deng (Bytedance)] [late] [miss]

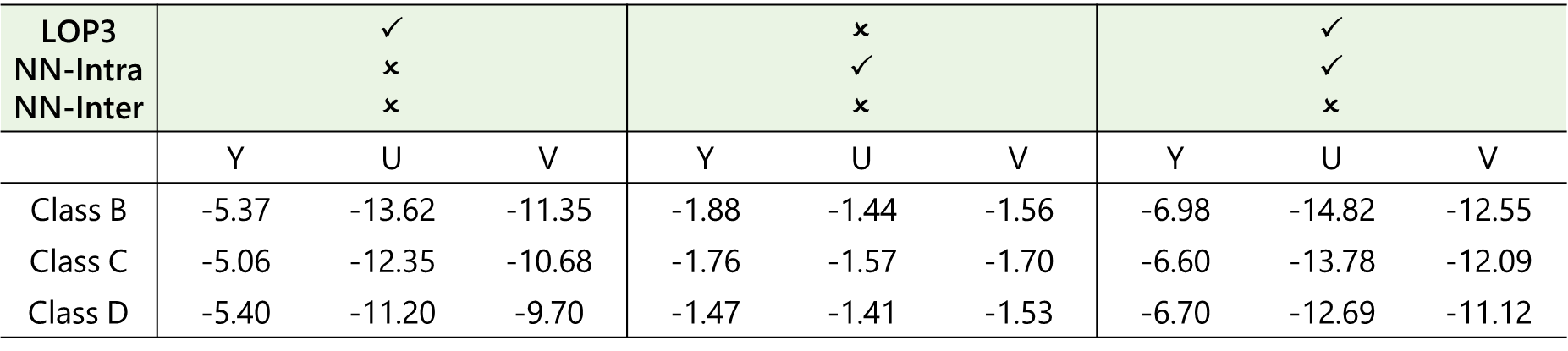
### EE1 related and beyond-EE contributions: Neural network-based video coding (10)

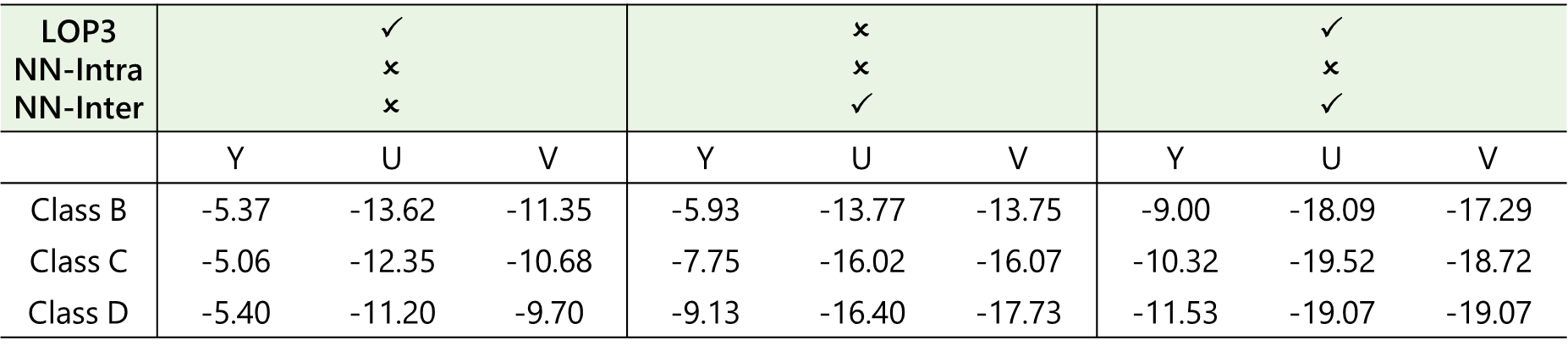
Contributions in this area were discussed at 1650–2005 on Thursday 16 Jan. 2025 (chaired by JRO).

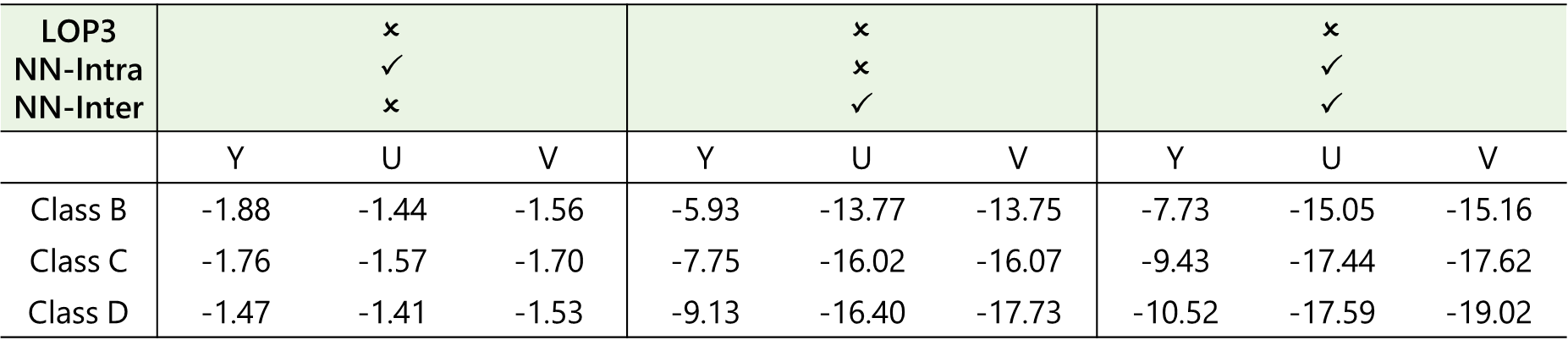
[JVET-AK0079](https://jvet-experts.org/doc_end_user/current_document.php?id=15050) EE1-Related: Combination Test of NNVC Tools and TRFS [Q. Qin, C. Jung (Xidian Univ.)]

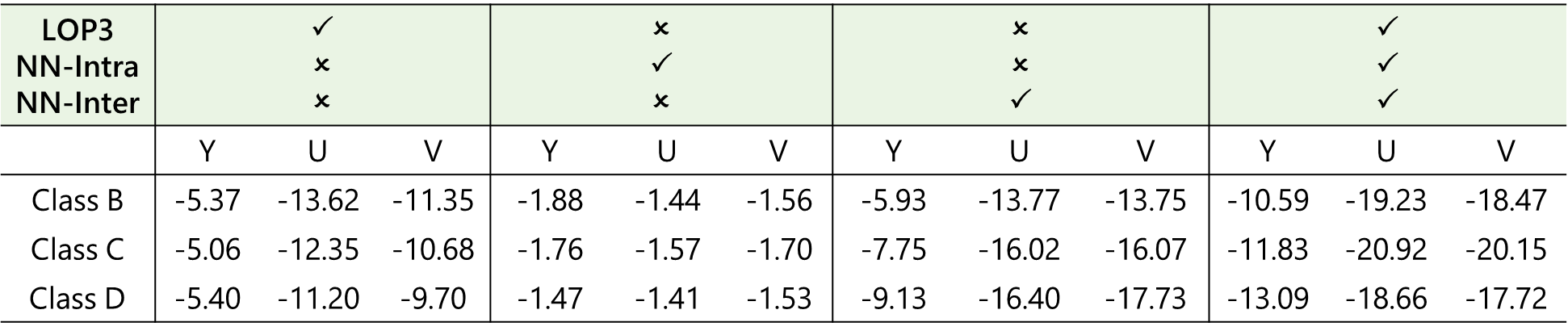
This contribution reports an exploration experiment on the combination of NNVC tools (i.e., LOP and NN-Intra) and TRFS (i.e., NN-Inter) to observe the complementarity and overlap by various combinations. Specifically, this contribution sets up four groups of experiments, including “LOP + NN-Intra”, “LOP + NN-Inter”, “NN-Intra + NN-Inter”, and “LOP + NN-Intra + NN-Inter”. All experiments take VTM-11.0\_NNVC-10.0 with NNVC tools disabled as the reference anchor. The following conclusions can be drawn from the experiments:

* **Complementarity of LOP3 and NN-Intra:** JVET has widely studied the coding performance achieved by LOP3 and NN-Intra tools. Although LOP3 and NN-Intra sequentially enhance the coding signals, the BD-rate improvements appear to be nearly additive. This behavior can be attributed to the limited proportion of I-frames directly influenced by NN-Intra in a to-be-encoded video sequence. Therefore, no performance overlap is observed when combining LOP3 with NN-Intra.
* **Interaction between LOP3 and NN-Inter:** The coding performance of combining LOP3 with NN-Inter is significantly lower than the sum of their individual contributions, suggesting a substantial performance overlap between these two coding tools. This is primarily attributed to the fact that LOP3 and TRFS perform signal enhancement in a series manner. Specifically, the decoded frames processed by TRFS are sourced from the decoded picture buffer (DPB), which in turn contains frames that have already been enhanced by LOP3. This dependency inherently limits the incremental gains achievable by their combination.
* **Complementarity of NN-Intra and NN-Inter:** In contrast, the combination of NN-Intra and NN-Inter yields nearly additive coding gains, highlighting their complementary roles in improving VVC performance. This is primarily because these two tools operate independently in parallel, targeting distinct aspects of the encoding process without significant functional overlap.
* **Combining LOP3, NN-Intra and NN-Inter:** Based on the above analysis, it is evident that combining LOP3, NN-Intra, and TRFS introduces considerable overlap in coding performance, particularly for chroma components (i.e., U and V). For example, in Class C sequences, the performance overlap for the luma (Y) component is approximately 2.74%, whereas the overlaps for the U and V components are notably higher, at 9.02% and 8.30%, respectively. This observation raises the need for further optimization to minimize redundancy when integrating these tools.





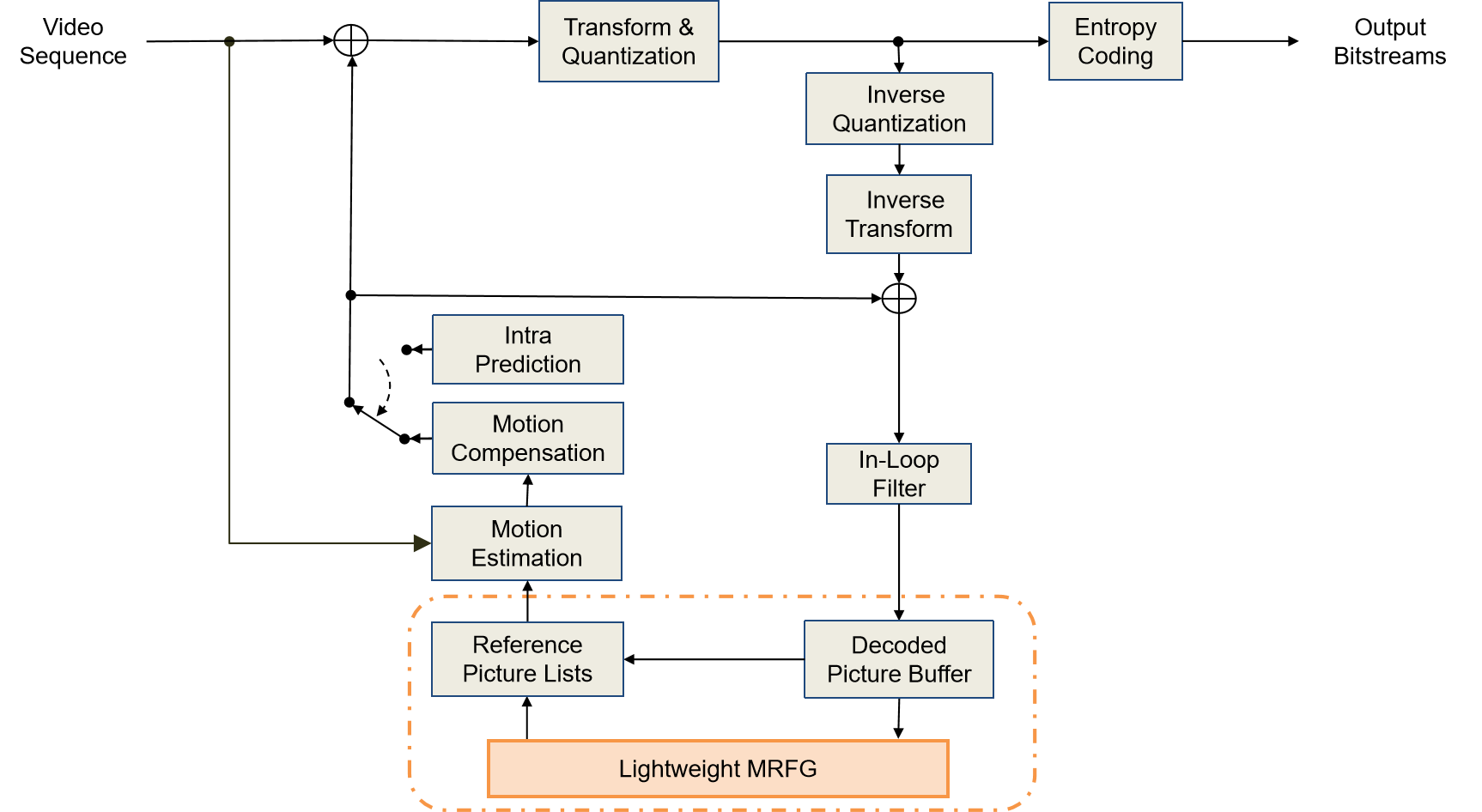


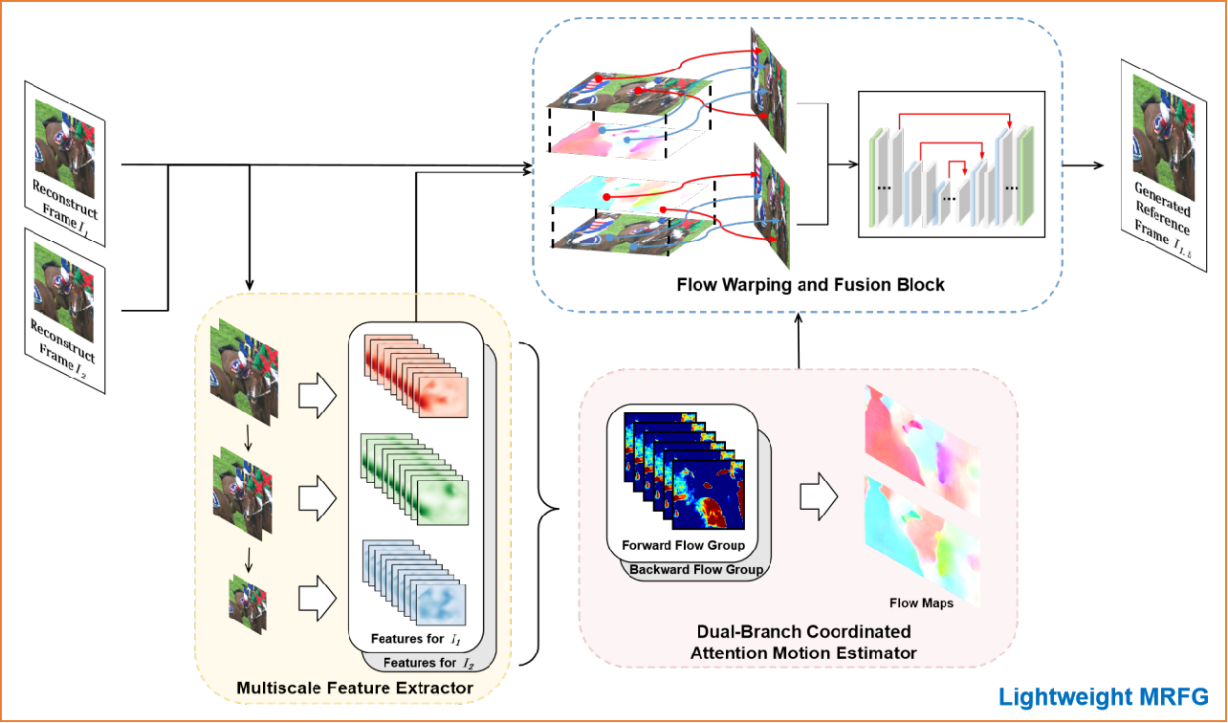


Interesting study showing overlap (not-additive gain) between LF and prediction networks in particular. As already the loop-filtered input is used for training predictors, other approaches of joint optimization might be necessary to consider.

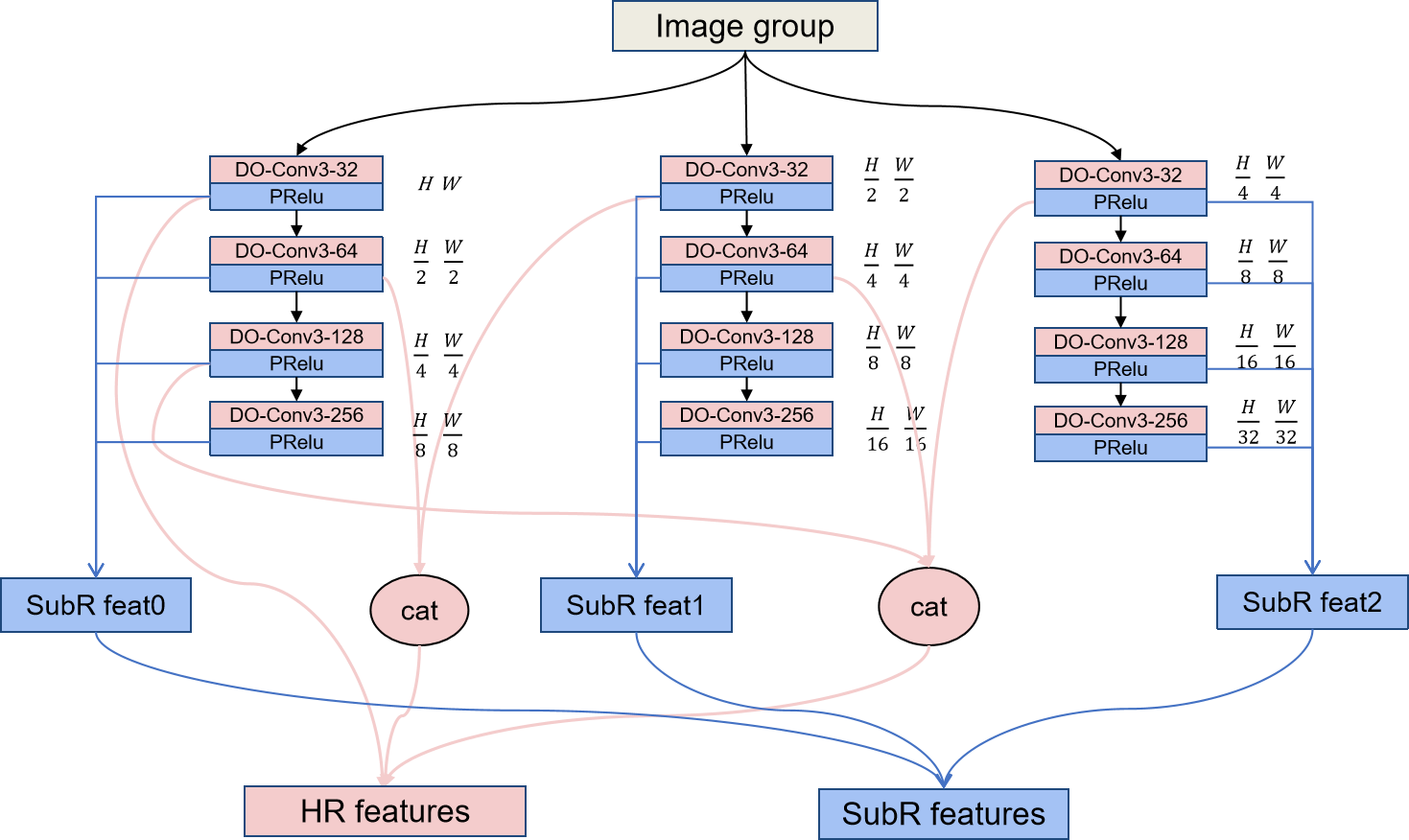
[JVET-AK0071](https://jvet-experts.org/doc_end_user/current_document.php?id=15042) AHG11: Lightweight Multiscale Reference Frame Generation for VVC Inter Coding [P. Li, C. Jung, Q. Qin (Xidian Univ.)]

This contribution presents lightweight multiscale reference frame generation for VVC inter coding, named LMRFG. LMRFG synthesizes a high-quality reference frames from two reconstructed frames in the decoded picture buffer (DPB) by learning optical flow prediction and warped feature fusion. The synthesized frame is inserted into the reference picture list (RPL) and considered as a candidate reference frame for VVC inter frame coding. For random access (RA) and low-delay B (LDB) configurations, LMRFG adopts the same network architecture for end-to-end optimization. Based on the anchor VTM-11.0\_NNVC-10.0 (NN-tools on), LMRFG achieves bitrate reductions of **{RA: 4.31%, 6.54%, 7.05%}** and **{LDB: 3.81%, 8.90%, 8.77%}** for {Y, U, V} components, respectively. Based on the anchor VTM-11.0\_NNVC-10.0 (NN-tools off), LMRFG achieves bitrate reductions of **{RA: 7.59%, 13.03%, 12.87%}** and **{LDB: 6.15%, 16.38%, 15.24%}** for {Y, U, V} components, respectively.

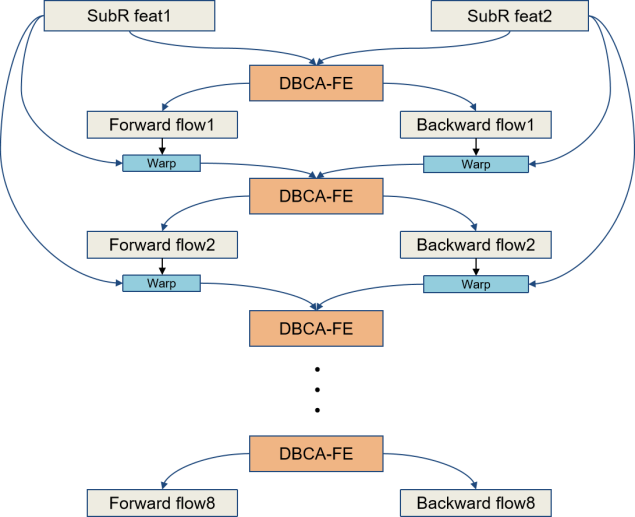
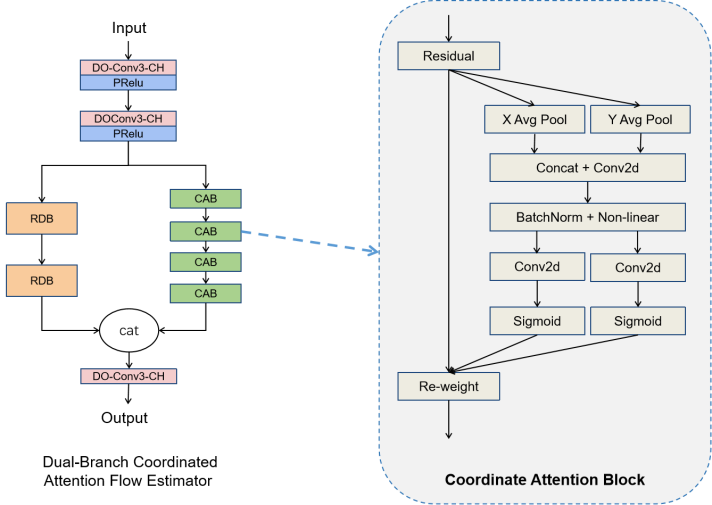




**Fig. 1.** Illustration of the proposed LMRFG framework integrated with VVC. LMRFG is mainly comprised of multiscale feature extractor (MFE), dual-branch coordinated attention motion estimator (DBCA-ME) and flow warping and fusion block (FWFB).

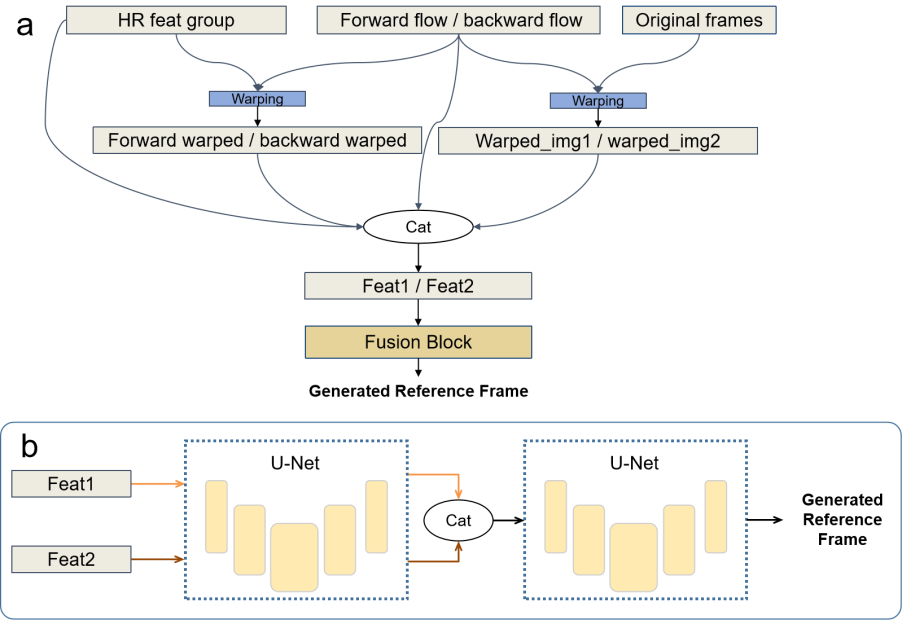


**Fig. 2.** Network architecture of the multiscale feature extractor (MFE).

(a) Overall network for DBCA-ME. (b) Details of dual-branch coordinated attention flow estimator.

**Fig. 3.** Network architecture of the dual-branch coordinated attention motion estimator (DBCA-ME).



**Fig. 4.** (a) Network architecture of the flow warping and fusion block (FWFB). (b) Fusion block in (a).

|  |  |  |
| --- | --- | --- |
| **Network Information in Inference Stage** | | |
| Mandatory | HW environment | |
| GPU Type | GPU (RA) GPU (LDB) |
| Framework | PyTorch v2.0.0 |
| Number of GPUs per Task | 0 |
|  |  |
| Number of Parameters (Each Model) | RA: 2.01M LDB: 2.01M |
| Total Number of Parameters (All Models) | 2 x 2.01M |
| Parameter Precision (Bits) | 32 |
| Memory Parameter (MB) | RA: 24.98M LDB: 24.98M |
| Multiply Accumulate (kMAC/pixel) | RA: 546.65K LDB: 546.65K |
|  |  |
| Optional | Total Conv. Layers |  |
| Total FC Layers |  |
| Total Memory (MB) |  |
| Batch size | 1 |
| Patch size | 256x256 |
| Changes to network configuration or weights required to generate rate points |  |
| Peak Memory Usage (Total) |  |
| Peak Memory Usage (per Model) |  |
| Border handling |  |
| Other information |  |
|  |  |

Gain higher than EE, complexity lower. Training set the same – Vimeo90K.

Training was done on sequences encoded with other NNVC tools off – according to proponents, that gave better results.

Not yet implemented in SADL.

Worse performance with larger resolutions – could this be caused by the fact that the NN may not be able to capture wider ranges of motion? It might be useful to analyse the properties of the generated reference picture compared to the current picture, how much rate is consumed by motion, residual etc.

Study in EE as replacement for the previous inter prediction method. Also study training with other material such as BVI, BVI-AOM.

See further notes under JVET-AK0255 on aspects of training.

[JVET-AK0119](https://jvet-experts.org/doc_end_user/current_document.php?id=15090) AHG11: NNLF LOP3 improvement with parallel 1x3/3x1 Backbone [T. Shao, P. Yin, S. McCarthy (Dolby), J. N. Shingala, A. Shyam, A. Suneja, S. Badya (Ittiam)]

This contribution proposes to replace the sequential 1x3/3x1 backbone block design in LOP3 with parallel 1x3/3x1 backbone design. It aims at reducing the number of sequential convolutional layers while achieving some coding efficiency. The same number of kMAC/pixel is maintained.

Implemented on top of NNVC10 using SADL, the fast stage 3 training results show that the BD-rate for fp32 is {-0.08%, -2.34%, -3.01%} under AI and {-0.15%, -2.49%, -2.95%} under RA when compared with NNVC10 anchor. The results show that the BD-rate for int16 is {-0.12%, -2.19%, -2.70%} under AI and {-0.09%, -2.22%, -3.03%} under RA when compared with NNVC10 anchor.

Why does chroma have so much larger gain? Not known

Less convolutional stages - lower latency is advantage

Can be expected to be applicable with LOP4.

Several experts expressed interest to investigate in EE on top of LOP4.

It was also suggested that a further study to implement such an approach in VLOP (not in EE) would be interesting.

[JVET-AK0146](https://jvet-experts.org/doc_end_user/current_document.php?id=15117) [AHG11] A Hybrid Framework Integrating End-to-End Learned Image Codec with Conventional Codec [N. Zou, A. Hallapuro, F. Cricri, H. Zhang, M. M. Hannuksela (Nokia)]

This contribution proposes a hybrid framework that integrates end-to-end learned image compression (LIC) methods with conventional compression techniques. The framework involves using LIC-coded intra frames and VTM-coded inter frames. Furthermore, the encoder decides whether to use the LIC-coded intra frames or the VTM-coded intra frames. With this hybrid framework, under the Random-Access configuration, the resulting BD-rates over NNVC-7.1 VTM (with NN tools off) anchor are reported to be as follows:

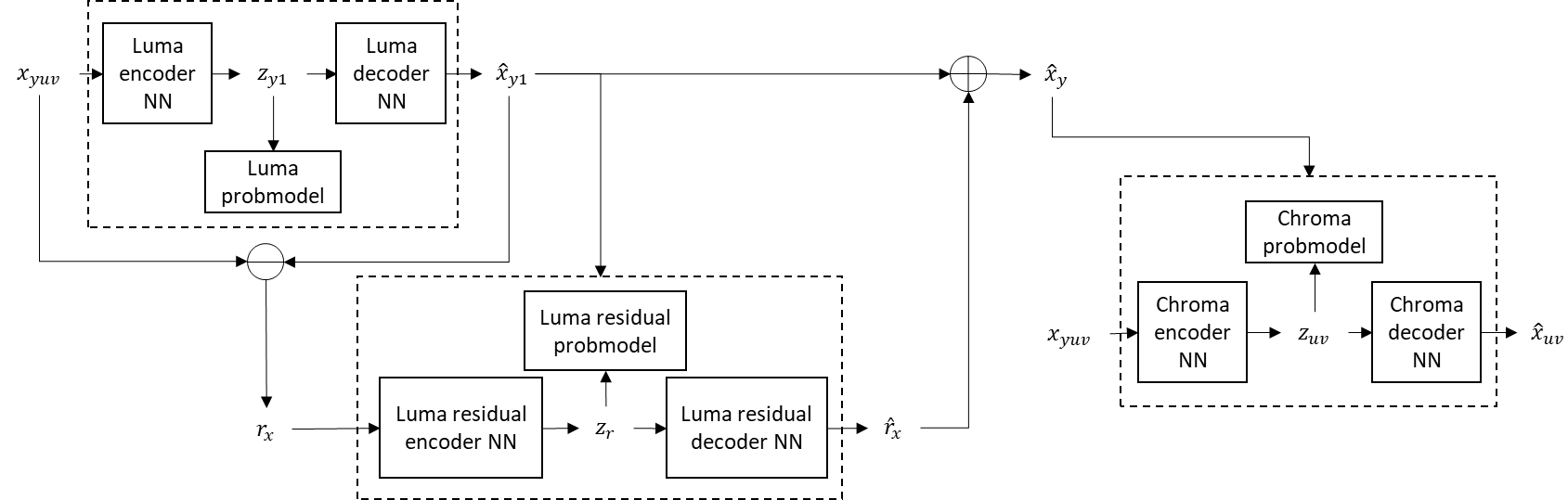
Class A1 0,00% (Y), 0,00% (Cb), 0,00% (Cr)

Class A2 -0,27 % (Y), -0,34 % (Cb), -0,35 % (Cr)

Class B -0,18 % (Y), -8,81 % (Cb), -8,35 % (Cr)

Class C 0,22 % (Y), -7,60 % (Cb), -7,32 % (Cr)

Class D -0,70 % (Y), -11,48 % (Cb), -11,34 % (Cr)



Very large network, 3.3 MMAC/s encoder, 8.5 for decoder, exact architecture not described.

Training uses BVI, DIV2K, JPEG-AI, CLIC and some other

Proponents are requested to upload the presentation (in v2 another version was provided)

Encoder can switch between conventional and learned method for the entire I frame.

Further study highly encouraged. More precise description should be provided, and complexity reduction, reduction of number of models is recommended.

[JVET-AK0175](https://jvet-experts.org/doc_end_user/current_document.php?id=15146) AhG11: Dimension-wise decomposed multiplier for content-adaptive loop-filter [Z. Xu, J. Konieczny (TCL)]

It is proposed to use a dimension-wise decomposed representation of layer multipliers, for the purpose of content adaptation of Neural Network-based Loop Filter (NNLF).

The performance of this approach under RA configuration is reported as followed.

* With respect to NNVC-9.1 with NNIntra + LOP.3 ON, the average gain is reported to be:

{-1.01%, -5.28%, -6.35%, EncT: XXX, DecT: XXX}

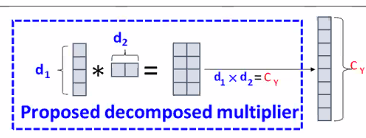
* With respect to NNVC-9.1 with NNIntra + CA-LOP.3 ON, the average gain is reported to be:

{-0.08%, -1.31%, -1.66%, EncT: XXX, DecT: XXX}

It is noted that the proposal is a modification of content-adaptive LOP3 (already in NNVC, disabled by default) by implementing dimension-wise decomposed representation of layer multipliers. The results in the first bullet above is comparing against the NNVC anchor, the second bullet shows results against NNVC anchor plus content adaptive LOP3 from NNVC enabled. This is the actual gain of the proposal

Preliminary results compared to NNVC 10 are also presented.

It was asked which scripts were used for content adaptation. A modified version of the script used in NNVC for content-adaptive LOP.



The BD rate gain comes from the fact that the decomposed multiplier requires less parameters.

Investigate in EE, reporting results relative to the existing content adaptive filter. Not only BD rate should be compared, but also the impact on the quality of the filter output, e.g. by comparing the average PSNR for the GOPs for which the adaptation is done.

This should be done on top of adaptive LOP4 from JVET-AK0311.

[JVET-AK0177](https://jvet-experts.org/doc_end_user/current_document.php?id=15148) AHG 11: Neural Network Coded Reference Frame for Intra Coding [F. Brand, T. Solovyev, E. Alshina (Huawei)]

This contribution presents a method to combine end-to-end trained image compression methods with long-proven methods from VTM. The method comprises using an end-to-end coded image as reference image for a modified I-Frame. The encoder can then decide whether and how to use the NN-reference frame. In this method data transfer from a GPU/NPU is only necessary on full frame level into the decoded picture buffer, thus solving a major problem in realizing neural-network-based components in hardware while using specialized hardware. Compared to the previous contribution JVET-AJ0208 the method was harmonized with LMCS, showing similar gains. This contribution presents two variants, one where the encoder checks for each frame, whether the NN-reference frame is beneficial and should be used and one where this decision is based purely on the QP. In the first case gains of -4.21%/-7.07%/-3.63% in AI can be achieved with currently 128% encoder runtime. In the second case the gains are -3.29%/-0.43%/0.10% with 48% encoder runtime.

*In the current version, the runtimes are not reliable, since potentially multithreading was used on the cluster. Runtimes affected by this are marked with an asterisk (\*) in the tables. We will update the runtimes before the presentation.*

(new version not yet uploaded – update abstract from v2)

Used for QP>22, otherwise VTM intra.

It was commented that more sequence adaptation might be better. In some cases, luma and chroma are quite unbalanced.

The “intra difference coding” was newly implemented in VTM. It was pointed out that multi-layer VVC could be used, however RDO would need to be modified and aligned between NN and VVC. Frame-level RDO is used, with first pass comparing cost of VVC and NN

Investigated with ECM? No.

RA results? Ongoing.

Why not JPEG-AI? Not optimized for PSNR.

Would be interesting also using other metrics.

Further study.

[JVET-AK0242](https://jvet-experts.org/doc_end_user/current_document.php?id=15231) AHG11: Enhancing super-resolution with residual in NNVC [T. Yang, X. Li, W.-X. He, Y.-Q. Zhu, Q. Liu (HUST), Z.-Y. Lv (vivo)] [late]

This contribution proposes a method in which residual information is incorporated as an additional input to the super-resolution (SR) model, aiming to further improve the coding performance of the SR model. The proposed model is specifically designed for 2.0x upscaling.

All tests are conducted using the NNVC11.0 Anchor with NN tools (Intra Pred + LOP4), and enabling the singleRatio configuration.

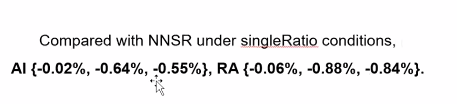
The overall BD-rate savings are reported as follows:

The mode of singleRatio: AI { -0.69%, 1.77%, 2.16%} and RA { -0.83%, 1.76%, 1.90%}.

Model has 4.8 kMAC/p (slightly higher than current SR from NNVC)

It was asked which residual is used. This is the difference between prediction (inter/intra) and reconstruction before deblocking, which should be identical to the result of the inverse transform as far as rounding/clipping is ignored. Could there be an effect of LMCS?

Results compared to current NNSR:



Gain is comparably low, and more memory would be required (however also needed for NNLF). Results with same or lower complexity would be requested. Higher gain in chroma would be more convincing.

Would it be possible to also support ratio 1.5x?

Further study on the aspects above before starting EE.

[JVET-AK0311](https://jvet-experts.org/doc_end_user/current_document.php?id=15300) AHG11/AHG14: Content-adaptive LOP4 filter [M. Santamaria, R. Yang, F. Cricri, M. M. Hannuksela, D. Bugdayci Sansli, A. Hallapuro, J. Lainema, H. Zhang (Nokia)] [late]

This contribution reports the content adaptation per random-access segment for LOP4 loop-filter. This tool has previously been studied for LOP2, LOP3 and VLOP loop-filters. The content-adaptive parameters are signalled via a specifically-designed Adaption Parameter Set (APS) called Neural Network Filter Update (NNFU). It is reported that applying this approach in the RA configuration yields average coding gains of -0.96% (Y), -4.42% (Cb), and -4.57% (Cr) compared to NNVC-11.0 with NN intra tool ON and LOP4 tool ON.

Optimization script and strategy is same as before, just adapted to LOP4. Slightly better gain than the gain from adaptive LOP3 (JVET-AI0111).

Decision: Adopt JVET-AK0311 (non-CTC)

As a general comment, it would be beneficial to make an investigation about the visual impact of NNLF (both adaptive and non-adaptive).

A demo on real-time decoding of adaptive NNLF is announced for Friday during the Workshop.

[JVET-AK0312](https://jvet-experts.org/doc_end_user/current_document.php?id=15301) AHG11: Conditional loop-filter [M. Santamaria, F. Cricri (Nokia)] [late]

This contribution studies using the LOP4 loop-filter conditionally to generate the filtered luma and chroma together or independently. For that purpose, the usage of the loop-filter is signalled separately for luma and chroma components at block level or frame level. The following PSNR BD-rates are reported with respect to NNVC-v11rc: 0.00% (Y), -0.34% (U), -0.38% (V) in RA, 0.03% (Y), -3.93% (U), -3.59% (V) in LDB, 0.07% (Y), -3.86% (U), -4.02% (V) in LDP and -0.01% (Y), -0.04% (U), 0.02% (V) in AI.

Decoder run time is increased because there may be additional cases where the LF is enabled only for luma or only for chroma. The SADL implementation in such cases would run for all components, and ignore the output that is not intended.

It was commented that SADL already has a feature for not running stages in an NN when certain output is not needed. Could be used for that purpose.

The separate signalling is done at slice level and block level by performing separate RDO decision.

Investigate in EE. In this context, the reduction of decoding time should also be investigated. This could be achieved by adapting SADL; also combination with the approach from EE1-1.5.2 should be investigated. It can be expected that in general more blocks would be filtered than in the case of joint signalling, but by avoiding some filtering of blocks where the benefit is not large this could be reduced.

### SADL and NNVC implementation, CTC (4)

Contributions in this area were discussed at 0900–1110 on Friday 17 Jan. 2025 (chaired by JRO).

[JVET-AK0134](https://jvet-experts.org/doc_end_user/current_document.php?id=15105) AhG14: SADL update [F. Galpin (InterDigital)] [late]

This contribution presents updates in the Small AdHoc Deep Learning (SADL) library. The update contains both some fixes and new features.

**New in version v11**

***Improvements***

* allow reloading models
* more accurate mac count
* speed up for lop/vlop
* refactoring conv2d
* refactoring unit tests

***New***

* analysis in/out sizes
* packed sparsity
* v5 format
* model info embedded in the model

***Fixes***

* fix dump slice
* fix stack overflow
* fix mismatch onnx conversion
* fix gcc 13
* fix clang

It was asked why the decoding runtime increased in LOP4 compared to LOP3. This is still under investigation, needs to be identified whether further optimization outside or inside of SADL needs to be done. Theoretically, it should have same runtime.

All changes in SADL are agreed.

[JVET-AK0151](https://jvet-experts.org/doc_end_user/current_document.php?id=15122) AhG14: On the impact of explicit full SIMD implementation for 1x3 and 3x1 DW convolution [Y. Li, D. Rusanovskyy, M. Karczewicz (Qualcomm)]

In this document, the performance of an explicit full SIMD implementation for the 1x3, 3x1depth-wise convolution versus the compiler-optimized partial SIMD implementation with auto-vectorization is compared. The results show an improvement of both encoding and decoding time for the extended implementation.

It is suggested to change the current auto-vectorization to the full explicit SIMD implementation for these operators in SADL.

It was commented that this is beneficial, though only for x86 processors.

Decision(SW): Adopt JVET-AK0151 (merge request already issued).

[JVET-AK0255](https://jvet-experts.org/doc_end_user/current_document.php?id=15244) [AHG11] Response to Call for training materials for neural network-based video coding tool development [F. Zhang, D. Bull, J. Nawala, Y. Jiang, X. Zhu, J. Sole, E. Alshina] [late]

This contribution proposes to extend NNVC training data set and use BVI -AOM instead of BVI-DVC. Compression performance improvement for AI based algorithms is expected. Copyright conditions are more flexible than for BVI-DVC.

Total of 239 sequences, 78 new compared to BVI-DVC. The licensing statement would allow JVET establishing its own storage of the data. 39 from BVI-DVC are not included.

It was commented that the licensing allows development of standards, but not deployment. This remains an open issue.

It was suggested to investigate retraining LOP4 once with adding new sequences from BVI-AOM, and once replacing BVI-DVC by BVI-AOM (which would mean that not only the last step of training would be necessary). Establish an EE on that.

For the EE on inter prediction, it is recommended to use the BVI-DVC extended by new sequences from BVI-AOM for the comparison of performance against the Vimeo90Ktriplet based training.

CTC would need to be changed in the next meeting, provided that benefit of the new materials is shown.

It was pointed out that contributions using this material for training need to make reference to this contribution.

Expression of thanks to UB and Netflix[JVET-AK0257](https://jvet-experts.org/doc_end_user/current_document.php?id=15246) AHG14: The extension of SADL library [N. Fu, W. Bao, Z. Chen (Wuhan Univ.)] [late]

This contribution presents the extensions in the Small AdHoc Deep Learning (SADL) library from Wuhan University. These extensions include layers such as Resize and GridSample.

Changes are as follows:

* Extend the Resize operator to support Bicubic interpolation mode.
* Extend the GridSample operator to support Bicubic interpolation mode.

(addition to previously available interpolation modes nearest neighbor and bilinear)

Decision(SW): Adopt JVET-AK0257 (merge request already issued).

[JVET-AK0323](https://jvet-experts.org/doc_end_user/current_document.php?id=15312) Neural Network Analysis for Next-Generation Video Compression Standard [T. Hsieh, W.-J. Chien, V. Seregin, M. Karczewicz (Qualcomm)] [late]

Neural networks (NNs) have been a hot topic for many applications, like autonomous driving, object recognition, natural language processing, etc. The question comes to the point: will NNs be realistic for video coding applications? Some complexity analysis is the keyhole that allows us to peek into the feasibility for NNs in future video compression standards. Currently, in the standard community, several video coding tools utilizing NNs have been investigated, among them are intra prediction and in-loop filtering (ILF) which are analyzed in this contribution. There are mainly two ways to deploy NNs for video coding applications, one is using GPU/NPU on the market, and the other is using in-house designed ASIC. The details are illustrated in this contribution.

Presentation deck to be provided.

Hardware architectures: External NN processor (two separate SoCs), or combined on one chip.

Analysis for intra prediction and loop filter

Estimate that chip area for NN computation may be more than 60x compared to conventional video ASIC.in case of intra prediction. 2-chip solution not practical due to latency in sequential block dependency (4x4 worst case, requiring 5 cycles for processing). Sparsity by zero weights may not be simple to implement. Intra is not practical on single chip, either, e.g. due to power consumption.

It was commented that in current NN intra pred. sparsity is designed in a way such that it is grouped, as a rule 16 zero values go together.

It was asked how intra NN would compare to MIP. This was not analysed.

Loop filter is less critical for two-SoC solution, latency not so critical, but high cost of computation still a problem.

For VLOP2, chip area would be similar to entire video codec (multi-standard).

It is to be noted that the analysis is not based on a real implementation study, but uses many assumptions that might be differently resolved by alternative hardware architectures.

More information about “what is possible, or what will be possible within a certain time frame” would be welcome.

## AHG6/AHG12: Enhanced compression beyond VVC capability (74)

### Summary and BoG reports

Contributions in this area were discussed at 1825–2010 on Tuesday 14 Jan. 2025 (chaired by JRO), at 0900–1315 and at 1430–1800 on Wednesday 15 Jan. 2025 (chaired by JRO).

[JVET-AK0024](https://jvet-experts.org/doc_end_user/current_document.php?id=15190) 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)]

1. **List of tests**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Tests** | **Tester** | **Cross-checker** |
| **1 Intra prediction** | | | |
| 1.1a | Extended OBMC for IBC/IntraTMP coded blocks | Y. Kidani  (KDDI) X. Li  (Alibaba) | R. Yu (Qualcomm) D. Kim (ETRI) |
| 1.1b | Test 1.1a + extended OBMC for CCP coded blocks | X. Li  (Alibaba) |  |
| 1.2a | Improvement on non-MPM | G. Wang  (vivo) | J. Fu (PKU) |
| 1.2b | Most dominant intra prediction mode | M. Hong  (LGE) | M. Coban  (Qualcomm) |
| 1.2c | Test 1.2b + excluding improbable intra modes | M. Hong  (LGE) | M. Coban  (Qualcomm) |
| 1.2d | Test 1.2a + Test 1.2c | G. Wang  (vivo),  M. Hong  (LGE) |  |
| 1.3 | IntraTMP merge candidates enrichment | M. Radosavljević (InterDigital) | D. Ruiz Coll,  J.-K. Lee (Ofinno) |
| 1.4 | IntraTMP signalling improvement | F. Le Léannec  (InterDigital) | withdrawn |
| 1.5 | Test 1.3 + Test 1.4 | F. Le Léannec  (InterDigital) | withdrawn |
| 1.6 | Extended IntraTMP merge candidate list with an improved fusion mode | D. Ruiz Coll  (Ofinno) | [W. Lim](mailto:woong.lim@etri.re.kr),  [S.-C. Lim (ETRI)](mailto:sclim@etri.re.kr) |
| 1.7 | Block vector guided DIMD | L. Zhang  (OPPO) | N. Yan  (Kwai) |
| 1.8 | Non-adjacent DIMD fusion modes for TMRL intra mode candidate list construction | V. Rufitskiy,  A. Filippov  (TCL) | S. Blasi (Nokia) |
| 1.9 | Weighted OBIC | P. Andrivon  (Ofinno) | F.Le Léannec  (InterDigital) |
| 1.10a | Bilateral filtering for intra prediction | W. Yin  (Bytedance) | V. Shchukin (Ericsson)  C. Ma  (Kwai) |
| 1.10b | Test 1.10a + 6-tap cubic interpolation filter for TIMD | W. Yin  (Bytedance) | V. Shchukin (Ericsson)  C. Ma  (Kwai) |
| 1.11 | 8-tap interpolation filter for angular intra prediction | T. Dong  (TCL) | S.H. Park (KNU) |
| 1.12 | On MPM with matrix-based position dependent intra prediction | Z. Sun  (OPPO) | X. Li  (Alibaba) |
| 1.13 | Multiple filter taps for EIP | C. Zhou  (vivo) | [Z. Deng (Bytedance)](mailto:zhipin.deng@bytedance.com) |
| 1.14 | Block vector guided EIP | Z. Xie  (OPPO) | Z. Zhang  (Alibaba) |
| 1.15 | Test 1.13 + Test 1.14 | C. Zhou  (vivo)  Z. Xie  (OPPO) | Z. Zhang  (Alibaba) |
| 1.16a | Adaptive intra filter for template prediction | Z. Zhang  (Alibaba)    C. Zhou  (vivo) | Z. Xie  (OPPO) |
| 1.16b | Template-based filter selection for intra prediction | Z. Zhang  (Alibaba)    C. Zhou  (vivo) | Z. Xie  (OPPO) |
| 1.16c | Test 1.16a+Test 1.16b | Z. Zhang  (Alibaba)    C. Zhou  (vivo) | Z. Xie  (OPPO) |
| 1.17 | Test 1.10a + Test 1.11b + Test 1.16b | W. Yin  (Bytedance)  T. Dong  (TCL)    Z. Zhang  (Alibaba)    C. Zhou  (vivo) | Z. Xie  (OPPO) |
| 1.18 | Test 1.7 + Test 1.9 | L. Zhang  (OPPO)  P.Andrivon (Ofinno) | F.Le Léannec  (InterDigital) |
| 1.19 | Test 1.8 + Test1.9 | V.Rufitskiy (TCL)  L. Zhang  (OPPO) | R. G. Youvalari (Xiaomi) |
| **2** **Inter prediction** | | | |
| 2.1a | OBMC modification for GPM blocks | R. Yu  (Qualcomm) | Z. Deng  (Bytedance) |
| 2.1b | OBMC operation order for block and subblock boundaries | R. Yu  (Qualcomm) | Z. Deng  (Bytedance) |
| 2.1c | Test 2.1a + Test 2.1b | R. Yu  (Qualcomm) | Y. Kidani  (KDDI) |
| 2.2a | Additional temporal affine candidates | L. Zhao (Bytedance) | Y. Zhang (Qualcomm) |
| 2.2b | Additional synthetic affine candidates | L. Zhao (Bytedance) | Y. Zhang (Qualcomm) |
| 2.2c | Test 2.2a + Test 2.2b | L. Zhao (Bytedance) | Y. Zhang (Qualcomm) |
| 2.3 | Regression-based GPM with intra and inter prediction | K. Jia  (Alibaba) | Y. Wang  (Bytedance) C. Zhou  (vivo) |
| 2.4 | GPM extension | H. Zhang  (OPPO) | Y. Ahn (LGE) |
| 2.5 | TMVP candidate selection | S. Hong  (Nokia) | R. G. Youvalari (Xiaomi) |
| 2.6 | Template matching padding | N. Neumann  (RWTH) | X. Li  (Google) |
| **3** **Transform and coefficients coding** | | | |
| 3.1a | Multiple kernel set selection for intra LFNST/NSPT with subsampled DIMD for transform set selection | M. Coban  (Qualcomm),  F. Wang  (OPPO),  C. Bonnineau  (Interdigital),  L. Zhao  (Bytedance) | M. Hong  (LGE) |
| 3.1b | Test 3.1a with different types of transform sets | M. Coban  (Qualcomm),  F. Wang  (OPPO),  C. Bonnineau  (Interdigital),  L. Zhao  (Bytedance) | M. Hong  (LGE) |
| 3.2 | LFNST/NSPT set derivation for CCP coded block | H. Huang  (OPPO) | M. Hong  (LGE) |
| 3.3a | Advanced SBT | G. Laroche  (Canon) | F. Le Léannec  (InterDigital) |
| 3.3b | Test 3.3a with position granularity set to 4 instead to 1 | G. Laroche  (Canon) |  |
| 3.3c | Test 3.3a not applied after SBT | G. Laroche  (Canon) |  |
| 3.3d | Test 3.3a with only the position is inferred | G. Laroche  (Canon) |  |
| 3.4 | Last significant coefficient position signalling with secondary prefix | F. Le Léannec  (InterDigital) | P. Nikitin (Qualcomm)  C. Nguyen (Dolby) |
| 3.5a | Coefficient significant and gtX flags coding with probability adaptation | M. Balcilar  (InterDigital) | P. Nikitin (Qualcomm) |
| 3.5b | Coefficient significant and gtX flags coding with increased CABAC engine’s range resolution | M. Balcilar  (InterDigital) | P. Nikitin (Qualcomm) |
| 3.6a | Test 3.4 + Test 3.5a | F. Le Léannec  (InterDigital) | P. Nikitin (Qualcomm) |
| 3.6b | Test 3.4 + Test 3.5b | F. Le Léannec  (InterDigital) | P. Nikitin (Qualcomm) |
| 3.7 | Improved LFNST/NSPT kernel set selection for SGPM | S. Puri  (InterDigital) | I. Zupancic  (Nokia) |
| 3.8a | Implicit MTS extension | M. Karczewicz  (Qualcomm) | M. Salehifar (Bytedance) |
| 3.8b | Test 3.8a reusing existing LUT | M. Karczewicz  (Qualcomm) | M. Salehifar (Bytedance) |
| **4 In-loop filtering** | | | |
| 4.1 | Restrictions on ALF coefficient values | V. Shchukin  (Ericsson),  P. Bordes  (InterDigital) | I. Jumakulyyev (Nokia) |
| 4.2 | Using Laplacian information in ALF | I. Jumakulyyev  (Nokia) | V.Shchukin (Ericsson)  V. Rufitskiy (TCL) |
| 4.3a | Boundary-aware offset refinement for ALF | W. Yin  (Bytedance) | N. Hu  (Qualcomm) |
| 4.3b | Test 4.3a + boundary-aware offset refinement for DBF | W. Yin  (Bytedance) | N. Hu  (Qualcomm) |
| 4.4 | Improvement to ALF scaling factors | N. Hu  (Qualcomm) | withdrawn |
| 4.5a | ALF joint optimization (encoder only) | H. Wang (Qualcomm) | W. Yin  (Bytedance) |
| 4.5b | Encoder only ALF joint optimization tested in VTM | H. Wang  (Qualcomm) | W. Yin  (Bytedance) |
| 4.5c | Increase of the number of ALF APSs | H. Wang  (Qualcomm) | W. Yin  (Bytedance) |
| 4.5d | Test 4.5a + Test 4.5c | H. Wang  (Qualcomm) | W. Yin  (Bytedance) |
| 4.5e | Test 4.5a + reuse of ALF control information | H. Wang  (Qualcomm) | W. Yin  (Bytedance) |
| 4.5f | Test 4.5d + reuse of ALF control information | H. Wang  (Qualcomm) | W. Yin  (Bytedance) |
| 4.6a | Huffman coding for CCALF + simulated annealing for CCALF | K. Takada  (Sharp) | W. Yin (Bytedance) |
| 4.6b | Test 4.1 + Test 4.6a | K. Takada  (Sharp),  V. Shchukin  (Ericsson) | W. Yin (Bytedance) |
| 4.7 | Temporal ALF | L. Xu  (OPPO) | W. Yin  (Bytedance) |
| 4.8.a1 | NN-based ILF with ALF | D. Rusanovskyy  (Qualcomm) | F.Galpin (InterDigital ) |
| 4.8.a2 | NN-based ILF with ALF (Complexity constrained) | D. Rusanovskyy  (Qualcomm) | F.Galpin (InterDigital ) |
| 4.8b | NN-based ILF with VLOP model from NNVC | D. Rusanovskyy  (Qualcomm) | F.Galpin (InterDigital) |
| 4.9a | ALF-CCCM | P. Astola  (Nokia) | R. G. Youvalari (Xiaomi) |
| 4.9b | ALF-CCCM without the use of reference pictures | P. Astola  (Nokia) | R. G. Youvalari (Xiaomi) |

**Description of tests**

***Intra prediction***

**Test 1.1: Extended OBMC for non-inter blocks (**[**JVET-AK0076**](https://jvet-experts.org/doc_end_user/current_document.php?id=15047)**)**

In ECM, OBMC is applied to inter coded blocks. In the tests, OBMC is extended to non-inter coded blocks. The extended OBMC method is applied to sub-blocks on top and left boundaries of the current block. A predictor is generated using the neighbouring block information and blended with the prediction samples of the current block.

In Test 1.1a, OBMC is extended to IBC/IntraTMP coded blocks. If a neighbouring block is IBC/IntraTMP coded, the block vector information of the neighbouring block is used for generating the predictor. If a neighbouring block is intra mode (neither IBC nor IntraTMP) coded, an intra prediction mode is derived by applying DIMD-liked method to the neighbouring reconstructed samples of the current block for generating the predictor.

In Test 1.1b, OBMC is additionally extended to CCP coded blocks. The OBMC is further applied when the current block is coded by a CCP mode and the neighbouring block is coded by a non-CCP mode. An intra prediction mode derived by applying DIMD-liked method is used to generate the predictor for Cb and Cr, respectively.

Test 1.1a: Extended OBMC for IBC/IntraTMP coded blocks

Test 1.1b: Test 1.1a + extended OBMC for CCP coded blocks

**Test 1.2a: Improvement on non-MPM (**[**JVET-AK0086**](https://jvet-experts.org/doc_end_user/current_document.php?id=15057)**)**

In ECM, intra modes are coded using primary MPM list (6 modes), secondary MPM list (22 modes) and the remaining modes (45 modes).

In the test, the non-MPM list are constructed as follows:

1. The first non-planar mode in the MPM list is selected as the starting intra-picture mode.
2. From the starting intra mode, non-MPM list is filled in turns with one forward mode and then one backward with the same step, where the step increases by one in each turn. Mode is skipped if it has already been included in the primary or secondary MPM lists.
3. If a mode satisfies the PDP mode condition, it is swapped with the previous filled non-PDP mode.

Test 1.2a: Improvement on non-MPM

**Test 1.2bc: Intra mode coding based on HoG of neighbouring templates (**[**JVET-AK0059**](https://jvet-experts.org/doc_end_user/current_document.php?id=15030)**)**

In ECM, in DIMD prediction up to 7 intra modes are derived from reconstructed neighbouring samples. These directional predictors are then combined with a non-directional predictor (either planar or block vector-based) using weights calculated from the histogram of gradients (HoG).

Test 1.2b targets intra mode signalling where up to 8 intra prediction modes are derived based on the HoG from neighbouring templates. The HoG computation follows the same process as in DIMD, but template area differs. If the block size is less than 256, a 3-line template area is used, otherwise, template area is extended to 4 lines. Up to 8 intra prediction modes are derived by the HoG based on predefined template areas as follows:

1. Up to 4 intra prediction modes are derived from above and left template area
2. Up to 2 intra prediction modes are derived from left template area
3. Up to 2 intra prediction modes are derived from above template area
4. Neighbouring modes of derived modes (1) ~ (3) are added until up to 8 intra prediction modes are identified.

After constructing up to 8 intra prediction modes, they are ordered by computing their template cost. The intra prediction mode with the minimum template cost is signalled before MPM information at each CU when block size is equal or less than 1024, and to avoid redundancy, this mode is excluded from MPM, secondary MPM, and non-MPM.

In Test 1.2c, on top of Test 1.2b, some improbable intra modes are excluded. After computing the HoG, 20 improbable intra modes with small HoG amplitude are excluded from MPM, SMPM, and non-MPM list. If the signalling using the smallest template cost mode is applied, the HoG which was utilized to derive that mode is reused to exclude the intra prediction modes. If the signalling is not applied due to block size restriction, the HoG of DIMD is used to exclude the intra prediction modes.

When both the signalling and the excluding improbable intra modes are applied, non-MPM mode size reduces from 45 (=67 – 22) to 24 (=67 – 22 – 1(smallest template cost mode) – 20(excluding intra prediction modes)).

Test 1.2b: Most dominant intra prediction mode

Test 1.2c: Test 1.2b + excluding improbable intra modes

**Test 1.3: IntraTMP merge candidates enrichment (**[**JVET-AK0220**](https://jvet-experts.org/doc_end_user/current_document.php?id=15209)**)**

This test evaluates modifications for BV merge list generation, where new BV candidates are added.

The distance for non-adjacent spatial merge candidates is increased from 4 to 5, where up to 5 rows/columns of spatial neighboring blocks (in terms of the CU height/width) of the current CU can be used in merge BV candidate construction. This is achieved for camera captured content only. The spatial neighbouring blocks of the current block are identified using 13 adjacent blocks (as in OBIC). Total number of BV candidates is eventually limited to 36.

In ECM, merge BVs are collected from local and non-local PUs coded in IBC or Intra TMP mode. In this test BV merge list is generated from all the spatial merge candidates that contain BVs (BV-based prediction is currently used in several modes and submodes in ECM software).

When constructing merge BV candidates, BVs from HMVP list are used. HMVP candidates are added after spatial (local and non-local) candidates and after ARBVP candidate construction. Up to 25 candidates from HMVP list are added to the merge candidates.

The complexity due to newly introduced above aspects is reduced by following elements:

* Block vector refinement employs a sparse search withing refinement area, followed by a last refinement step around best found positions.
* A distance-based criteria is used to discard some BV candidates close to already existing candidates in merge list under construction.
* Top-left template is no more considered in IntraTMP template cost computations.

**Test 1.6: Extended IntraTMP merge candidate list with an improved fusion mode (**[**JVET-AK0161**](https://jvet-experts.org/doc_end_user/current_document.php?id=15132)**)**

In ECM, in addition to the regular IntraTMP mode, a TMP fusion mode can blend up to 5 candidate reference blocks according to weight parameters. Two different fusion models are available to derive the weight parameters of the fusion mode, which are based on the template SAD cost and the Gaussian solver methods. An intra\_tmp\_fusion\_weight\_type flag is signalled to the decoder, indicating the TMP fusion model used to derivate the weight parameters in the blending process.

For the neighbour blocks encoded as regular IntraTMP (non-fusion) mode, if the IntraTMP index is not the index 0, the first BV candidate in the IntraTMP list is also included in the IntraTMP merge and AR-BVP list.

In the test, the adjacent and non-adjacent neighbour blocks are checked in the IntraTMP merge and AR-BVP list construction. Currently, if those blocks were encoded using the IntraTMP fusion mode, only the first BV fusion candidate is selected to be added to these lists. The reminder fusion BV candidates, up to 5 adjacent BVs in the IntraTMP refined list, are not checked. In the test, the entire set of fusion BV candidates when a neighbouring block is coded in IntraTMP fusion mode are evaluated.

This test also modifies IntraTMP fusion model signalling. Instead of signalling the fusion model with a flag, the fusion model for IntraTMP groups 1 and 2 is derived on the decoder side based on the template matching cost of the fused templates, which are calculated using the respective weight parameters.

To alleviate the complexity impact, sparse searching in regions R4 and R5 of IntraTMP search range is skipped if the top-left position of those regions in not valid, and if the bottom-right position of those regions is valid, the whole validation process is bypassed for such regions. Merge candidates with the template cost above a threshold (set equal to the template cost of the last candidate in the IntraTMP sparse list) are skipped and not considered for the IntraTMP and the AR-BVP list construction.

Test 1.6: Extended IntraTMP merge candidate list with an improved fusion mode

**Test 1.7: Block vector guided DIMD (**[**JVET-AK0060**](https://jvet-experts.org/doc_end_user/current_document.php?id=15031)**)**

In this test, neighbour reconstructed samples pointed at by BVs, instead of the spatial adjacent samples, are used to build HoG for DIMD as shown in the next figure. Up to 5 block vectors are obtained by the same searching process as in IntraTMP. With each block vector, a reference area in the current picture is determined and its samples are used to derive the intra modes and corresponding amplitudes. To reduce complexity, overlapping reference areas are clustered.

A diagram of a graph and direction

Description automatically generated with medium confidence

Two intra modes with the highest amplitudes are selected from the HoG and the prediction of the current block is the blending of those 2 predictors and a non-directional predictor. The non-directional predictor is the blending of up to 5 BV-based predictors obtained using the block vectors.

The BV guided DIMD mode is signalled as a sub-mode of DIMD after OBIC. At encoder, to reduce complexity, the mode is checked by a full RD only for block sizes larger than 128 with width and height greater than 4, otherwise, it is compared with other intra modes by the SATD.

Test 1.7: Block vector guided DIMD

**Test 1.8: Non-adjacent DIMD for TMRL (**[**JVET-AK0104**](https://jvet-experts.org/doc_end_user/current_document.php?id=15075)**)**

In the test, DIMD is applied to non-adjacent reference lines, where HoG is calculated in the area determined by a reference suing 2x2 edge operator. Intra prediction modes obtained using the HoG and fusion candidates are added to the list of intra prediction modes for TMRL lines. DC mode is not included into TMRL list.

Each candidate added to a TMRL candidate list consists of a pair of

* an intra-prediction mode (or a set of modes if DIMD fusion is used in the case of non-adjacent DIMD applied to TMRL), and
* non-adjacent reference lines (reference line index).

To compare template matching costs, which is estimated by using the HAD metric in TMRL, each non-adjacent reference line used in TMRL is projected onto the adjacent reference line of a current block according to the angular mode direction of a TMRL. If the fusion process is enabled for a candidate from the TMRL list, a projection of each mode is obtained and, then, these projections are combined according to the weights assigned to each angular mode and reference line belonging to this set.

**Test 1.9: Weighted OBIC (**[**JVET-AK0056**](https://jvet-experts.org/doc_end_user/current_document.php?id=15027)**)**

In ECM, OBIC derived intra prediction modes are based on the sample-wise occurrences of the neighbouring intra prediction modes. Similarly to DIMD, a histogram of occurrence (HoC) is built and up to five most represented angular modes are selected. For neighbouring blocks using more than one intra prediction mode to generate a predictor (e.g. TIMD, DIMD…), each corresponding intra mode is used with the same level of contribution in the HoC construction. The selected modes are blended with planar mode or a block-vector based mode to generate the OBIC predictor.

In the test, instead of using the same level in HoC construction, the occurrence value is weighted for the case when multiple intra prediction modes are used from a neighbour block, weights (*k*i) and the location dependency state (used on the fusion process) are reused to scale the relevant mode occurrence as follows.

HoC(Mi) += *ki* . *w.* *h*

In case a HoC entry is built from an intra mode with different location dependency states, the most represented location dependency state is selected.

Additionally, each contributor to the HoC is weighted by a scaling factor depending on the distance between the current block and the considered neighbouring block where the distance weight is selected from {4, 2, 1, 0}.

**Test 1.18: Combination of Test 1.7 and Test 1.9 (**[**JVET-AK0283**](https://jvet-experts.org/doc_end_user/current_document.php?id=15272)**)**

Test 1.18: Test 1.7 + Test 1.9

**Test 1.19: Combination of Test 1.7 and Test 1.8 (**[**JVET-AK0284**](https://jvet-experts.org/doc_end_user/current_document.php?id=15273)**)**

Test 1.19: Test 1.7 + Test 1.8

**Test 1.10: Bilateral filtering for intra prediction (**[**JVET-AK0118**](https://jvet-experts.org/doc_end_user/current_document.php?id=15089)**)**

In Test 1.10a, intra prediction samples are further filtered by bilateral filter to reduce the noise, bilateral filter is kept the same as BIF used in the loop-filtering stage. The available above and left neighbouring reconstructed samples are used in the bilateral filtering process for the current block.

In Test 1.10b, on top of Test 1.10a, a 6-tap cubic interpolation filter is always applied for the blocks coded with TIMD mode.

Test 1.10a: Bilateral filtering for intra prediction

Test 1.10b: Test 1.10a + 6-tap cubic interpolation filter for TIMD

**Test 1.11: 8-tap interpolation filter for angular intra prediction (**[**JVET-AK0087**](https://jvet-experts.org/doc_end_user/current_document.php?id=15058)**)**

In the test, 8-taps interpolation filter is used for intra prediction reference samples if the following conditions are met:

1. Intra mode prediction angle has sufficient angular distance from horizontal or vertical modes. The threshold to determine that is dependent on the block size. The threshold is different for normal (66 modes) and high (131 modes in TIMD) angular mode precision.
2. Block size is smaller than a given threshold, the threshold is different for normal (66 modes) and high (131 modes in TIMD) angular mode precision.

Otherwise, the existing 6-taps interpolation filter is used.

Test 1.11a: 8-tap filter with the norm 8192.

Test 1.11b: 8-tap filter with the norm 256 (reduced precision).

**Test 1.16: On intra interpolation filter (**[**JVET-AK0102**](https://jvet-experts.org/doc_end_user/current_document.php?id=15073)**)**

In the test, intra interpolation filter is selected based on TM cost as follows.

In Test 1.16a, instead of always choosing 4-tap DCTIF interpolation filter for template prediction, the MDIS condition is also applied in intra filter selection for the template prediction. The 6-tap DCTIF is always applied for TMRL mode.

In Test 1.16b, the intra interpolation filter is determined by using TM cost of the current coding block. That is, 6-tap DCTIF, 4- and 6-tap smoothing interpolation filters are used to perform intra prediction on the template area, and the one results in the minimum TM cost is selected as the intra interpolation filter for the prediction of the current block. If the 6/4-tap smoothing interpolation filter is selected, [1, 2, 1]/4 reference filter is disabled to avoid applying smoothing filter twice for intra prediction.

Test 1.16a: Adaptive intra filter for template prediction

Test 1.16b: Template-based filter selection for intra prediction

Test 1.16c: Test 1.16a + Test 1.16b

**Test 1.17: Combination of Test 1.10a, Test 1.11b, and Test 1.16b (**[**JVET-AK0148**](https://jvet-experts.org/doc_end_user/current_document.php?id=15119)**)**

Test 1.17: Test 1.10a + Test 1.11b + Test 1.16b

**Test 1.12: On MPM with matrix-based position dependent intra prediction (**[**JVET-AK0061**](https://jvet-experts.org/doc_end_user/current_document.php?id=15032)**)**

In ECM, matrix-based position-dependent intra-prediction (PDP) replaces modes 0, 1, (2+2k) for blocks up to 16x16, and 0, 1, (2+4k) for other blocks.

In the test, some of the non-PDP intra prediction modes (IPM) in the MPM/Non-MPM list are converted to PDP modes as follows:

If is not added to the MPM/Non-MPM list, will be further considered for MPM/Non-MPM list, where offset is equal to 2 for blocks with width and heights less or equal to 16 or 4, otherwise.

Additionally, PDP prediction, instead of the conventional angular prediction, is used to calculate the template cost for MPM mode sorting for PDP modes.

Test 1.12a: Converting non-PDP to PDP in the MPM and non-MPM lists

Test 1.12b: Using PDP prediction for MPM sorting

Test 1.12c: Test 1.12a + Test 1.12b

**Test 1.13: Multiple filter taps for EIP (**[**JVET-AK0084**](https://jvet-experts.org/doc_end_user/current_document.php?id=15055)**)**

In ECM, EIP filter length is 15. In the test, a 9-tap filter is introduced and an index indicating which filter length is used is signalled. For EIP mode using 9-tap filter, the prediction is calculated as follows:

9-tap filter is applied to multi-model EIP for blocks with area larger than or equal to 64, and smaller than or equal 256.

Test 1.13: Multiple filter lengths for EIP

**Test 1.14: Block vector guided EIP (**[**JVET-AK0062**](https://jvet-experts.org/doc_end_user/current_document.php?id=15033)**)**

In the test, BV is used to determine the reference area for calculating EIP filter parameters instead of directly using the adjacent spatial reference area and the mode is signalled with a flag as a sub-mode of EIP.

BV is derived from the sparse searching process of the IntraTMP in the current ECM.

Only square EIP filter shape is allowed in this mode.

Test 1.14: Block vector guided EIP

**Test 1.15: Combination of Test 1.13 and Test 1.14 (**[**JVET-AK0149**](https://jvet-experts.org/doc_end_user/current_document.php?id=15120)**)**

Test 1.15: Test 1.13 + Test 1.14

1.1a is straightforward and matching according to cross-checkers. Tradeoff with encoding time still in reasonable range. 1.1b is not attractive.

Decision: Adopt JVET-AK0076 test 1.1a

Among the 1.2x tests, 1.2c has most attractive tradeoff. Implementation is straightforward and results confirmed by cross-checker.

Decision: Adopt JVET-AK0059 test 1.2c

1.3 and 1.6 are having rather small gain but increase complexity by additional processing at encoder and decoder. No action on these.

1.7 and 1.8 provide some compression gain, but also require significantly more encoder/decoder processing. Run times reported are not in the range of reasonable tradeoff, and reported to be even higher by cross-checkers. Further study (EE) recommended as some experts found the ideas interesting, but the proposals would not be taken into consideration without significant decrease of complexity/runtime, without significantly affecting the compression benefit.

1.9 was supported by several experts, as it is simple to implement and has small gain without increasing run time.

Decision: Adopt JVET-AK0056 test 1.9.

For 1.10a, it was commented by one expert that some misalignment with requests from previous meeting notes (processing order with PDPC etc.) was found in the new implementation, and conditions under which it is applied (certain modes, certain block shapes) appear a bit arbitrary to him. Generally, the tradeoff would be still acceptable, even though a similar concept had been rejected during VVC development, as it was assumed that the latency in the intra prediction loop would be critical. This might not be so relevant in an exploration, however.

Several experts supported the adoption.

Decision: Adopt JVET-AK0118 test 1.10a

Test 1.11b shows a very good tradeoff (in comparison with some other proposals), providing compression gain without significant complexity increase

Decision: Adopt JVET-AK0087 test 1.11b

Test 1.16b standalone does not have enough compression benefit justifying the additional complexity. The combination 1.17 shows would not change this assessment.

Several independent experts supported adoption of test 1.12c which has attractive tradeoff. Even though the element of 1.12b has some complexity increase, its standalone performance would still in the margin of acceptable tradeoff.

In terms of this exploration, the method is assessed to be acceptable.

Decision: Adopt JVET-AK0061 test 1.12c.

Test 1.13-1.15: The combination might have acceptable tradeoff after further encoder optimization (1.15\*). It was asked if a similar encoder optimization (skipping some checks) could also be done to EIP without modification. This is probably the case, but no results are available for such a configuration. Several experts found the ideas interesting and supported adoption.

It was commented that 1.13 has only marginal gain while increasing encoder time, and also RA results are still missing. Overall, it does not seem very attractive. It was argued by proponents that the main motivation is reduction of decoding time (99.1%).

It was concluded to continue the study of 1.14 in EE, providing results with the encoder optimization for the method standalone, and also compare that against a version with same encoder optimization without changing EIP.

Partial results in RA were presented, indicating luma gain of -0.01/-0.05/-0.04% for classes B/C/D. It was concluded to continue EE study on 1.13, also study the combination with 1.15\* again in ECM16 (to be compared against the version with enc. optimization without changing EIP.

***Inter prediction***

**Test 2.1: OBMC modifications (**[**JVET-AK0212**](https://jvet-experts.org/doc_end_user/current_document.php?id=15201)**)**

In ECM, when two GPM partitions are coded with non-affine inter, the GPM blending is carried out first, and then OBMC is applied on top of blended GPM samples.

In Test 2.1a, OBMC is applied separately on top of each GPM partition, and then blending is applied on top of OBMC modified samples of the two partitions. If the corresponding partition contains subblock motion, OBMC is applied on those subblock boundaries as well.

In Test 2.1b, OBMC order is swapped, i.e. it is applied on subblock boundaries (if subblocks are present within the CU) first, and then OBMC is applied on block boundaries.

Test 2.1a: OBMC modification for GPM blocks

Test 2.1b: OBMC operation order for block and subblock boundaries

Test 2.1c: Test 2.1a + Test 2.1b

**Test 2.2: Enhanced derivation of affine merge candidates (**[**JVET-AK0095**](https://jvet-experts.org/doc_end_user/current_document.php?id=15066)**)**

In ECM, temporal affine candidates are derived to enable the inheritance of affine motion from collocated reference pictures, where CPMVs of the current block are derived by scaling the collocated CPMVs to the reference picture with the scaling factor closest to 1. To generate bi-directional affine candidates for a non-low-delay B picture, the collocated CPMVs referring to only one reference list is scaled to the corresponding reference pictures in the two reference lists of the current frame.

In Test 2.2a, additional affine temporal candidates are derived by scaling the collocated CPMVs to the reference picture with index 0, and the collocated CPMVs referring to both lists can be utilized to obtain the temporal affine candidates. In addition, when ARMC is applied, all temporal affine candidates are sorted based on template costs, and candidates with the minimum costs are inserted into the subblock merge list.

In Test 2.2b, after the subblock merge list is constructed, additional synthetic affine candidates are introduced by combining the motion of two existing affine candidates, where the CPMVs and reference index of a candidate in List 0 are coupled with those of another candidate in List 1 to generate a new affine candidate. At most 5 synthetic candidates can be appended to the end of the current subblock merge list. The final subblock merge list is reordered when ARMC is applied.

The maximum allowed candidate number in the final subblock merge list stays unchanged and the two aspects are only applied to non-low-delay B pictures.

Test 2.2a: Additional temporal affine candidates

Test 2.2b: Additional synthetic affine candidates

Test 2.2c: Test 2.2a + Test 2.2b

**Test 2.3: Regression-based GPM with intra and inter prediction (**[**JVET-AK0101**](https://jvet-experts.org/doc_end_user/current_document.php?id=15072)**)**

In ECM, GPM partitions can be predicted using intra or inter prediction. However, in regression-based GPM only inter prediction can be used.

In the test, regression-based GPM is extended that two partitions can be prediction using inter and intra prediction. A flag is signalled to indicate the intra-inter prediction usage.

Intra prediction mode candidates are the first 6 modes in the MPM list and the MV candidates for the inter prediction are from regular GPM MV candidate list. Then, an intra-inter pair list is constructed by combining these intra prediction mode and MV candidates.

For each pair, two integer blending matrices (*W0* and *W1*) are derived. The matrix derivation method is the same as that of the current inter-inter regression-based GPM. Then the intra-inter pair list is re-ordered with the template cost, and the index of the selected intra-inter pair is signalled. The final predicted samples are generated by weighing inter and intra predicted samples with *W0* and *W1* matrices.

Besides, an SPS flag is signalled to enable or disable this mode. This flag is set to disable this mode for large QP (not smaller than 37) and for low delay configuration.

Test 2.3: Regression-based GPM with intra and inter prediction

**Test 2.4: GPM extension (**[**JVET-AK0063**](https://jvet-experts.org/doc_end_user/current_document.php?id=15034)**)**

In the test, several modifications are performed for GPM modes.

Firstly, CU size restriction of GPM is modified as “cbWidth >= 8 && cbHeight >= 8 && cbWidth <= 128 && cbHeight <= 128”. The restriction of width /height ratio “cbHeight < 8 \* cbWidth && cbWidth < 8 \* cbHeight” is removed. With this modification, CUs of size cbWidth \* cbHeight = 2m \* 2n (m, n∈2…7) can apply the geometric partition mode.

Secondly, a condition on the area of the block is introduced to determine whether GPM-MMVD is allowed. If cbWidth \* cbHeight >= 2048, GPM-MMVD mode is not allowed for the current block and the threshold of regression-based GPM is changed from 1.25 to 1.05. The method of shape-based GPM split mode is also applied for the additional blocks.

Test 2.4: GPM extension

**Test 2.5: TMVP candidate selection (**[**JVET-AK0185**](https://jvet-experts.org/doc_end_user/current_document.php?id=15156)**)**

In ECM, TMVP candidates are derived from either bottom right or center position relative to the current block.

In the test, if both positions contain valid motion information and the candidate list is not complete, the motion vectors from both positions are included.

Test 2.5: TMVP candidate selection

**Test 2.6: Template matching padding (**[**JVET-AK0085**](https://jvet-experts.org/doc_end_user/current_document.php?id=15056)**)**

In the test, the outside area of a reference picture is derived by template matching. First, a template is defined adjacent to the target block. Then a search is performed and the predictions associated with the best BVs are averaged to form the prediction for the target block. The lowest TM cost (SAD) multiplied by 1.5 is used as a threshold for BVs selection, a maximum number of candidates is set to 4.

The search starts with distance of 3 and is gradually increases until one of the candidates reaches SAD cost value of 128 or the maximum distance of 12 is reached.

The target block size is set to 12x1 in the direction parallel and perpendicular to the picture boundary respectively. The template size is set to 16x3 in those dimensions, and it is centered adjacent to the target block.

The method is applied to RAPs and low-delay pictures. For low-delay pictures it replaces motion compensated padding.

Test 2.6: Template matching padding



2.1, even though having complexity impact by needing additional processing (which explains the increased encoding and decoding time) is supported by several experts including cross-checkers. The combination 2.1c, as well as standalone results show that the benefit of 2.1b is negligible for RA. Results for LB (where gain is only appearing with the combination) lack explanation.

Decision: Adopt JVET-AK0212 2.1a (only for RA in CTC)

Further study of 2.1a-c on LB in EE, where changing the order of operations in 2.1b may have some benefit, try to produce more explainable results, and reduce the encoding run time.

For 2.2.x, cross-checker confirms results, as well as appropriateness of implementation and description, and supports 2.2c. Results are also consistent over various classes.

Overall, gains in the combination are additive, and tradeoff gain/runtime is acceptable.

Decision: Adopt JVET-AK0095 test 2.2c

2.3 extends regression based GPM to intra prediction, allowing all combinations of inter and intra with regression. Results are confirmed by cross-checkers. It was asked why this change is not applied for LB? According to proponents, tradeoff is not good in that case.

Generally agreed that this is useful and tradeoff for RA is acceptable.

Decision: Adopt JVET-AK0101 test 2.3 (CTC only for RA)

Partial results for LB were presented in a follow-up review on Monday 20 Jan. This increases runtime and loses compression gain – no action onLB CTC is deemed useful.

For 2.4, some restrictions are released to allow usage of GPM for more block sizes and block shapes. Results and run time confirmed by cross-checkers.

Tradeoff is better than in the proposal of last meeting, but still not acceptable. Further study recommended if it can be further improved (outside of EE).

2.5 has some benefit only in LB (with acceptable tradeoff), but no loss or increase of runtime in RA. Results are homogeneous over classes (where in the classes some sequences also may have loss). Results confirmed by cross-checkers, which also express that the modification is straightforward.

Decision: Adopt JVET-AK0185 test 2.5 (in CTC for both RA and LB)

Test 2.6 is only applied for LB pictures, therefore gain in RA is marginal (but no harm). Cross-checker confirms results and supports proposal, as tradeoff for LB is attractive.

It was asked what the implementation complexity would be (in comparison with the motion compensated padding that it would replace).

According to proponents and cross-checker, due to some more optimization of TM (early stopping criterion) it would be less complex than MC padding. If that would be the case, it would be a candidate for adoption.

It was also commented that the runtime increase in case of more subpicture boundaries (rather than picture boundaries as in CTC) would be higher.

The results in smaller resolution classes indicate more gain in BD rate (0.28% in class D) and higher reduction of decoder run time (0.5%) which indicates its benefit over MC padding.

In RA pictures, though the complexity seems to be lower than MV padding, the method would generate higher BD rate; therefore it is not proposed to be used there.

Several independent experts supported the proposal, and no objection was raised.

Decision: Adopt JVET-AK0085 test 2.6

It was suggested to perform an analysis about complexity (number of different types of operations per boundary sample) in comparison to MC padding in the context of the tool assessment of AHG7.

***Transform and coefficient coding***

**Test 3.1: Multiple transform set selection for intra LFNST/NSPT with subsampled DIMD for transform set selection (**[**JVET-AK0217**](https://jvet-experts.org/doc_end_user/current_document.php?id=15206)**)**

In ECM, multiple transform set selection (MTSS) for intra LFNST/NSPT is applied to IntraNN and inter GPM modes.

In the test, MTSS is extended to DIMD, OBIC, TIMD, SGPM, MIP, EIP, and IntraTMP modes. Secondary intra mode for selection of LFNST/NSPT transform kernels is derived from the second highest HoG of DIMD with subsampling applied to reduce the complexity in HoG derivation.

The derivation methods for transform sets are summarized below for Test 3.1a:

|  |  |  |
| --- | --- | --- |
|  | IPM to derive the 1st transform set | IPM to derive the 2nd transform set |
| DIMD, OBIC, TIMD | The 1st IPM used in the mode | The 2nd IPM used in the mode |
| SGPM, MIP, EIP, IntraTMP | The 1st IPM derived by HoG of prediction | The 2nd IPM derived by HoG of prediction |

In Test 3.1b, an alternative scheme of selecting 2nd transform set is introduced where for NSPT blocks, if the difference between the 1st IPM and 2nd IPM is less than equal to 2, the 1st IPM is used to select the kernels from set0 (regular intra mode kernels) instead of 2nd IPM as in Test 3.1a. Multiple NSPT kernel sets assigned to different type of prediction modes (set0: regular intra modes, set1: TIMD/DIMD/OBIC/SGPM/MIP/EIP, set2: IntraTMP/inter) was adopted to ECM-15.0.

MTSS can be applied to CUs with number of samples equal or larger than 128. For CUs less than 256 samples, only the first secondary set candidate can be used, for larger CUs all 3 candidates can be used.

Test 3.1a: MTSS for intra LFNST/NSPT with subsampled DIMD for transform set selection

Test 3.1b: Test 3.1a with different types of transform sets

Test 3.1a\*: Test 3.1a with encoder optimization

Test 3.1b\*: Test 3.1b with encoder optimization

**Test 3.2: LFNST/NSPT set derivation for CCP coded block (**[**JVET-AK0064**](https://jvet-experts.org/doc_end_user/current_document.php?id=15035)**)**

In ECM, LFNST/NSPT transform set for CCP coded chroma blocks is derived by a collocated luma intra prediction mode.

In the test, DIMD is used to derive intra prediction mode of the current block based on the CCP predicted samples. A mode with the largest histogram amplitude is used to determine the LFNST/NSPT transform set. When JCCR is applied to the current block, according to the JCCR mode, either individual or both chroma component predictors are used in the HoG derivation.

Test 3.2: LFNST/NSPT set derivation for CCP coded blocks

**Test 3.3: Advanced SBT with direction and position inference (**[**JVET-AK0131**](https://jvet-experts.org/doc_end_user/current_document.php?id=15102)**)**

In this test, a position and direction of rectangular shape SBT modes is derived at decode side without signalling. For each subblock size and direction, the best position of the transform part (grey area show in the next figure) is evaluated in the searching process by shifting SBT partition with a granularity of 1 or 4 samples within a block. Then, based on the best position, the direction (horizontal or vertical) for each subdivision is determined.

The position selected is the position of the subblock which maximizes the sum of gradients obtained on the corresponding inter predictor block, the direction is selected as the one that maximizes the sum of gradients of each best position.

A screenshot of a computer

Description automatically generated

The implicit SBT mode usage is indicated by a flag at TU level, if implicit mode is enabled only SBT rectangular shape choice is signalled without the position and direction. The implicit SBT mode is also allowed to be used for SBT TU with non-zero residual to obtain further smaller subblock residual.

The implicit SBT mode is applied only to luma component and is not allowed for intra slices and inter CIIP and GPM modes, it is also disabled for LFNST, MTS, and for SCC by SPS flag.

Test 3.3a: Advanced SBT

Test 3.3b: Test 3.3a with position granularity set to 4 instead of 1

Test 3.3c: Test 3.3a not applied after SBT

Test 3.3d: Test 3.3a with only the position is inferred

**Test 3.4: Last significant coefficient position signalling with secondary prefix (**[**JVET-AK0097**](https://jvet-experts.org/doc_end_user/current_document.php?id=15068)**)**

In ECM, last position (x, y) is signalled using context coded prefix and bypass coded suffix. In the test, a context coded secondary prefix flag is added to indicate whether the suffix is equal to the max value, and if it is not then suffix is signalled.

Prefix and suffix signalling is kept unchanged.

This method is applied when the transform block size is higher or equal to some value (e.g. 8) and prefix value has maximum allowed value in the considered transform block.

In Test 3.4a (first trade-off), a single CABAC context is used to code the secondary prefix, respectively for X and Y coordinates.

In Test 3.4b (second trade-off), the secondary prefix is coded with a separate context for each block size, and the first bin of the suffix part is context coded with a separate context by block size.

Low QPs (27, 22, 17, 12) results are also provided in the contribution.

Test 3.4a: Secondary prefix for last position with shared contexts

Test 3.4b: Secondary prefix for last position with separate contexts

**Test 3.5: High Resolution Arithmetic Engine (**[**JVET-AK0105**](https://jvet-experts.org/doc_end_user/current_document.php?id=15076)**)**

In ECM, CABAC engine uses 15 bits to represent probability. It is argued that the accuracy is not enough to represent probability of significant and gtX bins.

In Test 3.5a, the probability is assumed to be 1, when the probability of a given context is above some threshold (e.g. 32720) and is assumed to be 0, when the probability of a given context is lower than some certain threshold (e.g. 48). In those cases, bins are not encoded assuming they are equal to 1 or 0, but context probability is updated. At encoder side, it is ensured that this assumption always holds, and incorrect cases are invalidated and not selected at encoder.

In Test 3.5b, resolution of the CABAC engine is increased, where instead of 9-bit 10-bit CABAC range variable is used.

Test 3.5a: Significant coefficient and gtX flags coding with probability adaptation

Test 3.5b: Significant coefficient and gtX flags coding with increased CABAC range resolution

**Test 3.6: Combinations of Tests 3.4 and 3.5 (**[**JVET-AK0098**](https://jvet-experts.org/doc_end_user/current_document.php?id=15069)**)**

Test 3.6a: Test 3.4a + Test 3.5a

Test 3.6b: Test 3.4a + Test 3.5b

Test 3.6c: Test 3.4b + Test 3.5a

Test 3.6d: Test 3.4b + Test 3.5b

**Test 3.7: Improved LFNST/NSPT kernel set selection for SGPM (**[**JVET-AK0222**](https://jvet-experts.org/doc_end_user/current_document.php?id=15211)**)**

In the test, instead of using split direction to determine LFNST transform kernel set for SGPM, two intra modes are derived by DIMD process applied to SGPM prediction. These two derived intra modes are compared to SGPM intra modes, and if the derived mode is close to one of SGPM intra modes with some threshold, then the derived mode is used to determine LFNST/NSPT kernel, otherwise the split direction mode is used as in ECM.

Test 3.7: Improved LFNST/NSPT kernel set selection for SGPM

**Test 3.8: Implicit MTS extension (**[**JVET-AK0187**](https://jvet-experts.org/doc_end_user/current_document.php?id=15158)**)**

In ECM, implicit MTS infers a separable transform pair of an intra coded block from a LUT utilizing the TU size and the prediction mode index. However, the current method restricts the usage of a LUT to regular (or PDP) intra modes while non-regular intra modes continue using DCT-II.

In the test, additional LUTs are introduced for each, DIMD, TIMD, MIP, TMP, SGPM, EIP, and IntraNN coded blocks. A primary and secondary prediction modes are given for each intra mode type, denoted as ipm1 and ipm2, respectively. They are set as the first and second highest HoG of the prediction block for SGPM, MIP, EIP and TMP, respectively. For DIMD, OBIC, TIMD, and PDP they are set as the 1st and 2nd IPM of each mode or derived through HoG using neighbouring reconstructed pixels if 1stand 2nd IPM are equal or 2nd is unavailable.

The optimal transform pair is selected from the LUT depending on the TU size, the primary prediction mode, and the prediction mode difference defined as diff = abs(ipm1 – ipm2).

In Test 3.8a, the current LUT utilized for regular intra (or PDP) is replaced by a new LUT supporting the mode difference dependency and extended to other intra modes.

In Test 3.8b, the method of Test 3.8a is employed where the current LUT utilized for regular intra (or PDP) is kept unchanged.



3.1b more attractive than 3.1a. With encoder optimization in 3.1b\* (which in its major part giving speedup could not be applied to current ECM), the tradeoff would be acceptable both for AI and RA. Results and method confirmed and supported by cross-checkers.

It is noted that the word “intra” in the title of 3.1 is expressing that the proposal only modifies the intra case, whereas MTSS is applied to both inter and intra.

Several independent experts supported the proposal.

See further notes under JVET-AK0329 shows that there is no interaction problem between 3.1b\* and 3.7.

Test 3.2 has some gain mainly for chroma without impacting runtime. It is only relevant for CCP coded blocks. RA results cannot be expected to improve much beyond the numbers given above, as only one QP point is missing.

Cross-checkers confirm matching of results they obtained so far, and also support the proposal’s adoption.

Method straightforward, good tradeoff for small gain. Support expressed also by other non-proponents.

Decision: Adopt JVET-AK0064 test 3.2.

Test 3.3x implicitly determines best position of SBT in a rectangular block by gradient analysis from the reference block.

It was asked if the parsing of transform coefficients and inverse transform can still be conducted independently, as the e.g. the orientation is not known – this cannot be done independent from the prediction. Similar dependencies may however also exist in other tools in ECM.

Discussion was performed about the conflict between granularity of 1 and the deblocking filter granularity of 4x4. However, in the current implementation, no deblocking is applied at the adaptive SBT boundary, regardless if the granularity is 1 or 4. This is OK in terms of consistent decoding process.

Changing the granularity of deblocking is not recommended, anyway.

Tradeoff of none of the variants is attractive from current results, though better already than in the proposal of last meeting. Further study for reducing encoding time in EE.

3.4x and 3.5x (combined in 3.6x) target further optimizations of coefficient coding

3.4a 1 additional context coded bins, 3.4b 10 additional context coded bins, 3.5x no additional context (3.5 is only targeting higher precision in CABAC).

Most gain in 3.6 comes from 3.4b. The additional gain from 3.5 is relatively low, which seems to indicate that the current precision of ECM CABAC is sufficient.

Decision: Adopt JVET-AK0097 test 3.4b.

3.7 requires additional DIMD for kernel selection (which is however also used in other places in ECM) and does not increase worst case. Encoder/decoder run times are constant. Small gain consistent over all classes in AI and RA; in LB slightly higher chroma gain.

Cross-checkers confirm results and support the proposal, as well as another independent expert.

See further notes under JVET-AK0329 shows that there is no interaction problem between 3.1b\* and 3.7.

Test 3.8x improves implicit MTS (non-CTC, would generate losses). Compared to current implicit MTS, relevant gain is achieved with reasonable tradeoff). Test 3.8b uses the same lookup table as current implicit MTS, whereas test 3.8a uses a new one with larger memory requirement. Both are almost equal in performance. Also some more operations necessary in 3.8a. 3.8b is even slightly better in LB.

Decision: Adopt JVET-AK0187 test 3.8b.

***In-loop filtering***

**Test 4.1: Restrictions on ALF coefficient values (**[**JVET-AK0123**](https://jvet-experts.org/doc_end_user/current_document.php?id=15094)**)**

In the test, ALF coefficients are limited to power of 2, e.g. m\*2b, where m and b are integers in the ranges of and , the range of coefficients is , where BALF is the precision factor with the corresponding mantissa precision BM.

For each configuration, the range of allowed mantissa precisions is hardcoded. For the chroma components, mantissa precision is utilized, i.e., only pure powers of two are allowed, but for the luma component, the mantissa precision is configured as . For luma, the encoder signals one bit in the APS to indicate which mantissa precision is used, or . The choice of value for luma is based on RD-cost.

Coefficients are signaled using Huffman coding and 18 Huffman codes (16 for luma and 2 for chroma) are utilized. The frequencies for Huffman code are collected by compression of BVI-DVC sequences using ECM-based encoder.

For each set of ALF coefficients (i.e. corresponding to different classes) a scaling factor K is added, so coefficients are represented as and each set of coefficients is signalled in APS with the corresponding scale factor.

Simulated annealing is used at encoder to find ALF coefficients. The simulated annealing process is launched several times with different initial solutions. First, the existing coordinate descent optimization method from ECM is run, and the first simulated annealing run is using that as a starting point. Then simulated annealing is run 99 more times, where each run takes the output of the previous run as its starting point.

A random number generator is extensively utilized in the simulated annealing approach. There is a variety of pseudo-random generators that can be used, for example, Xorshift. To provide an encoder reproducibility, the same hardcoded pseudo-random number generator is used with a fixed seed for each simulated annealing launch.

Encoder optimization is also applied, separate test results are provided.

Test 4.1a: Restrictions on ALF coefficient values

Test 4.1b: Encoder optimization

**Test 4.6: Huffman coding and simulated annealing for CCALF (**[**JVET-AK0198**](https://jvet-experts.org/doc_end_user/current_document.php?id=15185)**)**

In this test, Huffman coding and simulating annealing methods of Test 4.1 are applied to CCALF, coefficients of CCALF are already restricted to a power of 2.

The parameters for the early termination are reduced from EE2-4.1 to accelerate the computation time. If there are no newly accepted solutions for iterations in a row, the simulated annealing procedure is stopped, where is the number of coefficients of CCALF.

Huffman code is derived from each coefficient and its frequency of occurrence. The frequencies are collected by compressing BVI-DVC sequences using ECM-based encoder.

Test 4.6a: Huffman coding for CCALF + simulated annealing for CCALF

Test 4.6b: Test 4.1a + Test 4.6a

**Test 4.2: Using Laplacian information in ALF (**[**JVET-AK0091**](https://jvet-experts.org/doc_end_user/current_document.php?id=15062)**)**

In the test, Laplacian of post-SAO buffer in deriving online ALF filters is introduced. At block boundaries, Gaussian pre-smoothing is applied before computing Laplacian. For deriving discrete Laplacian values 2D kernel [0,1,0; 1,-4,1; 0,1,0] is employed.

New filter shape, 5x5 cross shape (taps #36 - #40) is applied to Laplacian buffer as shown in the next figure.

A crossword puzzle with numbers

Description automatically generated

**Test 4.3: Boundary-aware offset refinement for in-loop filters (**[**JVET-AK0121**](https://jvet-experts.org/doc_end_user/current_document.php?id=15092)**)**

In ECM, in-loop-filters are performed in the following order: DBF, BF/SAO/CCSAO in parallel, and ALF. The output sample values of one stage will be directly used as the input of the next stage without any refinement.

In the test, an offset between the input and output sample value of one in-loop-filter is further refined by a non-linear function, the refinement parameters depend on whether a sample is located at a block boundary or not. The proposed method is applied to ALF in Test 4.3a and applied to both DBF and ALF in Test 4.3b.

The refined offset denoted as is derived from the original offset *O* as:

,

where {*A, B, C*} are parameters equal to {6, 31, 5} or {5, 31, 5}, which is selected at the encoder and signalled to the decoder in SPS.

The partitioning information is utilized to determine the boundary label. Samples located at a CU or TU boundary is labeled as “boundary” while other samples are labeled as “non-boundary”. The parameter K is different for the boundary and non-boundary samples.

A flag is signalled in the slice header to control the method, if enabled parameter set index is further signalled in the slice header.

Test 4.3a: Boundary-aware offset refinement for ALF

Test 4.3b: Test 4.3a + boundary-aware offset refinement for DBF

**Test 4.5: Reuse of ALF control information (**[**JVET-AK0176**](https://jvet-experts.org/doc_end_user/current_document.php?id=15147)**)**

In the test, ALF control information, including on/off decision and ALF APS usage information, from a previously encoded picture is reused. To indicate this mode, a flag is signalled if ALF is enabled for a picture, ALF control information is reused and not signalled.

At encoder side, ALF filter is derived for a group of pictures jointly which share the ALF on/off decision. The method is only applied to non-reference pictures, bitstreams between the first and last pictures are not written out until the joint ALF optimization of the group is done.

For CTC in RA, a group contains at most 4 non-reference pictures, so the longest delay is 4 pictures at the encoder side, and there is no delay at decoder side. This test increases the number of APSs from 8 to 12.

Test 4.5a: ALF joint optimization (encoder only)

Test 4.5b: ALF joint optimization (encoder only) for VTM-11.0ecm15.0.

Test 4.5c: Increase of the number of ALF APSs to 12

Test 4.5d: Test 4.5a + Test 4.5c

Test 4.5e: Test 4.5a + reuse of ALF control information

Test 4.5f: Test 4.5d + reuse of ALF control information

**Test 4.7: Temporal ALF (**[**JVET-AK0065**](https://jvet-experts.org/doc_end_user/current_document.php?id=15036)**)**

In this test, a temporal adaptive loop filter is introduced that uses the reconstructed samples in the reference pictures. Like ALF, the filter coefficients and related information are transmitted in APS which is specific for this temporal filter.

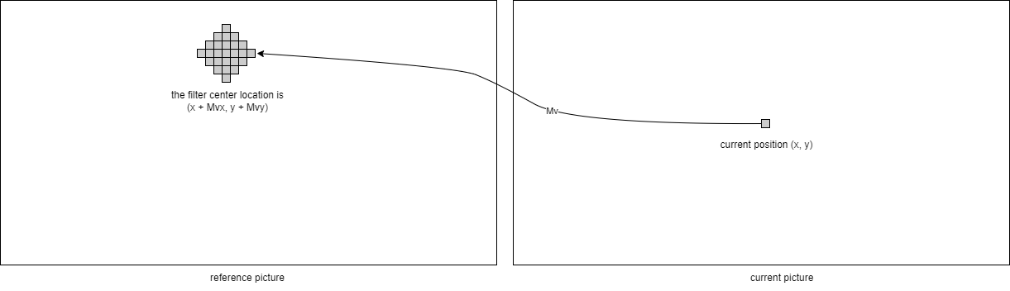
It generates offsets to the luma ALF’s output shown in the next figure, the usage for each CTB is signalled by a flag.

A black background with white squares and black text

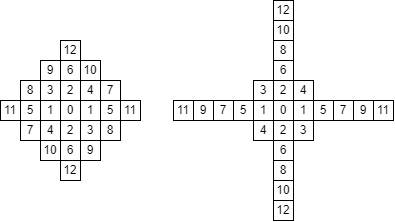
Description automatically generated

The method consists of 6 filtering modes:

1. Uni-filtering mode, which uses the rounded MV0 in the reference picture identified by ref\_idx\_l0 in List 0 to locate the filter inputs (an shown in the next figure). The positions without existing MV0 will be skipped.
2. Uni-filtering mode, which uses the rounded MV1 in the reference picture identified by ref\_idx\_l1 in List 1 to locate the filter inputs. The positions without existing MV1 will be skipped.
3. Bi-filtering mode, which uses the rounded MV0 and MV1 in the reference pictures identified by ref\_idx\_l0 in List 0 and ref\_idx\_l1 in List1 to locate the filter inputs. The reconstructed samples from both filter inputs are averaged before feeding into the temporal filter. The positions without existing MV0 and MV1 will be skipped.
4. Uni-filtering mode, it uses the collocated reconstructed samples in the closest reference picture as the filter inputs.
5. Uni-filtering mode, it uses the collocated reconstructed samples in the second closest reference picture as the filter inputs.
6. Bi-filtering mode, it uses the collocated reconstructed samples in the first and the second closest reference pictures as the filter inputs. Like the third mode. The reconstructed pixels from the two reference pictures are averaged before feeding into the filter.



Two filter shapes having 13 coefficients is used shown in the next figure.



Test 4.7: Temporal ALF

**Test 4.8:  Integration of NN-based ILF in ALF (**[**JVET-AK0183**](https://jvet-experts.org/doc_end_user/current_document.php?id=15154)**)**

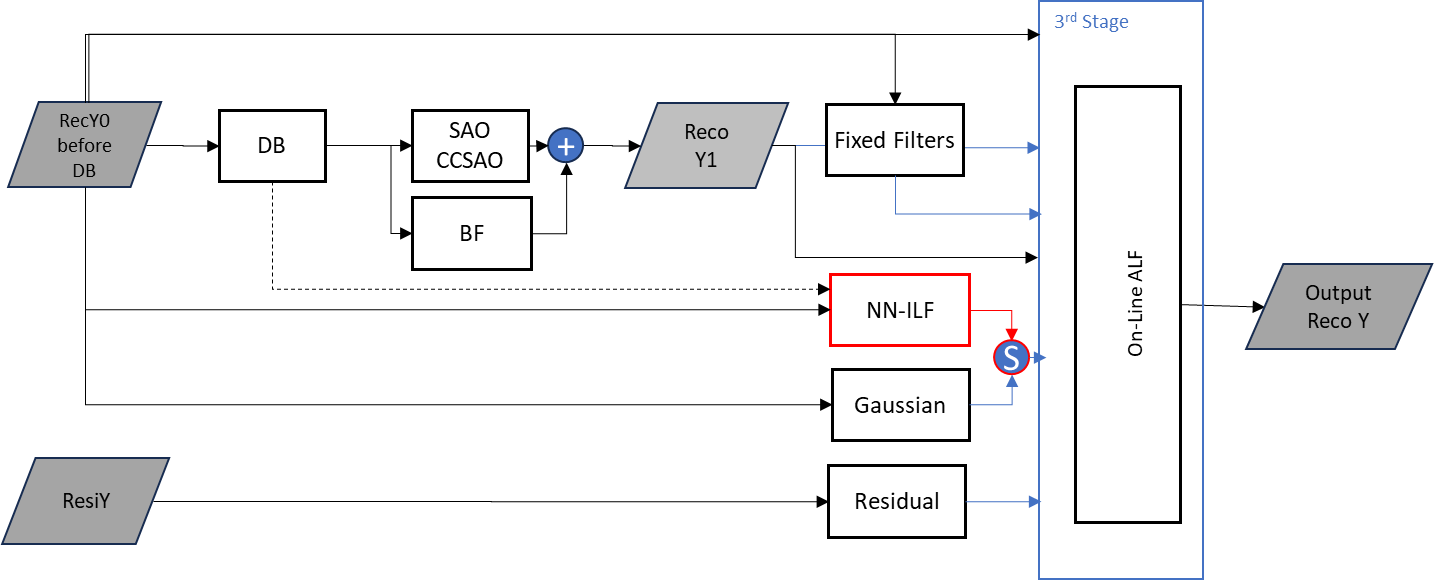
NNVC VLOP architecture is shown in the next figure. In that design, Y, Cb and Cr samples are being filtered jointly. Luma branch employs a cascade of 8 backbones blocks, with chroma branch comprising a single backbone block. Each backbone block implements 3D convolution with separable 1x3, 3x1 and 1x1 kernels.

A diagram of a computer

Description automatically generated

In the test, chroma data is not used for NN input and no chroma branch processing is performed. Modified NN luma-only filters was retrained on DIV2K, BVI and TVD ECM12 coded dataset following the NNVC training guidelines. Resulting filter has complexity of 4.2 kMAC/pixel and 13.4K parameters of the side information.

Inference interface of NNVC ILF was implemented in ECM, with NN-ILF being introduced as an extension of the ECM-ALF filter. NN-filter operates on input reconstructed pixels in parallel to existing fixed, residual, and Gaussian filters. Simplified block diagram of the ECM-ALF extension with NN-ILF is shown below.



Complexity-aware RDO was introduced, where NN filtering is disabled for blocks, if provided distortion reduction is found to be below a threshold.

Test 4.8a1: NN-based ILF with ALF

Test 4.8a2: NN-based ILF with ALF with encoder optimization

Test 4.8b: NN-based ILF with joint YCbCr VLOP model from NNVC. In this test, chroma data is also used as an input as in NNVC, but NN-ILF output chroma pixels are not used.

**Test 4.9:  ALF-CCCM (**[**JVET-AK0088**](https://jvet-experts.org/doc_end_user/current_document.php?id=15059)**)**

In the test, an additional cross-component filtering scheme is examined where cross-component filters are derived at the decoder-side as shown in the next figure.

A diagram of a computer system

Description automatically generated

Each CTU is divided into non-overlapping blocks and for each block the cross-component filter coefficients are derived using SAO outputs. The output samples of luma ALF are used as input to the CCCM like filtering. To obtain a correction signal SAO chroma samples are subtracted from the CCCM like filtering output samples. The correction is weighted by 0.5 and added to the ALF chroma output, the reconstruction becomes *chromaRec =* *alfChroma + ccAlf + 0.5\*cccmCorrection*.

The encoder decides the best block size for each CTU using a rate-distortion optimization (RDO) loop. There are eight possible blocks sizes 4x4, 8x2, 2x8, 8x8, 16x16, 32x32, 64x64, 128x128.

For each CTU, the encoder’s RDO decides the best cross-component model from eight possible models, they differ in the number of samples, non-linear term, and biases. The nonlinear and bias terms are the same as in the 7-tap CCCM model.

CCCM solver is used for deriving the filter coefficients. The 6-tap down sampling filter (used in CCCM and CCLM) is used for mapping the co-located luma into the chroma grid.

For inter coding, a picture may inherit the block sizes and model types for all CTUs from a reference picture. The reference picture is derived from the L0 and L1 lists. Only reference pictures with ALF-CCCM present in at least one CTU are considered. The reference picture with the smallest POC distance to the current POC is selected. If activated, the picture level inheritance will skip the CTU-level signalling completely for the current picture.

Test 4.9a: ALF-CCCM

Test 4.9a\*: ALF-CCCM disabled for I-slices

Test 4.9b: Test 4.9a without the use of reference pictures



Gain in 4.1a comes from saving signalling, which requires another step of encoder optimization. Same optimization applied without the normative change in 4.1b does not give benefit. Results are confirmed by cross-checker who supports the proposal.

Benefit from 4.6a is not obvious. In LB, the losses in luma and gains in chroma also likely would compensate. Combination 4.6b does also nor show benefit compared to 4.1a standalone.

Only 4.1a shows acceptable tradeoff

Decision: Adopt JVET-AK0123 test 4.1a.

Test 4.2 shows some benefit for RA and LB. Cross-checker reports matching of results so far and expresses support for the proposal. Another crosscheck with complete results also reports match, and also expresses support.

It was reported that the only relevant change is introducing a new 5x5 shape for Laplacian at SAO input

Decision: Adopt JVET-AK0091 test 4.2.

Test 4.3 is asserted to be a simple extension of current approach with scaling factor. Tradeoff is similar as in 4.2 (gain in RA and LB), though with lower gain. Among the two variants, 4.3b (applying for both ALF and DBF has slightly higher gain.

Support expressed by cross-checkers (confirming results that were obtained so far were matching, and also expressing that the change is simple and straightforward), and other experts.

Decision: Adopt JVET-AK0121 test 4.3b.

Test 4.5 is not applied for LB, as it is introducing at least one frame lookahead delay. Only applied to non-reference frames, but this still requires multi-frame optimization.

With encoder-only method, already most of the gain can be achieved (0.07% vs. 0.10% comparing 4.5a against 4.5e).

It would not be relevant for CTC, and also not for an anchor in a Call, as typically optimization over multiple frames would be forbidden in the latter.

No consensus reached for inclusion – except for proponents, two experts supported, two opposed.

Test 4.7 uses filtered samples (coming from the integer position of the motion compensation applied to the reference picture), and generates a residual that is added to the ALF output. The gain is close to 0.1% in both RA and LB, and the tradeoff is acceptable. Cross-checkers report matching of results, and support the proposal.

It was asked if the proponents also investigated using the prediction directly, or if they considered to use fixed filters.

It was commented that this is bringing a new aspect to ECM, which could have more potential for being further optimized.

Decision: Adopt JVET-AK0065 test 4.7.

Test 4.8 uses a NN-based filter (VLOP without chroma and retrained in tests a1 and a2, same as VLOP in b) and blends it with Gaussian filter as input to ALF.

a2 has RDO at encoder which enforces less usage and therefore reduces decoder run time, while having less coding gain. It was pointed out that a similar approach was investigated in EE1-1.5.

No training crosscheck on tests a with the retrained model.

Complexity of a1/a2 is 4.2 kMAC/pix, 13K parameters, approx.. 40 convolutional layers.

Joint filtering is difficult to implement in ECM. It was found beneficial to conduct separate processing.

Filter is implemented in SADL, potentially runtime could be further reduced by planned SADL optimization.

Best tradeoff for 4.8a2, but still too high increase in decoder runtime (>30% for 0.6% gain in RA).

It was suggested to include the 4.8b method (VLOP1) in ECM as optional tool, also with the option of replacing it with any subsequent VLOP version. This would have the benefit to do investigations on combinations of EE1 and EE2 tools.

It was agreed that having an interface in ECM that allows testing of loop filters developed in EE1 together with EE2 tools is beneficial (this should be an optional part of ECM, not in CTC, but for studies in AHGs). It is asserted that the software developed in EE2-4.8b is a valuable step in this direction. The interface should not only contain the input/output of NNLF, but also control mechanisms at block level, HLS like flags, and encoder decision parts. Relevant experts were asked to discuss this offline and report back. See further notes under JVET-AK0337.

It was also commented that it would be desirable for this interface working for both luma and chroma inputs/outputs.

Further investigate 4.8a2 in EE, targeting significant reduction in decoding time.

Test 4.9x shows good benefit for chroma in particular for RA and LB (less in AI). Test 4.9a was the original proposal, 4.9b (not using reference pictures, e.g. not inheriting ALF parameters from them) was requested as an additional information in EE. The latter shows more loss in luma (likely due to more parameters to be sent), but also more gain in chroma.

It was commented that a problem of the method is additional latency, as chroma filtering cannot be started before luma filtering is done at decoder. It is however argued that ECM may have already similar or bigger problems in terms of parallelism/pipelining support than this one.

It was also commented that encoder parallelization may be difficult to implement.

Several experts expressed concerns on these aspects, considering that the main benefit is in LB. It was suggested to study the method for one more round in EE, also testing a configuration where the chroma processing is aligned in loop position with other chroma filters, such that the latency/parallelism problems could be resolved, and see how much of the gain would be retained.

Further study in EE.

### EE2 contributions: Enhanced compression beyond VVC capability (39)

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-AK0056](https://jvet-experts.org/doc_end_user/current_document.php?id=15027) EE2-1.9: Weighted OBIC [P. Andrivon, M. Blestel, N. Zouidi (Ofinno)]

[JVET-AK0275](https://jvet-experts.org/doc_end_user/current_document.php?id=15264) Cross-check of JVET-AK0056 EE2-1.9: Weighted OBIC [F. Le Léannec (InterDigital)] [late]

[JVET-AK0059](https://jvet-experts.org/doc_end_user/current_document.php?id=15030) EE2-1.2: Intra mode coding based on HoG of neighboring templates [M. Hong, J. Nam, J. Lim, S. Kim (LGE)]

[JVET-AK0306](https://jvet-experts.org/doc_end_user/current_document.php?id=15295) Crosscheck of JVET-AK0059 (EE2-1.2: Intra mode coding based on HoG of neighboring templates) [M. Coban (Qualcomm)] [late] [miss]

[JVET-AK0060](https://jvet-experts.org/doc_end_user/current_document.php?id=15031) EE2-1.7: Block vector guided DIMD [L. Zhang, Y. Yu, H. Yu, D. Wang (OPPO)]

[JVET-AK0267](https://jvet-experts.org/doc_end_user/current_document.php?id=15256) Crosscheck of JVET-AK0060 (EE2-1.7: Block vector guided DIMD) [N. Yan (Kwai)] [late]

[JVET-AK0061](https://jvet-experts.org/doc_end_user/current_document.php?id=15032) EE2-1.12: On MPM with matrix-based position dependent intra prediction [Z. Sun, Y. Yu, H. Yu, L. Xu, D. Wang (OPPO)]

[JVET-AK0231](https://jvet-experts.org/doc_end_user/current_document.php?id=15220) Crosscheck of JVET-AK0061 (EE2-1.12: On MPM with matrix-based position dependent intra prediction) [X. Li (Alibaba)] [late]

[JVET-AK0062](https://jvet-experts.org/doc_end_user/current_document.php?id=15033) EE2-1.14: block vector guided EIP [Z. Xie, Y. Yu, H. Yu, D. Wang (OPPO)]

[JVET-AK0299](https://jvet-experts.org/doc_end_user/current_document.php?id=15288) Crosscheck of JVET-AK0062 (EE2-1.14: block vector guided EIP) [Z. Zhang (Alibaba)] [late]

[JVET-AK0063](https://jvet-experts.org/doc_end_user/current_document.php?id=15034) EE2-2.4: GPM extension [H. Zhang, F. Wang, Y. Yu, H. Yu, D. Wang (OPPO)]

[JVET-AK0266](https://jvet-experts.org/doc_end_user/current_document.php?id=15255) Crosscheck of JVET-AK0063 (EE2-2.4: GPM extension) [Y. Ahn (LGE)] [late]

[JVET-AK0064](https://jvet-experts.org/doc_end_user/current_document.php?id=15035) EE2-3.2: LFNST/NSPT set derivation for CCP coded block [H. Huang, Y. Yu, H. Yu, D. Wang (OPPO)]

[JVET-AK0280](https://jvet-experts.org/doc_end_user/current_document.php?id=15269) Crosscheck of JVET-AK0064 (EE2-3.2: LFNST/NSPT set derivation for CCP coded block) [M. Hong (LGE)] [late]

[JVET-AK0307](https://jvet-experts.org/doc_end_user/current_document.php?id=15296) Crosscheck of JVET-AK0064 (EE2-3.2: LFNST/NSPT set derivation for CCP coded block) [M. Coban (Qualcomm)] [late] [miss]

[JVET-AK0065](https://jvet-experts.org/doc_end_user/current_document.php?id=15036) EE2-4.7: Temporal ALF [L. Xu, Y. Yu, H. Yu, N. Song, D. Wang (OPPO)]

[JVET-AK0230](https://jvet-experts.org/doc_end_user/current_document.php?id=15219) Crosscheck of JVET-AK0065(EE2-4.7: Temporal ALF) [W. Yin (Bytedance)] [late]

[JVET-AK0076](https://jvet-experts.org/doc_end_user/current_document.php?id=15047) EE2-1.1: Extended OBMC for non-inter blocks [Y. Kidani, H. Katou, K. Kawamura (KDDI), X. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba)]

[JVET-AK0269](https://jvet-experts.org/doc_end_user/current_document.php?id=15258) Crosscheck of Test 1.1a in JVET-AK0076 (EE2-1.1: Extended OBMC for non-inter blocks) [D. Kim, S.-C. Lim (ETRI)] [late]

[JVET-AK0293](https://jvet-experts.org/doc_end_user/current_document.php?id=15282) Crosscheck of Test 1.1a in JVET-AK0076 (EE2-1.1: Extended OBMC for non-inter blocks) [R. Yu (Qualcomm)] [late]

[JVET-AK0084](https://jvet-experts.org/doc_end_user/current_document.php?id=15055) EE2-1.13: Multiple filter taps for EIP [Z. Lv, C. Zhou, G. Wang (vivo)] [late]

[JVET-AK0291](https://jvet-experts.org/doc_end_user/current_document.php?id=15280) Crosscheck of JVET-AK0084 (EE2-1.13: Multiple filter taps for EIP) [Z. Deng (Bytedance)] [late] [miss]

[JVET-AK0085](https://jvet-experts.org/doc_end_user/current_document.php?id=15056) EE2-2.6: Template Matching Padding [N. Neumann (Nokia), M. Wien (RWTH Aachen Univ.)]

[JVET-AK0244](https://jvet-experts.org/doc_end_user/current_document.php?id=15233) Crosscheck JVET-AK0085 on EE2-2.6 Template Matching Padding [X. Li (Google)] [late]

[JVET-AK0086](https://jvet-experts.org/doc_end_user/current_document.php?id=15057) EE2-1.2a: Improvement on non-MPM [G. Wang, C. Zhou, Z. Lv (vivo)]

[JVET-AK0245](https://jvet-experts.org/doc_end_user/current_document.php?id=15234) Crosscheck of JVET-AK0086 (EE2-1.2a: Improvement on non-MPM) [J. Fu (PKU)] [late]

[JVET-AK0087](https://jvet-experts.org/doc_end_user/current_document.php?id=15058) EE2-1.11: 8-tap interpolation filter for angular intra prediction [T. Dong, A. Filippov, V. Rufitskiy, J. Konieczny, H. Qin, J. Zhang (TCL)]

[JVET-AK0303](https://jvet-experts.org/doc_end_user/current_document.php?id=15292) Crosscheck of JVET-AK0087 (EE2-1.11: 8-tap interpolation filter for angular intra prediction) [S. Park (KNU)] [late] [miss]

[JVET-AK0318](https://jvet-experts.org/doc_end_user/current_document.php?id=15307) Crosscheck of JVET-AK0087 (EE2-1.11b: 8-tap interpolation filter for angular intra prediction) [G. Verba (Qualcomm)] [late] [miss]

[JVET-AK0088](https://jvet-experts.org/doc_end_user/current_document.php?id=15059) EE2-4.9: ALF-CCCM [P. Astola, I. Jumakulyyev, D. Bugdayci Sansli, J. Lainema (Nokia)]

[JVET-AK0248](https://jvet-experts.org/doc_end_user/current_document.php?id=15237) Crosscheck of JVET-AK0088 (EE2-4.9: ALF-CCCM) [R. G. Youvalari (Xiaomi)] [late]

[JVET-AK0091](https://jvet-experts.org/doc_end_user/current_document.php?id=15062) EE2-4.2: Using Laplacian information in ALF [I. Jumakulyyev, D. Bugdayci Sansli, I. Zupancic, J. Lainema (Nokia)]

[JVET-AK0129](https://jvet-experts.org/doc_end_user/current_document.php?id=15100) Crosscheck of JVET-AK0091 (EE2-4.2: Using Laplacian information in ALF) [V. Shchukin (Ericsson)] [late]

[JVET-AK0314](https://jvet-experts.org/doc_end_user/current_document.php?id=15303) Crosscheck of JVET-AK0091 (EE2-4.2: Using Laplacian information in ALF) [V. Rufitskiy (TCL)] [late]

[JVET-AK0095](https://jvet-experts.org/doc_end_user/current_document.php?id=15066) EE2-2.2: Enhanced derivation of affine merge candidates [L. Zhao, K. Zhang, Z. Deng, N. Zhang, Y. Wang, W. Yin, M. Salehifar, L. Zhang (Bytedance)]

[JVET-AK0304](https://jvet-experts.org/doc_end_user/current_document.php?id=15293) Crosscheck of JVET-AK0095 (EE2-2.2: Enhanced derivation of affine merge candidates) [Y. Zhang (Qualcomm)] [late]

[JVET-AK0097](https://jvet-experts.org/doc_end_user/current_document.php?id=15068) EE2-3.4: Last significant coefficient position signaling with secondary prefix [F. Le Léannec, M. Balcilar, C. Salmon-Legagneur, F. Galpin (InterDigital)]

[JVET-AK0259](https://jvet-experts.org/doc_end_user/current_document.php?id=15248) Cross-check of JVET-AK0097 (EE2-3.4: Last significant coefficient position signaling with secondary prefix) [P. Nikitin (Qualcomm)] [late] [miss]

[JVET-AK0298](https://jvet-experts.org/doc_end_user/current_document.php?id=15287) Crosscheck of JVET-AK0097 (EE2-3.4: Last significant coefficient position signaling with secondary prefix) T. N. Canh (Dolby Labs) [late]

[JVET-AK0098](https://jvet-experts.org/doc_end_user/current_document.php?id=15069) EE2-3.6a/3.6b: combinations of tests 3.4 and 3.5a/3.5b [F. Le Léannec, M. Balcilar, C. Salmon-Legagneur, F. Galpin (InterDigital)]

[JVET-AK0261](https://jvet-experts.org/doc_end_user/current_document.php?id=15250) Cross-check of JVET-AK0098 (EE2-3.6a/3.6b: combinations of tests 3.4 and 3.5a/3.5b) P. Nikitin (Qualcomm)

[JVET-AK0101](https://jvet-experts.org/doc_end_user/current_document.php?id=15072) EE2-2.3: Regression-based GPM with intra and inter prediction [K. Jia, J. Chen, X. Li, R.-L. Liao, Y. Ye (Alibaba)]

[JVET-AK0268](https://jvet-experts.org/doc_end_user/current_document.php?id=15257) Crosscheck of JVET-AK0101 (EE2-2.3: Regression-based GPM with intra and inter prediction) [Y. Wang (Bytedance)] [late]

[JVET-AK0294](https://jvet-experts.org/doc_end_user/current_document.php?id=15283) Crosscheck of JVET-AK0101 (EE2-2.3: Regression-based GPM with intra and inter prediction) [C. Zhou (vivo)] [late]

[JVET-AK0102](https://jvet-experts.org/doc_end_user/current_document.php?id=15073) EE2-1.16: On intra interpolation filter [Z. Zhang, J. Chen, X. Li, R.-L. Liao, Y. Ye (Alibaba), C. Zhou, Z. Lv, G. Wang (vivo)]

[JVET-AK0309](https://jvet-experts.org/doc_end_user/current_document.php?id=15298) Crosscheck of JVET-AK0102 (EE2-1.16: On intra interpolation filter) [Z. Xie (OPPO)] [late]

[JVET-AK0104](https://jvet-experts.org/doc_end_user/current_document.php?id=15075) EE2-1.8: Non-adjacent DIMD for TMRL [V. Rufitskiy, A. Filippov, T. Dong, H. Qin, J. Konieczny, K. Ding (TCL)]

[JVET-AK0297](https://jvet-experts.org/doc_end_user/current_document.php?id=15286) Crosscheck of JVET-AK0104 (EE2-1.8: Non-adjacent DIMD for TMRL) [S. Blasi (Nokia)] [late]

[JVET-AK0105](https://jvet-experts.org/doc_end_user/current_document.php?id=15076) EE2-3.5: High Resolution Arithmetic Engine [M. Balcilar, F. Le Leannec, C. Salmon-Legagneur, F. Galpin (InterDigital)]

[JVET-AK0260](https://jvet-experts.org/doc_end_user/current_document.php?id=15249) Cross-check of JVET-AK0105 (EE2-3.5: High Resolution Arithmetic Engine) [[P. Nikitin (Qualcomm)](mailto:pnikitin@qti.qualcomm.com)] [late] [miss]

[JVET-AK0118](https://jvet-experts.org/doc_end_user/current_document.php?id=15089) EE2-1.10: Bilateral Filtering for Intra Prediction [W. Yin, K. Zhang, Y. Wang, Z. Deng, N. Zhang, L. Zhao, M. Salehifar, L. Zhang (Bytedance)]

[JVET-AK0130](https://jvet-experts.org/doc_end_user/current_document.php?id=15101) Crosscheck of JVET-AK0118 (EE2-1.10: Bilateral Filtering for Intra Prediction) [V. Shchukin (Ericsson)] [late]

[JVET-AK0295](https://jvet-experts.org/doc_end_user/current_document.php?id=15284) Crosscheck of JVET-AK0118 (EE2-1.10: Bilateral Filtering for Intra Prediction) [C. Ma (Kwai)] [late]

[JVET-AK0121](https://jvet-experts.org/doc_end_user/current_document.php?id=15092) EE2-4.3: Boundary-Aware Offset Refinement for In-Loop-Filters [W. Yin, K. Zhang, Y. Wang, Z. Deng, N. Zhang, L. Zhao, M. Salehifar, L. Zhang (Bytedance)]

[JVET-AK0229](https://jvet-experts.org/doc_end_user/current_document.php?id=15218) Crosscheck of JVET-AK0121 (EE2-4.3: Boundary-Aware Offset Refinement for In-Loop-Filters) [N. Hu (Qualcomm)] [late]

[JVET-AK0123](https://jvet-experts.org/doc_end_user/current_document.php?id=15094) EE2-4.1: Restrictions on ALF coefficient values [V. Shchukin, P. Wennersten, J. Ström (Ericsson)]

[JVET-AK0172](https://jvet-experts.org/doc_end_user/current_document.php?id=15143) Crosscheck of JVET-AK0123 (EE2-4.1: Restrictions on ALF coefficient values) [I. Jumakulyyev (Nokia)] [late]

[JVET-AK0131](https://jvet-experts.org/doc_end_user/current_document.php?id=15102) EE2-3.3: Advanced SBT with direction and position inference [G. Laroche, P. Onno (Canon)]

[JVET-AK0276](https://jvet-experts.org/doc_end_user/current_document.php?id=15265) Cross-check of JVET-AK0131 EE2-3.3: Advanced SBT with direction and position inference [F. Le Léannec (InterDigital)] [late]

[JVET-AK0148](https://jvet-experts.org/doc_end_user/current_document.php?id=15119) EE2-1.17: Combination of EE2-1.16, EE2-1.11 and EE2-1.10 [Z. Zhang, J. Chen, X. Li, R.-L. Liao, Y. Ye (Alibaba), C. Zhou, Z. Lv, G. Wang (vivo), T. Dong, A. Filippov, J. Konieczny, V. Rufitskiy, H. Qin, K. Ding (TCL), W. Yin, K. Zhang, Y. Wang, Z. Deng, N. Zhang, L. Zhao, M. Salehifar, L. Zhang (Bytedance)]

[JVET-AK0296](https://jvet-experts.org/doc_end_user/current_document.php?id=15285) Crosscheck of JVET-AK0148 (EE2-1.17: Combination of EE2-1.16, EE2-1.11 and EE2-1.10) [C. Ma (Kwai)] [late]

[JVET-AK0310](https://jvet-experts.org/doc_end_user/current_document.php?id=15299) Crosscheck of JVET-AK0148 (EE2-1.17: Combination of EE2-1.16, EE2-1.11 and EE2-1.10) [Z. Xie (OPPO)] [late]

[JVET-AK0149](https://jvet-experts.org/doc_end_user/current_document.php?id=15120) EE2-1.15: a combination of EE2-1.13 and EE2-1.14 [Z. Lv, C. Zhou, G. Wang (vivo), Z. Xie, Y. Yu, H. Yu, D. Wang (OPPO)]

[JVET-AK0240](https://jvet-experts.org/doc_end_user/current_document.php?id=15229) Crosscheck of JVET-AK0149 (EE2-1.15: a combination of EE2-1.13 and EE2-1.14) [Z. Zhang (Alibaba)] [late]

[JVET-AK0161](https://jvet-experts.org/doc_end_user/current_document.php?id=15132) EE2-1.6: Extended IntraTMP merge candidate list with an improved fusion mode [D. Ruiz Coll, J.-K. Lee (Ofinno)]

[JVET-AK0271](https://jvet-experts.org/doc_end_user/current_document.php?id=15260) Crosscheck of JVET-AK0161 (EE2-1.6: Extended IntraTMP merge candidate list with an improved fusion mode) [W. Lim, S.-C. Lim (ETRI)] [late]

[JVET-AK0176](https://jvet-experts.org/doc_end_user/current_document.php?id=15147) EE2-4.5: Reuse of ALF control information [M. Karczewicz, H. Wang, N. Hu, V. Seregin (Qualcomm)]

[JVET-AK0226](https://jvet-experts.org/doc_end_user/current_document.php?id=15215) Crosscheck of JVET-AK0176 (EE2-4.5: Reuse of ALF Control Information) [W. Yin (Bytedance)] [late]

[JVET-AK0183](https://jvet-experts.org/doc_end_user/current_document.php?id=15154) EE2-4.8: Integration of NN-based ILF in ALF [D. Rusanovskyy, Y. Li, M. Karczewicz (Qualcomm)]

[JVET-AK0264](https://jvet-experts.org/doc_end_user/current_document.php?id=15253) Crosscheck of JVET-AK0183 (EE2-4.8: Integration of NN-based ILF in ALF) [F. Galpin, T. Poirier (InterDigital)] [late]

[JVET-AK0185](https://jvet-experts.org/doc_end_user/current_document.php?id=15156) EE2-2.5: TMVP Candidate Selection [S. Hong, L. Wang, K. Panusopone (Nokia)]

[JVET-AK0249](https://jvet-experts.org/doc_end_user/current_document.php?id=15238) Crosscheck of JVET-AK0185 (EE2-2.5: TMVP Candidate Selection) [R. G. Youvalari (Xiaomi)] [late]

[JVET-AK0187](https://jvet-experts.org/doc_end_user/current_document.php?id=15158) EE2-3.8: Implicit MTS extension [P. Garus, M. Karczewicz, M. Coban, V. Seregin, H. Wang, P. Nikitin (Qualcomm)]

[JVET-AK0285](https://jvet-experts.org/doc_end_user/current_document.php?id=15274) Crosscheck of JVET-AK0187 (EE2-3.8: Implicit MTS extension) [M. Salehifar (Bytedance)] [late]

[JVET-AK0198](https://jvet-experts.org/doc_end_user/current_document.php?id=15185) EE2-4.6: Huffman coding and simulated annealing for CCALF [K. Takada, S. Deshpande (Sharp), V. Shchukin (Ericsson)]

[JVET-AK0227](https://jvet-experts.org/doc_end_user/current_document.php?id=15216) Crosscheck of JVET-AK0198 (EE2-4.6: Huffman Coding and Simulated Annealing for CCALF) [W. Yin (Bytedance)] [late]

[JVET-AK0217](https://jvet-experts.org/doc_end_user/current_document.php?id=15206) EE2-3.1: Multiple transform set selection for intra LFNST/NSPT with subsampled DIMD for transform set selection [M. Coban, M. Karczewicz, P. Nikitin, P. Garus, B. Ray, V. Seregin (Qualcomm), F. Wang, Y. Yu, H. Yu, D. Wang (OPPO), C. Bonnineau, K. Naser, S. Puri, F. Le Léannec (InterDigital), L. Zhao, K. Zhang, L. Zhang (Bytedance)]

[JVET-AK0233](https://jvet-experts.org/doc_end_user/current_document.php?id=15222) Crosscheck of JVET-AK0217 (EE2-3.1: Multiple transform set selection for intra LFNST/NSPT with subsampled DIMD for transform set selection) [M. Hong (LGE)] [late]

[JVET-AK0220](https://jvet-experts.org/doc_end_user/current_document.php?id=15209) EE2-1.3: Intra TMP improvements [M. Radosavljević, T. Dumas, K. Naser, F. Le Léannec (InterDigital)] [late]

Initial version rejected as”placeholder”

[JVET-AK0243](https://jvet-experts.org/doc_end_user/current_document.php?id=15232) Crosscheck of JVET-AK0220 (EE2-1.3: Intra TMP improvements) [D. Ruiz Coll, J.-K. Lee (Ofinno)] [late]

[JVET-AK0222](https://jvet-experts.org/doc_end_user/current_document.php?id=15211) EE2-3.7: Improved LFNST/NSPT kernel set selection for SGPM [S. Puri, C. Bonnineau, K. Naser, F. Le Léannec (InterDigital)]

[JVET-AK0273](https://jvet-experts.org/doc_end_user/current_document.php?id=15262) Crosscheck of JVET-AK0222 (EE2-3.7: Improved LFNST/NSPT kernel set selection for SGPM) [I. Zupancic (Nokia)] [late]

[JVET-AK0283](https://jvet-experts.org/doc_end_user/current_document.php?id=15272) EE2-1.18: Combination of EE2-1.7 and EE2-1.9 [P. Andrivon, M. Blestel, N. Zouidi (Ofinno), L. Zhang, Y. Yu, H. Yu, D. Wang (OPPO)] [late]

[JVET-AK0277](https://jvet-experts.org/doc_end_user/current_document.php?id=15266) Cross-check of EE2-1.18: EE2-1.7 + EE2-1.9 [F. Le Léannec (InterDigital)] [late]

[JVET-AK0284](https://jvet-experts.org/doc_end_user/current_document.php?id=15273) EE2-1.19: Combination of EE2-1.7 and EE2-1.8 [V. Rufitskiy, A. Filippov, T. Dong, H. Qin, J. Konieczny, K. Ding (TCL), L. Zhang, Y. Yu, H. Yu, D. Wang (OPPO)] [late]

### EE2 related contributions (3)

Contributions in this area were discussed at 1820–1830 and 1950–2015 on Wednesday 15 Jan. 2025 (chaired by JRO).

[JVET-AK0184](https://jvet-experts.org/doc_end_user/current_document.php?id=15155) EE2-related: Additional results for NN-based ILF in ALF [D. Rusanovskyy, Y. Li, M. Karczewicz (Qualcomm), J. Li, Y. Li, C. Lin, K. Zhang, L. Zhang (Bytedance)] [late]

Contribution JVET-AJ0183 reports EE2-4.8 Test B results, where joint YCbCr VLOP filter of NNVC was tested for luma samples filtering in ECM. In this contribution, Test B results for NNVC LOP2 and HOP3 are reported to provide an initial assessment of the potential gain achievable by more complex NNVC ILF in ECM.

NN-ILF LOP and HOP models shared by JVET NNVC (JVET-AH0014) were used in this test, with no retraining over the ECM data set.

Preliminary comparative results (BD-rate change) vs ECM15 anchor are shown below.

Test B LOP (Partial results for completed sequences B\*, C, D, E):  
AI: -1.4%, 0.0%, 0.0% (Y, Cb, Cr) with 101% EncT and 433% DecT.  
RA: -1.4%, -0.0%, -0.0% (Y, Cb, Cr) with 101% EncT and x% DecT.

Test B HOP (Partial results for completed sequences of B\*, C, D, E):  
AI: -4.3%, 0.0%, 0.0% (Y, Cb, Cr) with 107% EncT and 5554% DecT.  
RA: -4.7%, -0.2%, -0.2% (Y, Cb, Cr) with 107% EncT and x% DecT.

The results could be interpreted to indicate that at most half of the luma gain obtained in NNVC may still be obtained with ECM. This view might change if also chroma filtering would be considered, which is known to perform better in NNLF, and strategy for luma/chroma optimization is different for ECM and NNVC.

Results for RA from encoder loop – currently mismatch between encoder and decoder.

Include this in the study for developing an interface for NNLF in ECM (see notes under EE2-4.8)

[JVET-AK0317](https://jvet-experts.org/doc_end_user/current_document.php?id=15306) EE2-related: Simplification of EE2-3.3 Advanced SBT with direction and position inference [Y. Zhang, J. Ye, H. Wang, V. Seregin, M. Karczewicz (Qualcomm)] [late]

This contribution presents the results of a simplified version of EE2 Tests 3.3a on Advanced SBT (ASBT) mode. Compared to the original EE2 test, simplification has been made on the process to search for best position and direction.

On top of ECM-15.0, the average BDR gains and runtimes are as follows:

-xx% -xx% -xx% xx% xx% for the RA configuration,

-xx% -xx% -xx% xx% xx% for the Low Delay B configuration.

Reduced runtime, and slight gain in RA.

Main aspect is simplification of maximum gradient identification.

Study in EE along with EE2-3.3.

[JVET-AK0329](https://jvet-experts.org/doc_end_user/current_document.php?id=15318) [EE2-3.1b\*][EE2-3.7]: combination of tests 3.1b\* and 3.7 [C. Bonnineau, K. Naser, S. Puri, F. Le Léannec (InterDigital), M. Coban, M. Karczewicz, P. Nikitin, P. Garus, B. Ray, V. Seregin (Qualcomm), F. Wang, Y. Yu, H. Yu, D. Wang (OPPO), L. Zhao, K. Zhang, L. Zhang (ByteDance)] [late]

This contribution reports the results of the combination of test EE2-3.1b\* and test EE2-3.7. The following results are obtained on top of ECM-15.0.

Here is a summary of the obtained results:

AI {-0.06%, 0.02%, 0.05%, 100.4%, 99.9%}

Was presented Mon. 20 Jan. 1330 (chaired by JRO)

The results indicate that there is no negative interaction between the two proposals. The encoder runtime is slightly decreased compared to EE2-3.1b\*, and the gain is slightly increased. The encoder runtime is not homogeneous over classes (class A1 is higher than other classes), but this was already the case in 3.1b\*.

Decision: Adopt JVET-AK0329.

[JVET-AK0331](https://jvet-experts.org/doc_end_user/current_document.php?id=15320) Crosscheck of JVET-AK0329 ([EE2-3.1b\*][EE2-3.7]: combination of tests 3.1b\* and 3.7) [A. Filippov (TCL)] [late]

### ECM modifications and software improvements beyond EE2 (33)

#### Intra and CIIP (6)

Contributions in this area were discussed at 1830–1950 on Wednesday XX Jan. 2025 (chaired by XXX).

[JVET-AK0089](https://jvet-experts.org/doc_end_user/current_document.php?id=15060) Non-EE2: Improvement on MPM [G. Wang, C. Zhou, Z. Lv (vivo)] [late]

In this contribution, an MPM list filling optimization method is proposed. Firstly, the Planar mode will not be fixed in the first position in the MPM list if current CU satisfies the PNN enabling condition. Secondly, besides DIMD mode, two dimdBlendModes instead of one are added and placed at the top of the MPM list.

On top of ECM-15.0, simulation results of the proposed method are:

AI: -0.05% -0.09% -0.04%, EncT: 100.1%, DecT: 99.8%.

The proponents believe that the gain would be retained also after adoption of EE2-1.2c into ECM16. Other experts believe that aspect 2 gain might be lost. According to proponents, aspect 1 gives still approx. 0.03%.

Several experts supported study in EE, considering that the tradeoff is reasonable. The two aspects should also be studied separately.

[JVET-AK0305](https://jvet-experts.org/doc_end_user/current_document.php?id=15294) Crosscheck of JVET-AK0089 (Non-EE2: Improvement on MPM)) [J. Fu (PKU)] [late]

[JVET-AK0090](https://jvet-experts.org/doc_end_user/current_document.php?id=15061) Non-EE2: IntraTMP complexity reduction [Z. Fan, T. Chujoh, T. Ikai (Sharp)]

This proposal aims to reduce the encoding and decoding time of ECM by decreasing the computational complexity of IntraTMP in two aspects. The results are as follows:

AI 0.04% 0.02% 0.01% 97.2% 96.9% RA 0.03% -0.03% 0.03% 99.9% 100.1%

It was commented that the suggested modifications are also implemented in software by variables.

More at the level of “fine tuning” an existing tool. It is not obvious that this is relevant for EE study.

[JVET-AK0250](https://jvet-experts.org/doc_end_user/current_document.php?id=15239) Crosscheck of JVET-AK0090 (Non-EE2: IntraTMP complexity reduction) [F. Wang (OPPO)] [late] [miss]

[JVET-AK0202](https://jvet-experts.org/doc_end_user/current_document.php?id=15191) AHG12: On cross-component intra prediction [Y.-J. Chang, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes to reduce the complexity of cross-component prediction (CCP) by adding some constraints to the existing CCP modes. It is also proposed to evaluate all CCP modes by SATD/SAD cost comparisons before rate distortion evaluation. The simulation results show 0.01%(Y), 0.06(U)%, 0.13%(V)} BD-rate with 95.5% EncT and 99.9% DecT under AI CTC, and {0.xx%(Y), 0.xx(U)%, 0.xx%(V)} BD-rate with xxx.x% EncT and xxx.x% DecT under RA CTC.

1) Since the CCPMerge list is constructed and reordered by template costs, it is natural that the first few candidates with lowest template costs can contribute more gains. It is proposed to constrain the CCPmerge fusion to the first few CCPmerge candidates only.

2) In another aspect, it is well-known that there are two types of CCP modes, i.e., single model and multi-model. The multi-model typically results in higher complexity than single model since it requires more CCP model derivations. Therefore, it is proposed to disable multi-model for some CCP modes while the derivation is using left-only template or top-only template.

3) It is further proposed to evaluate all CCP modes in one single list by SATD/SAD cost comparisons before rate distortion evaluation.

It was asked how much encoding time reduction is achieved by the non-normative aspect 3? This had not been investigated.

Several experts expressed interest to investigate in EE. It was commented that some adoption of the last meeting might give the change for this runtime reduction, in some cases even giving gains.

Results on the three aspects should also be reported standalone.

[JVET-AK0223](https://jvet-experts.org/doc_end_user/current_document.php?id=15212) Non-EE2: CCP merge mode with adjustment [Y. Wang, K. Zhang, W. Yin, Z. Deng, N. Zhang, L. Zhao, M. Salehifar, L. Zhang (Bytedance)]

This contribution presents a cross-component prediction (CCP) merge mode with adjustment. In the proposed method, a CCP model can be updated using an adjustment parameter, which is derived based on a template of the current block. On top of ECM-15.0, the simulation results are reported as below:

AI: {-0.01%, -0.40%, -0.27%; 101.1%, 100.4%};

RA\*: {0.00%, -0.29%, -0.16%; 105.7%, 100.3%}.

\* Two test points are copied from the anchor and running time for RA is not reliable.

Two additional RDO checks are needed.

Several experts found the method interesting and supported investigation in EE.

Current tradeoff in terms of encoding time not acceptable – try to improve, where also combination with JVET-AK0202 could give benefit.

[JVET-AK0234](https://jvet-experts.org/doc_end_user/current_document.php?id=15223) Non-EE2: Subblock-based CCCM [F. Pu, N. C. Thuong, P. Yin, S. McCarthy (Dolby)]

This contribution proposes subblock-based CCCM model for larger chroma prediction blocks to further improve the coding efficiency of CCCM. Subblock partitioning is applied according to the ratio between the width and height of the chroma PU. The maximum number of partitions is 4.

The proposed method is implemented on top of ECM-15.0. The BD-rate achieved over ECM-15.0 for all intra configuration is: { -0.03%, -0.34%, -0.33%, 100.8%, 100.6%}.

The determination of parameters requires decoding of previous subblocks, i.e. subblocks cannot be processed in parallel. As it is only applied to large blocks, this may be acceptable, not worse than in case of smaller blocks.

Split is always performed into 4 subblocks.

It was asked which subblock model would be used for CCP merge mode? This should not be a problem, as the subblocks would be considered as if they would have been smaller blocks.

Several experts found the method interesting and supported investigation in EE.

It was requested to also study the method without performing the TU split along with the CCM subblock split.

[JVET-AK0262](https://jvet-experts.org/doc_end_user/current_document.php?id=15251) EE2-related: On EIP filter shapes [K. Panusopone, M. He, S. Hong, L. Wang, J. Lainema (Nokia)] [late]

This contribution proposes to update the EIP filter shapes. The modified EIP filters include input samples above-right and below-left of the prediction sample. Simulation result, based on ECM-15.0 is reported as follows:

AI: -0.02%, -0.02%, 0.04%.

It was reported that the gain is additive with EE2-1.14 (version without encoder optimization of the latter)

Standalone tradeoff appears reasonable. Several experts expressed support

Investigate in EE along with 1.14

[JVET-AK0288](https://jvet-experts.org/doc_end_user/current_document.php?id=15277) Crosscheck of JVET-AK0262 (EE2-related: On EIP filter shapes) [Y. Yu, L. Xu (OPPO)] [late]

#### Inter (11)

Contributions in this area were discussed at 0900–1110 on Thursday 16 Jan. 2025 (chaired by JRO).

[JVET-AK0096](https://jvet-experts.org/doc_end_user/current_document.php?id=15067) Non-EE2: Additional inter merge candidates [N. Zhang, K. Zhang, Z. Deng, L. Zhao, Y. Wang, W. Yin, M. Salehifar, L. Zhang (Bytedance)]

In this contribution, additional spatial, pairwise-average, and multiple-average candidates are introduced for the merge mode. The proposed method was implemented on top of ECM-15.0 and the simulation results are as follows.

RA: Overall: { -0.04%, -0.06%, 0.02%, 100.4%, 100.2% }.

LB: Overall: { -0.12%, -0.19%, -0.21%, 100.5%, 101.5% }.

Number of pairwise candidates is increased from 4 to 8. This gives roughly half of the gain.

The decoder runtime in LB is increased due to additional TM necessary.

Several experts expressed interest to have this studied in EE.

The benefits of the three aspects should also be investigated separately.

[JVET-AK0108](https://jvet-experts.org/doc_end_user/current_document.php?id=15079) Non-EE2: CMVP extension for constructed affine merge candidates [C. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba)]

This contribution proposes to add affine chained motion vector prediction (CMVP) into affine merge list. Specifically, CMVP are applied to the constructed affine merge candidates CPMVs which are derived using the translation MVs of adjacent and non-adjacent neighbors. In the process of constructing affine candidates, the MVs at four adjacent or three non-adjacent vertices are accumulated with MVs or BVs and collectively construct new affine candidates. On top of ECM-15.0, simulation results are reported as follows:

RA : {-0.02%, 0.03%, 0.05%, 100.5%, 101.0%}, LDB : {-0.07%, -0.18%, 0.17%, 100.3%, 101.2%};

The decoding time is increased due to increase of number of candidates (additional TM necessary).

Interest was expressed by cross-checkers and other experts to study in EE. Interesting gain particularly in LB. To potentially get better tradeoff, it should also be tried to lower the number of additional candidates.

It was commented that the proposal may be in competition with new adoption EE2-2.2c.

[JVET-AK0320](https://jvet-experts.org/doc_end_user/current_document.php?id=15309) Crosscheck of JVET-AK0108 (Non-EE2: CMVP extension for constructed affine merge candidates) [Y. Kidani, K. Kawamura (KDDI)] [late]

[JVET-AK0111](https://jvet-experts.org/doc_end_user/current_document.php?id=15082) Non-EE2: On interpolation filter for template matching [Z. Dai, R.-L. Liao, J. Chen, X. Li, Y. Ye (Alibaba)]

This contribution proposes to replace the 2-tap bilinear interpolation filter in the template matching based tools with a 4-tap interpolation filter, two sets of filter coefficients are proposed. On top of ECM-15.0, the experimental results {Y, U, V, EncT, DecT} are shown as follows:

Filter#0: RA{-0.03%, 0.00%, 0.01%, 100.3%, 100.7%}, LD{-0.09%, -0.24%, 0.03%, 100.1%, 100.5%}.

Filter#1: RA{-0.02%, 0.02%, 0.00%, 100.4%, 100.8%}, LD{-0.05%, -0.07%, -0.08%, 99.9%, 100.4%}.

Interest was expressed by cross-checkers and other experts. Idea is straightforward, and tradeoff in terms of runtime would be acceptable (though it doubtlessly increases computational complexity and memory accesses).The filters are based on DCT-IF design. Filter 0 has higher cutoff and gives slightly better results. It was asked if the filters could be combined. It was also commented that the performance is generally better for lower resolution.

Study in EE. Both filters should be investigated, also to get better understanding about the individual benefits on a sequence by sequence and class by class basis.

Usage in bilateral matching should also be investigated.

[JVET-AK0251](https://jvet-experts.org/doc_end_user/current_document.php?id=15240) Crosscheck of JVET-AK0111 (Non-EE2: On interpolation filter for template matching) ]L. Xu (OPPO)] [late]

[JVET-AK0137](https://jvet-experts.org/doc_end_user/current_document.php?id=15108) Non-EE2: Improvements on inter AMVP candidate list construction [S. Park, H. Jeong, B. Jeon (SKKU)] [late]

This contribution proposes to put zero MV into the AMVP candidate list unless the list is already full so that the zero MV can be subjected to MVP refinement and reordering process. The proposed method achieves BDBR gain compared to ECM-15.0 as follows:

RA: -0.01% (Y) / -0.03% (Cb) / 0.01% (Cr), Enc: xxx.x%, Dec: xxx.x%.

LDB: -0.01% (Y) / -0.12% (Cb) / 0.25% (Cr), Enc: xxx.x%, Dec: xxx.x%

JVET-AK0254 is related and was presented before presenting this proposal. This is a study showing that the zero MV is often selected in AMVP candidate list.

The maximum size of the list is not increased, but one more candidate needs to be checked when the list is not full yet, at the last position of the list after reordering.

Significantly more gain in screen content classes, also loss in some classes.

No interest expressed for EE. As it is beneficial for screen content, it is suggested to further study if better position of the zero candidate in the list could be found dependent on content characteristics.

[JVET-AK0270](https://jvet-experts.org/doc_end_user/current_document.php?id=15259) Crosscheck of JVET-AK0137 (Non-EE2: Improvements on inter AMVP candidate list construction) [D. Kim, S.-C. Lim (ETRI)] [late]

[JVET-AK0162](https://jvet-experts.org/doc_end_user/current_document.php?id=15133) AHG12: On affine motion compensation [H. Huang, R. Yu, Z. Zhang, Y. Zhang, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution extends the usage of per-pixel based motion compensation for affine coded block. It is reported that {-0.06%(Y), 0.04(U)%, 0.01%(V)} BD-rate change with 99.0% EncT and 100.3% DecT can be achieved under RA CTC, and {-0.09%(Y), -0.12(U)%, 0.24%(V)} BD-rate change with 99.5% EncT and 101.6% DecT under LDB CTC.

It is proposed to apply pixel-wise MC in cases where BDOF is not applied.

It was asked if PROF would still be used? In some (very few) cases this would happen.

Acceptable tradeoff, interest expressed to investigate in EE.

Percentage of still using PROF should be reported.

[JVET-AK0186](https://jvet-experts.org/doc_end_user/current_document.php?id=15157) Non-EE2: Subblock-based spatial MVP [Z. Deng, K. Zhang, N. Zhang, L. Zhao, Y. Wang, W. Yin, M. Salehifar, L. Zhang (Bytedance)]

This contribution proposes a method of subblock-based spatial MVP, in which the subblock motion field of the current CU can be inherited from the motion of spatial neighbor blocks. The impact on coding efficiency and runtime of the proposed method over ECM-15.0 are reported for {Y, U, V, EncT, DecT} as below:

RA: { -0.01%, -0.02%, 0.02%, 100.1%, 100.0% }

LDB: { -0.10%, -0.21%, -0.05%, 100.2%, 101.1% }

Only spatially adjacent neighbors are considered as additional candidates.

It was commented that other candidates could be used, however those might be more complicated than the proposed method which is simple and straightforward extension of SBTmVP

Reasonable tradeoff, good gain for LDB. Several experts supported investigation in EE. As a variant, it was suggested also to test using a zero MV if one of the neighbors is intra coded.

[JVET-AK0192](https://jvet-experts.org/doc_end_user/current_document.php?id=15163) Non-EE2: On Sample-Based BDOF [M. Salehifar, Y. He, K. Zhang, L. Zhao, Y. Wang, N. Zhang, Z. Deng, W. Yin, L. Zhang (Bytedance)]

In this contribution, samples at half pixel position offset respect to the original sample are derived and used for gradient calculation in the sample-based BDOF, instead of using samples at full pixel position offset.

Simulation results on top of the ECM-15.0 are reported as:

RA: -0.04%, 0.02%, 0.03%, EncT: 102.1%, DecT: 102.2%

The process first applies BDOF to generate half sample positions, and then uses them to generate the final samples by a second BDOF.

The idea is asserted to be interesting, but in its current version overly complex. It is suggested to study alternative approaches (not in EE) to either perform everything in one BDOF step, or generating half-position samples in an alternative way.

[JVET-AK0199](https://jvet-experts.org/doc_end_user/current_document.php?id=15186) Non-EE2: MV refinement for TMVP [Z. Zhang, J.-L. Lin, Y. Zhang, H. Huang, V. Seregin, M. Karczewicz (Qualcomm)]

In this contribution, it is proposed to apply MV refinement for TMVP. The MV refinement is performed for a reference picture using collocated block as a template and the scaled to the reference picture collocated MV as a starting point, wherein, the scaling factor is (reference picture POC – collocated picture POC) / (collocated reference picture POC – collocated picture POC). A refined MV is determined by checking SATD cost between a reference template in the reference picture and the collocated block in the collocated picture. The refined MV which has the minimum SATD cost is scaled to derive TMVP MV for the current block, wherein, the scaling factor is (reference picture POC – current picture POC) / (reference picture POC – collocated picture POC). When a TMVP candidate has two MVs, it derives each refined MV independently. The proposed method was implemented on top of ECM-15.0 and the test results are summarized as follows.

RA: -0.04% (Y), 0.00% (U), -0.03% (V), 100.2% (EncT), 101.8% (DecT), 100.4% (EncVmPeak), 101.4% (DecVmPeak)

LDB: -0.24% (Y), -0.07% (U), -0.20% (V), 100.1% (EncT), 102.9% (DecT), 100.9% (EncVmPeak), 100.9% (DecVmPeak)

Refinement is always on, also used in subblocks.

Interesting gain (especially in LB), but also increased decoder runtime/complexity due to additional MV search in the colocated picture (which is not the reference picture to which the original MV points)

Several experts found it interesting to be investigated in EE. Improved strategies for decreasing the decoder run time should be investigated.

[JVET-AK0213](https://jvet-experts.org/doc_end_user/current_document.php?id=15202) AHG12: Extended BDOF usage for MV refinement [R. Yu, H. Huang, V. Seregin, M. Karczewicz (Qualcomm)]

In ECM-15.0, the bi-directional optical flow (BDOF) tool has been extended to be applicable for bi-predicted inter blocks with reference pictures from the same direction. The extended BDOF is used to adjust luma prediction sample values for inter blocks and refine motion vectors for affine blocks.

In this contribution, it is proposed to extend BDOF usage to also refine motion vectors for regular inter merge blocks with reference pictures from the same direction. The proposed modification is reported to be implemented on top of ECM-15.0. The BD-rate PSNR impact for the luma component under ECM common test configurations is reported to be -0.01% for RA and -0.07% for LDB. It is asserted that the modification is a straightforward extension, and it is proposed to investigate the modification in the next round of EE2.

RA encT 100.2% decT 100.4%, LB encT 100.6% decT 101.6%.

Gain in RA low, as this case rarely happens. Interesting gain in LDB

Straightforward to use BDOF in case of same reference direction also for regular merge. Interest was expressed for study in EE.

[JVET-AK0225](https://jvet-experts.org/doc_end_user/current_document.php?id=15214) Non-EE2: Extension on spatial and temporal merge candidates [J.-L. Lin, P.-H. Lin, Z. Zhang, V. Seregin, M. Karczewicz (Qualcomm)] [late]

This contribution proposes to use additional spatial blocks and temporal blocks along the CU boundary to derive spatial and temporal merge candidates. The proposed method was implemented on top of ECM-15.0 and the test results are as follows:

RA\*: -0.02% (Y), -0.06% (U), -0.04% (V), 100.3% (EncT), 100.8% (DecT), 100.0% (EncVmPeak), 100.0% (DecVmPeak)

LB\*: -0.09% (Y), -0.01% (U), 0.13% (V), 100.7% (EncT), 101.0% (DecT), 100.1% (EncVmPeak), 100.0% (DecVmPeak)

How many additional candidates? Up to 4 spatial

Decoding time increase due to more TM through extended list.

Interest expressed to study in EE also in combination with JVET-AK0096.

[JVET-AK0254](https://jvet-experts.org/doc_end_user/current_document.php?id=15243) Case study of candidates in AMVP candidate list [S. Park, H. Jeong, B. Jeon (SKKU)] [late]

This study investigates effectiveness of including zero MV in the AMVP candidate list. For this purpose, an oracle case is experimented by expanding the candidate list with adding all possible MVPs, including zero MV, without worrying about signaling cost of *mvp\_idx*. It is to see how well the current list construction process works. Comparison is made between the optimal MVP determined by a simplified RD calculation in the oracle case and that selected from the current AMVP candidate list. Experimental results, conducted with the first 65 frames on Class B, C, and D sequences from ECM-11.0, reveal that the conventional method matches the RD-optimal *mvp\_idx* in the oracle case in only 33% of the cases. Furthermore, a comparison of MVP types indicates that the conventional method has lower selection rate of zero MV compared to the oracle case. One least computation-intensive but yet effective modification of the AMVP candidate list construction is to incorporate zero MV into the AMVP candidate list.

It was suggested that a similar study of “oracle case” for regular merge list would be interesting.

#### GPM (2)

Contributions in this area were discussed at 1130–1150 on Thursday 16 Jan. 2025 (chaired by JRO).

[JVET-AK0103](https://jvet-experts.org/doc_end_user/current_document.php?id=15074) Non-EE2: Template matching based jointly reordering for GPM split modes and partition indexes [C. Ma, X. Xiu, X. Wang (Kwai)]

In this contribution, it is proposed to jointly reorder the split modes and motion information of the GPM modes, including the regular GPM, the GPM TM and the GPM MMVD, with template matching. The overall performance of the proposed method on ECM-15.0 is {-0.09%, -0.18%, -0.10%, 100.8%, 100.1%} in RA and {-0.06%, -0.08%, 0.09%, 100.7%, 100.4%} in LDB.

A flag is signalling whether the existing method or the proposed method is used.

Number of TM operations is not increased.

Reasonable gain and tradeoff. Interest expressed to investigate in EE. Results should also be reported when the method is always replacing the existing approach (without flag)

[JVET-AK0133](https://jvet-experts.org/doc_end_user/current_document.php?id=15104) Non-EE2: Improvements on the BV-based prediction in SGPM [J. Huo, Y. Fei, L. Wang, Y. Ma, F. Yang (Xidian Univ.)]

In ECM 15.0, SGPM employs BV-based prediction in which BV is from neighboring blocks. Besides BV information, IBC-RRIBC flipType is also an important parameter from neighboring blocks. In the current SGPM, the IBC-RRIBC flipType is not inherited which hamper efficient BV prediction. This contribution proposes to inherit the IBC-RRIBC flipType of neighboring blocks when SGPM from a merge candidate, and a flip-aware BV adjustment approach is applied to refine the BV candidate in SGPM. With the modification, the operation of SPGM is aligned with the operation in IBC/DBV/BVG-CCCM mode.

Since IBC-RRIBC is enabled for Class F and TGM, it is reported that the coding performance of ECM 15.0 is as follows:

For AI configuration:

Class F: -0.04 %, -0.07 %, 0.02 %, with xx% EncT, xx% DecT.

Class TGM: -0.08%, -0.04 %, -0.11 %, with xx% EncT, xx% DecT.

For RA configuration:

Class F: -0.11 %, -0.30 %, -0.27 %, with xx% EncT, xx% DecT.

Class TGM: -0.02%, 0.05 %, -0.06 %, with xx% % EncT, xx% DecT.

According to report of cross-checker, encoding/decoding time is not affected.

Cross-checking party expresses interest to investigate this in EE.

It is commented that small improvements of specific SCC tools are not of prior importance, this proposal seems to be useful if it does not increase runtime.

Investigate in EE.

[JVET-AK0253](https://jvet-experts.org/doc_end_user/current_document.php?id=15242) Crosscheck of JVET-AK0133 (Non-EE2: Improvements on the BV-based prediction in SGPM) [X. Li (Alibaba)] [late]

#### In-Loop Filters (7)

Contributions in this area were discussed at 1150–1300 on Thursday 16 Jan. 2025 (chaired by JRO).

[JVET-AK0113](https://jvet-experts.org/doc_end_user/current_document.php?id=15084) AHG12: On CCALF [L. Xu, Y. Yu, H. Yu, N. Song, D. Wang (OPPO)]

This document proposes two modifications to the CCALF in the current ECM-15.0. Firstly, The CCALF parameters are proposed to be signaled in a separate APS type different from that for ALF parameters. Secondly, multiple sets of CCALF parameters can be used for one slice. The proposed method is implemented on top of ECM-15.0 and the simulation results are summarized as below:

RA: 0.00%/-0.13%/-0.18%

LDB\*: -*0.08%/0.31%/-0.09%*

It was commented that optimization of HLS structures for a small benefit in saving bits is not of high importance at this stage of exploration. Naturally, highest gain is in class E which has the lowest bit rate.

[JVET-AK0308](https://jvet-experts.org/doc_end_user/current_document.php?id=15297) Crosscheck of JVET-AK0113 (AHG12: On CCALF) [W. Yin (Bytedance)] [late]

[JVET-AK0117](https://jvet-experts.org/doc_end_user/current_document.php?id=15088) Non-EE2: Neural network-based loop filtering for ECM [Y. Du, A. Li, J. Liu, C. Zhu, L. Luo, H. Guo (UESTC), Y. Huo, Y. Liu (Transsion)]

Different complexity in-loop filtering architectures have been proposed for Neural Network-based Video Coding (NNVC), such as the High-performance Operation Point (HOP, 460kMAC/Pixel) architecture in JVET-AD0380 [1], and the Low-performance Operation Point (LOP, 17kMAC/Pixel) and Very Low-performance Operation Point (VLOP, 5kMAC/Pixel) in JVET-AE0281 [2] and JVET-AH0051 [3]. Additionally, a proposal in AJ0210 [4] integrates neural network-based in-loop filtering with the ECM encoder.

In this contribution, we propose a neural network-based loop filtering method that enhances compressed images at multiple quality levels for ECM: low-QP, medium-low-QP, medium-high-QP, and high-QP. We utilize the DIV2K dataset along with the ECM encoder to encode images at 10 different QP values, resulting in four distinct quality levels of compressed reconstructed images. Subsequently, we train four separate loop filtering enhancement networks to perform filtering and enhancement on these compressed images.

The proposed method was tested using ECM CTC, which is applied during the loop-filtering phase to enhance compressed videos. The results reportedly provide the following BD-rate PSNR {Y, Cb, Cr} gain compared to the ECM15.0 anchor:

AI { xx.xx%, 0.00%, 0.00%, xxx%, xxx%},

It is proposed that this method be further studied in the next round of to EE2.

It was requested that the presentation deck is to be provided, as it has more information than the word doc.

Incomplete results – benefit cannot be assessed.

Seems to be in a similar range of complexity as LOP

Only reconstructed picture and QP is input. Four models for different QP ranges.

As chroma is not filtered, and no control information is sent, chroma BD rate is unchanged, as in that case it is a post filter.

Further study recommended.

[JVET-AK0132](https://jvet-experts.org/doc_end_user/current_document.php?id=15103) AHG12: Boundary strength adjustment for DBV mode [J. Huo, J. Fan, Y. Fei, J. Liu, Y. Ma, F. Yang (Xidian Univ.)]

Deblocking can provide a significant subjective benefit. To further improve objective benefits and subjective quality, this contribution proposes to adjust the boundary strength parameter for DBV mode in deblocking. The proposed method provides an objective benefit in ECM with deblocking in RDO. The simulation results on top of ECM-15.0 are as follows {for Y, U, V, EncT, DecT}:

LDB\* {0.00%, -0.40%, 0.04%, xx.x%, xx.x%}

Incomplete results in RA, no relevant change in LB – benefit cannot be assessed.

It was recommended to proponents that in case of deblocking subjective quality is more important.

[JVET-AK0252](https://jvet-experts.org/doc_end_user/current_document.php?id=15241) Crosscheck of JVET-AK0132 (AHG12: Boundary strength adjustment for DBV mode) [X. Li (Alibaba)] [late]

[JVET-AK0200](https://jvet-experts.org/doc_end_user/current_document.php?id=15187) Non-EE2: CCSAO with reused CTU control and extended edge classifier [C.-W. Kuo, X. Xiu, X. Wang (Kwai)]

In this contribution, it is proposed to store historical CCSAO CTU control information for use in future frames. Compared to ECM-15.0 anchors, the BD-rate impact is summarized as below (incomplete 5 points copied from anchor).

Y U V EncT DecT:

RA\* -0.03% -0.08% -0.08% 99.6% 99.2%

LB\* -0.06% -0.20% 0.06% 99.8% 99.2%

Only 3 points of large resolution classes are missing – likely that final results will be identical.

On/off flags and reuse of classifier index/offset parameters are stored for each CTU in each reference frame. Reuse of colocated CTU info.

How many bit per CTU? 2 bit according to proponents.

Several experts expressed interest to investigate this in EE

[JVET-AK0224](https://jvet-experts.org/doc_end_user/current_document.php?id=15213) Non-EE2: Cross-Chroma Adaptive Loop-Filter [W. Yin, K. Zhang, H. Liu, Y. Wang, Z. Deng, N. Zhang, L. Zhao, M. Salehifar, L. Zhang (Bytedance)]

In this contribution, a method of cross-chroma adaptive loop-filter is proposed. Specifically, the Cb reconstruction samples are used to filter Cr samples, while the Cr reconstruction samples are similarly used to filter Cb samples. The proposed method is applied individually to Chroma-ALF (Test #1) and CC-ALF (Test #2) for evaluation.

On top of ECM-15.0, simulation results of the proposed method are reported as below:

Test #1:

AI: 0.01%, -0.19%, -0.19%, 100.1%, 99.9%.

RA: 0.00%, -0.28%, -0.29%, 100.2%, 100.2%.

LB: 0.01%, -0.78%, -0.81%, 100.3%, 99.9%.

Test #2:

AI: 0.01%, -0.13%, -0.19%, 100.0%, 100.1%.

RA: 0.01%, -0.23%, -0.48%, 99.9%, 100.0%.

LB: -0.01%, -0.81%, -0.99%, 99.8%, 100.2%.

The samples from the other chroma component before deblocking are used as input.

Interesting gain in chroma. Several experts supported investigation in EE.

Tests 1 and 2 should be investigated separately and in combination, to see if feeding the additional information from one chroma component into the other is sufficient in one of the stages.

[JVET-AK0237](https://jvet-experts.org/doc_end_user/current_document.php?id=15226) Non-EE2: CCALF with dual-chroma inputs [C. Ma, X. Xiu, C.-W. Kuo, X. Wang (Kwai)]

In ECM15.0, for one chroma component, CCALF takes luma reconstruction signal, luma residual signal and the corresponding chroma reconstruction signal as input. This contribution proposes to additionally utilize the other chroma reconstruction signal as the input to the CCALF of one chroma component. On top of ECM-15.0, simulation results are reported as below:

3x3 cross shape:

AI : {0.01%, -0.15%, -0.20%, 99.6%, 98.3%},

RA : {0.00%, -0.20%, -0.33%, 99.4%, 98.9%},

LDB : {-0.03%, -0.71%, -0.64%, 98.8%, 98.6%};

5x5 cross shape:

AI : {0.01%, -0.15%, -0.25%, 99.4%, 98.2%},

RA : {0.00%, -0.25%, -0.47%, 99.8%, 99.3%},

LDB : {0.00%, -0.69%, -0.95%, 100.0%, 100.0%};

Interesting gain in chroma. Several experts supported investigation in EE.

Also combination with JVET-AK0224 should be tested. It should be considered if a relation exists with the continued EE2-4.9.

[JVET-AK0327](https://jvet-experts.org/doc_end_user/current_document.php?id=15316) AHG12: On adaptive clipping [R. G. Youvalari, M. Abdoli, A. Tissier, F. Plowman (Xiaomi)] [late]

Reviewed Sat. 18 Jan. 1500

This contribution proposes to extend the usage of adaptive clipping of luma samples of ECM to the intermediate luma samples used for CC-ALF. The BD-Rate impact of the proposed for different configurations are as below:

AI: { 0.00%, 0.00%, 0.00%, 100.2%, 100.3%}

RA: { 0.00%, -0.0x%, -0.0x%, xxx.x%, xxx.x%}

*LB: {-0.04%, -0.14%, 0.02%, 100.2%, 100.5%}\**

Not relevant for EE, as this is not a coding tool to be understood or more carefully checked. Contribution came very late, incomplete results, no cross-check.

Further study. Cross-check to be provided at this meeting, or a future meeting.

[JVET-AK0334](https://jvet-experts.org/doc_end_user/current_document.php?id=15323) Non-EE2: Non-linear clipping for CCALF [N. Song, L. Xu, Y. Yu, H. Yu, D. Wang (OPPO)] [late]

This contribution proposes to extend the usage of adaptive clipping of ALF to the intermediate input value used for CC-ALF. The BD-Rate impact of the proposed for different configurations are as below:

AI: 0.01%/ -0.19%/ -0.20%

RA: 0.xx%/0.xx%/0.xx%

LDB.\*: *0.01%/-0.67%/-0.45%*

Was presented Mon. 20 Jan. 1400 (chaired by JRO)

Some interest was expressed to investigate this in EE, but considering that there is loss in luma in AI (though might be compensated by chroma gain), and incomplete results on RA and LB, no action can be taken at this meeting.

Proponents are suggested to perform further study (outside of EE), and potentially communicate with proponents of JVET-AK0327 to identify why results are different (which can currently only be stated for the AI results).

[JVET-AK0336](https://jvet-experts.org/doc_end_user/current_document.php?id=15325) Crosscheck of JVET-AK0334 (Non-EE2: Non-linear clipping for CCALF) [Z. Dai (Alibaba)] [late]

#### Entropy coding, transforms, quantization, and transform coefficient coding (4)

Contributions in this area were discussed at 1420–1610 on Thursday 17 Jan. 2025 (chaired by JRO).

[JVET-AK0135](https://jvet-experts.org/doc_end_user/current_document.php?id=15106) AhG12: CABAC contexts retraining [F. Galpin (InterDigital)] [late]

This contribution proposes to update context initialization parameters for all slice types. All CABAC parameters (initial probabilities parameters, window sizes, adaptive weights and rate offsets) are retrained and updated. It is reported that on top of ECM-15.0, the overall coding performance impact for {Y, U, V} is {-0.05% -0.06% 0.05%} {-0.07% -0.01% -0.06%} {-0.02% -0.35% -0.33%} in AI, RA and LDB configurations respectively.

Unlike previous attempt of retraining, the new results have gain also in LDB. Training was done with CTC sequences and QP points plus classes TGM/F/D.

Training was performed using the script from ECM.

Decision: Adopt JVET-AK0135

[JVET-AK0328](https://jvet-experts.org/doc_end_user/current_document.php?id=15317) Crosscheck of JVET-AK0135 (AhG12: CABAC contexts retraining) [M. Blestel (Ofinno)] [late]

[JVET-AK0214](https://jvet-experts.org/doc_end_user/current_document.php?id=15203) Non-EE2: IntraNN NSPT set [G. Verba, M. Coban, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes to use NSPT kernels from the second set of NSPT kernels for IntraNN-coded blocks and to increase the number of non-zero coefficients for several block shapes. The method was implemented on top of ECM-15.0 and experimental results following CTC are reported as follows:

AI: -0.13% (Y) -0.12% (U) -0.08% (V) 101.1% (EncT) 100.4% (DecT)

RA:

It was asked why not using set 3? Answer: Set 3 is for inter, set 2 for non-directional intra modes, and gave better results in a preliminary experiment.

The aspect of increasing number of non-zero coefficients to all cases (also other modes and other sets of kernels, which were retrained) should be investigated separately.

According to report of proponents, retraining was performed using BVI and DIV2K (same sets as used for NSPT split)

Investigate in EE.

Following aspects should be tested separately:

* Using the second set instead of first set for IntraNN without any other change
* Retraining NSPT without increasing number of non-zero coefficients
* Retraining NSPT with increased number of non-zero coefficients

It would be desirable to share the scripts used for training with cross-checkers.

So far, it was never been a requirement to provide training cross-check in the EE2 exploration, and it does not seem appropriate to establish such a requirement at a relatively late stage of this exploration.

It needs to be discussed whether such a requirement should be set up for a call for proposals, and which rules would apply (e.g. size of the model / number of parameters) to follow such a requirement.

[JVET-AK0215](https://jvet-experts.org/doc_end_user/current_document.php?id=15204) Non-EE2: MTS with reduced memory [B. Ray, V. Seregin, M. Karczewicz (Qualcomm)]

In this contribution, max Intra MTS size syntax is added to SPS, in the test it is set to 64 to reduce the memory requirement.

Simulation results on top of the ECM-15.0 are reported as follows:

AI: 0.01%, -0.04%, 0.07%, EncT: 99.1%, DecT: 99.8%

The proposal has two aspects:

* Implementing two flags for turning off MTS extension, and controlling the max Intra MTS size (this was requested by AHG7 for its investigations)
* Proposing to limit max Intra MTS size to 64 (very small loss, reduction of runtime and memory)

The first aspect is agreed.

For the second aspect, a comment is made that in terms of the ECM exploration, the benefit is minor.

It was also pointed out that ECM software has a bug when setting max MTS TU size to 64. AHG6 needs to work on a solution to fix that. It was requested to issue a bug report.

Decision: Adopt JVET-AK0215 first aspect (two flags), non-CTC

[JVET-AK0279](https://jvet-experts.org/doc_end_user/current_document.php?id=15268) Non-EE2: On coefficient level binarization in Transform Skip [M. Abdoli, R. G. Youvalari, F. Plowman, A. Tissier (Xiaomi)] [late]

Residual coefficient coding in Transform Skip (TS) is associated with a TU-level context bin budget that determines whether the typical binarization (i.e. sigFlag, gt1, parity flag, gtX) is used along with CABAC contexts, or alternatively, the Rice mode with no binarization and no CABAC context coding is used. This contribution proposes to modify the binarization of the Rice coding mode in the case the context bin budget runs out. In the new binarization of the Rice coding mode, a significance flag is coded for each TS coefficient using a CABAC context. The proposed modification impacts around 1% of Coefficient Groups (CG) in QP22 of class F and TGM and the experiments in ECM-15.0 under CTC show the following results (\*):

*All-Intra*:

Natural content: { 0.00%, -0.03%, -0.01%, 101.0%%, 101.9%}

Class F: {-0.15%, -0.13%, -0.33%, 100.2%, 102.0 %}

Class TGM: {-0.47%, -0.48%, -0.50%, 100.0 %, 99.7 %}

*Random-Access*:

Class F: {-0.22%, -0.09%, -0.17%, 101.6 %, 101.6 %}

Class TGM: {-0.32%, -0.33%, -0.27%, 100.6 %, 99.3 %}

*Low-delay B*:

Class F: {-0.25%, -0.26%, -1.17%, xxx.x %, xxx.x %}

Class TGM: {-0.06%, 0.00%, -0.07%, xxx.x %, xxx.x %}

(\*) Among all reported run-times, only EncT of All-Intra in Class F and TGM are reliable.

Several experts expressed interest to study this as an alternative method for sticking to a maximum budget of context coded bins, even though in other elements of ECM, such a constraint is not observed.

It is commented that having a different method of budget limitation for TS and regular coding would not be desirable.

It was asked how much better the results would be when the budget limitation would be completely disabled. According to proponent, it would not be much better.

It was suggested that such a study should also be performed in a low QP range, where the case of budget limitation occurs more frequently, and also in regular transform coefficient coding, where it would also occur with camera captured content in low QP.

It was further commented that using the proposed method in regular coefficient coding could be problematic due to interleaving of context coded and bypass coded bins.

Study in EE, CTC and additional QP points 2, 7, 12 and 17 for AI with every 64th frame, only for TS TC coding. In a follow-up discussion on Monday 20 Jan. it was suggested to also perform a comparison in the low QP range, studying a case without any constraint for the budget of context coded bins.

#### Other (3)

Contributions in this area were discussed at 1610–1630 on Thursday 16 Jan. 2025 (chaired by JRO), and 1520–1600 on Saturday 18 Jan. 2025 (chaired by JRO).

[JVET-AK0171](https://jvet-experts.org/doc_end_user/current_document.php?id=15142) Non-EE2: Picture-level mirroring and rotation [D. Mieloch, M. Lorkiewicz, J. Stankowski (PUT)]

The document presents the results of experiments on improving video coding efficiency by picture-level mirroring and rotation. In order to evaluate the maximum possible gain from such transformations, all test sequences were encoded multiple times, for each possible mirroring, rotation, and their combinations. Then, the selection of the best scenario was done independently for each frame, leading to optimal mode selection and the best achievable BD-rate improvement from mirroring and rotation transformations. The results were obtained using ECM-15.0 software, and the maximum average BD‑rate change was ‑0.46%, 0.49%, and 0.61% for Y, U, and V components in AI configuration.

It is commented that mirroring and rotation might change the chroma sampling position, such that tools that take it into consideration might need to be modified or a control parameter might need to be changed such that the tool would still operate as intended. This might explain the increase of chroma BD rate. CCLM, CCALF and CCSAO might be affected by this.

It was suggested to investigate whether the decrease of BD rate in luma is coming with a decrease in rate – if that was the case, this would be an indication that the chroma PSNR is worse, except for vertical mirroring.

In the reference picture buffer, it would be needed to be brought back to original orientation, if the pictures and other information such as MV would be used for prediction.

If rotation is performed, it would mean that a new video sequence needs to be started, as width/height are changing.

[JVET-AK0173](https://jvet-experts.org/doc_end_user/current_document.php?id=15144) Non-EE2: Adaptive picture-level vertical mirroring [D. Mieloch, M. Lorkiewicz, A. Dziembowski, J. Stankowski, D. Karwowski (PUT)]

The document proposes a method for improving video coding by introducing adaptive picture-level vertical mirroring. The method involves an automatic frame pre-processing step for determining whether the mirroring should be applied without the necessity to perform a brute force search or apply additional RDO checks. The proposed method was implemented in the ECM-15.0 software and achieved an average BD-rate change of -0.23%, -0.26%, and -0.44% for Y, U, and V components in AI configuration, with no change in a decoder runtime.

The approach is based on a pre-analysis of dominant diagonal edge direction.

In JVET-AK0171, the gain of the vertical mirroring is -0.46% luma, but encoding time doubles due to two-pass encoding.

JVET-AA0142 and JVET-AB0165 proposed a similar approach, also addressing the luma/chroma correspondence.

Investigate in EE both for AI and RA. Also report how often the orientation is switched per sequence (which might be showing whether the previous picture’s analysis could be used), and whether there are large variations in gain between sequences. Might not be useful for low delay due to additional delay, also additional memory is required.

[JVET-AK0219](https://jvet-experts.org/doc_end_user/current_document.php?id=15208) Non-EE2: Chroma partition prediction in separate tree condition [P.-H. Lin, J.-L. Lin, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes a method of chroma partition prediction for separate tree condition. In ECM, separate tree is applied on I-slice to allow luma CTU and chroma CTU use different partition tree structure. In the proposed method, the split of a chroma block is predicted by the decoded luma region.

On top of ECM-15.0 [1], simulation results are reported as follows:

AI {-0.03% Y, -0.89% U, -0.87% V, 102.0% EncT, 99.9% DecT }

The proposed method contains two aspects.

The first aspect proposes to set the partition parameter MaxTTChromaISlice from 32 to 64, which is aligned with MaxBTChromaISlice.

The second aspect proposes to disable some available splits of a chroma block by analyzing the partition of decoded luma region. For each chroma CTU, the 2 variables average luma BT depth and average luma QT depth are calculated from all the blocks in the collocated luma CTU. TT split of a chroma block is disallowed when the two following conditions are met:

* The current BT depth of a chroma block is larger than the average luma BT depth
* The current QT depth + 1 is smaller than the average luma QT depth

RA not finalized yet; only used in I slices, not for local dual tree.

Encoder run time increase due to the first aspect

Second aspect would give less gain without the first aspect

Investigate in EE. Following items should be reported:

* First aspect standalone (non-normative)
* Second aspect standalone (normative with syntax change, without increasing MaxTTChromaISlice)
* Both aspects together as proposed
* Both aspects together, but second aspect in a non-normative variant (encoder would speed up by deriving the chroma partitioning)

### CTC for EE2/ECM and general ECM improvements (1)

Contributions in this area were discussed at XXXX–XXXX on XXday XX Jan. 2025 (chaired by XXX).

See also discussion under 4.7 and 4.15.

[JVET-AK0337](https://jvet-experts.org/doc_end_user/current_document.php?id=15326) Report on NN-ILF interface for ECM [D. Rusanovskyy, Y. Li, M. Karczewicz (Qualcomm), F. Galpin (InterDigital), J. Ström (Ericsson)+more] [late]

Contribution JVET-AK0183 reported results of EE2-4.8 Test B, where an initial implementation of the NN-ILF interface developed in NNVC was integrated into ECM15 and used for luma filtering. This interface was tested with NNVC trained VLOP1 filter (results in JVET-AK0183), and NNVC trained LOP2 and HOP3 filters (results in JVET-AK0184).

EE2-4.8 Test B, improved to resolve RA mismatch, was further evaluated on-site and off-site of the JVET-AK meeting, and its performance with VLOP1 filter was jointly assessed by three multiple participating laboratories with matching results achieved for currently completed QP points.

As a result of evaluation, it is concluded that initial NN-ILF interface implementation proposed in JVET-AK0183 is adequate for testing of luma filtering with the NNVC trained ILF models (VLOP1, LOP2, HOP3), and can further be improved to enable testing of the newest NNVC models, e.g. LOP4 and HOP5. The bug causing crash in RA had been resolved, but only filtering luma.

Implementation is for 128x128 block size. For 256x256 CTU, it would be operated in 4 sub-blocks.

Several experts expressed opinion that this is a very valuable step forward.

In the EE2 discussion, it had been suggested considering the option of including such an interface (when the bugs found in EE2-4.8 were resolved, and also support for chroma part would be available) into ECM as non-default option for further study.

During the discussion, some concern was raised about including SADL library in ECM, which would be a consequence. Experts familiar with SADL commented that this was not a big problem. Excluding SADL would mean that filters developed in EE1 could not easily be tested together with EE2 tools, which is the purpose of this interface.

Further study is recommended to also support chroma part of NNVC loop filters, and also test the functionality of the interface with LOP and HOP, such that it could be included in ECM by the next meeting.

# High-level syntax (HLS) and related proposals (XX)

## AHG9: Aspects on SEI messages in VSEIv4 and related (35)

### General comments (1)

Contributions in this area were discussed at 1810–1830 on Tuesday 14 Jan. 2025 (chaired by Y.-K. Wang).

[JVET-AK0160](https://jvet-experts.org/doc_end_user/current_document.php?id=15131) AHG9: On VVC interface of object mask information and annotated regions SEI messages [J. Chen, B. Chen, Y. Ye (Alibaba), M. M. Hannuksela (Nokia)]

Chaired by Ye-Kui Wang on 14 January 2025 at 18:10.

This contribution proposes the following items:

1. It is proposed to add the interface of OMI SEI message in VVC spec., which is currently missing.

Two alternative options are proposed. The second option needs change to VSEI and VVC text.

1. It is proposed to remove the interface variables ConfWinLeftOffset and ConfWinTopOffset from the OMI and AR SEI messages and clarify that the coordinates indicated by these SEI messages are relative to the top-left luma sample of the cropped decode picture. The changes proposed in this item are asserted to be editorial.

This part relates to option 2.

Regarding AR SEI message JVET-AE0054 is a previous related proposal but was only for AR SEI. That was marked for further study.

It was agreed that we need interface text (first part).

A comment was made that the change related to including conformance cropping (using cropped decoded picture) makes things clear and consistent (e.g. for AR) and so option 2 is preferred.

It was asked if the change to AR part proposed is editorial and it was confirmed by author to be the case.

Two non-proponents supported option 2.

Decision: Adopt option 2.

Regarding AR SEI message JVET-AE0054 is a previous related proposal but was only for AR SEI. That was marked for further study.

It was agreed that we need interface text (first part).

A comment was made that the change related to including conformance cropping (using cropped decoded picture) makes things clear and consistent (e.g. for AR) and so option 2 is preferred.

It was asked if the change to AR part proposed is editorial and it was confirmed by author to be the case.

Two non-proponents supported option 2.

Decision: Adopt option 2.

### SEI processing order and processing order nesting SEI message aspects (5)

Contributions in this area were discussed at 1530–1915 on Tuesday 14 Jan. 2025 (chaired by J. Boyce).

[JVET-AK0055](https://jvet-experts.org/doc_end_user/current_document.php?id=15026) AHG9: Semantics of the SPO SEI message [Y.-K. Wang, J. Xu, L. Zhang (Bytedance), K. Yang, Y. Li, Y. Xu (SJTU)]

This contribution proposes the following changes to the semantics of the SPO SEI message:

1. Add a constraint to require that the value of po\_sei\_importance\_flag[ i ] shall be equal to 0 when po\_sei\_payload\_type[ i ] is equal to any value in spoPropertySeiList, i.e., the i-th type of SEI message indicates a property.

Justification: It is asserted that an SEI message type that only indicates a property should never determine that a processing chain must be ignored when the post processor does not recognize the SEI message type.

A participant questioned if this change was necessary, and it doesn’t seem harmful to let the encoder have this flexibility.

A participant noted that there are two flags to which this might apply.

No action.

1. Add a constraint to require that a processing chain shall not include both an ERP SEI message type and a GCMP SEI message type.

Justification: It is asserted that it does not make sense for a processing chain to include an RWP SEI message but not an ERP or GCMP SEI message.

A participant questioned if there was a restriction against having both ERP and GCMP SEI messages in the same scope. If we do have such a constraint, it would be good to be consistent.

It was questioned whether it would be more appropriate for such a constraint to be present in in the video coding spec (VVC, HEVC, AVC) rather than in VSEI. The proponent suggests that if a video coding spec refers to both ERP and GCMP, the constraint could be applied in VSEI.

Discussed on 18 January 2025 at 14:30 PM. Chaired by S. Deshpande.

The proponent reported that he has confirmed that semantic constraint like this already exists and so having it in SPO also makes sense. Support was expressed to include this constraint.

It was asked if additional projection type is added in future how it should be handled. It was answered that similar approach (and new constraint) can be added at that time or it can be added now in a generic way to make it future proof.

Decision: Adopt #2 as noted above to make it future proof.

1. Add a constraint to require that, when a processing chain includes an RWP SEI message type, it shall also include either an ERP or GCMP SEI message type.

Justification: It is asserted that it does not make sense for a processing chain to include an RWP SEI message type but not either an ERP or GCMP SEI message type.

It was suggested that the wording should not preclude use of a different projection format defined in the future with RWP. It is possible to change constraints in the future if/when a different projection format is defined, but we need to remember to do that.

It was suggested that it would be desirable to impose such constraints in a similar way to how the constraints are imposed in the absence of SPO.

It was noted that sometimes an SEI message has the same name in more than one video coding spec but has a different definition.

Decision: Adopt #3

1. Due to that certain order in omnidirectional processing at the decoder side needs to be followed, the following constraints are proposed:
   1. When a processing chain includes both an RWP SEI message type and an ERP SEI message type, the indicated processing order for the RWP SEI message type shall be later than that for the ERP SEI message type.
   2. When a processing chain includes both an RWP SEI message type and a GCMP SEI message type, the indicated processing order for the RWP SEI message type shall be later than that for the GCMP SEI message type.
   3. When a processing chain includes both an ERP SEI message type and a frame packing arrangement (FPA) SEI message type, the indicated processing order for the ERP SEI message type shall be later than that for the FPA SEI message type.
   4. When a processing chain includes both a GCMP SEI message type and an FPA SEI message type, the indicated processing order for the GCMP SEI message type shall be later than that for the FPA SEI message type.
   5. When a processing chain includes an ERP SEI message type, an FPA SEI message type, and an RWP SEI message type, the indicated processing order for the ERP SEI message type shall be later than that for the FPA SEI message type, and the indicated processing order for the FPA SEI message type shall be later than that for the RWP SEI message type.
   6. When a processing chain includes a GCMP SEI message type, an FPA SEI message type, and an RWP SEI message type, the indicated processing order for the GCMP SEI message type shall be later than that for the FPA SEI message type, and the indicated processing order for the FPA SEI message type shall be later than that for the RWP SEI message type.

It was suggested that the abstract wording is not consistent with the language in the spec attachment.

Decision: Adopt #4 as described in the spec attachment. A new version of this contribution to be uploaded which fixes the abstract to be consistent with the spec text.

1. Many SEI messages include a persistence cancel flag in the syntax. For example, the NNPFA SEI message syntax includes the nnpfa\_cancel\_flag, which, when equal to 1, cancels the persistence of the target NNPF established by any previous NNPFA SEI message with the same nnpfa\_target\_id. Currently, it is allowed for the SEI prefix indication of an SEI message type in the SPO SEI message to have the persistence cancel flag equal to 1. However, it is asserted that that does not make sense.

It is therefore proposed that, when an SPO SEI message includes an SEI message type and the SEI prefix indication for the SEI message type includes the persistence cancel flag, it is required that the value of the persistence cancel flag included in the SEI prefix indication for this SEI message type shall be equal to 0. Specifically, the following constraints are proposed:

The detailed proposed text changes, marked relative to JVET-AJ2006-v3, are included in an attachment to this contribution.

It was suggested that a note may be able to used in place of the amount of text, although a note may be difficult to provide correct language for.

No action.

[JVET-AK0112](https://jvet-experts.org/doc_end_user/current_document.php?id=15083) AHG9: On the SPO SEI message [Y. Gao, P. Wu, Y. Bai, S. Xie, M. Jia, W. Niu, C. Huang (ZTE)]

Chaired by S. Deshpande on 14 January 2025 at 18:25.

It is asserted that the current design of poSeiList derivation related to sub-chains in JVET-AJ2006 has the following problem:

1. When there is no SEI message in the processing chain specified by this SPO SEI message such that po\_sei\_importance\_flag[ i ] is equal to 0 and po\_sei\_processing\_degree\_flag[ i ] is equal to 1, the lists PoSeiList derived from the derivation process described in JVET-AI2006 and JVET-AJ2006 are inconsistent in some cases.
2. When there is more than one SEI message in a sub-chain such that po\_sei\_importance\_flag[ i ] is equal to 1 and po\_sei\_processing\_degree\_flag[ i ] is equal to 0, the sub chain processing result derived from JVET-AJ2006 may not be consistent with the initial intention of the content provider.

This contribution proposes the following modifications to address the above asserted problems:

Change the Interpretation of the value combination of po\_importance\_flag[ i ]=1 and po\_processing\_degree\_flag[ i ] =0 to the following:

“ The i-th SEI message type belongs to a sub-chain. when the decoding system cannot interpret or does not support the i-th SEI message type ,the derivation of PoSeiList associated with this sub-chain is terminated.”

No comments or conflicting opinions were voiced regarding the two cases which show different behavior of JVET-AI2006 and JVET-AJ2006.

It was commented that JVET-AK0165 also addresses the nested sub-chains aspect (similar to this proposal), but the proposed solution is different than that in JVET-AK0112.

It was decided to make the decision after discussing that proposal. See notes for that below.

It is asserted that the current design of poSeiList derivation related to sub-chains in JVET-AJ2006 has the following problem:

1. When there is no SEI message in the processing chain specified by this SPO SEI message such that po\_sei\_importance\_flag[ i ] is equal to 0 and po\_sei\_processing\_degree\_flag[ i ] is equal to 1, the lists PoSeiList derived from the derivation process described in JVET-AI2006 and JVET-AJ2006 are inconsistent in some cases.
2. When there is more than one SEI message in a sub-chain such that po\_sei\_importance\_flag[ i ] is equal to 1 and po\_sei\_processing\_degree\_flag[ i ] is equal to 0, the sub chain processing result derived from JVET-AJ2006 may not be consistent with the initial intention of the content provider.

This contribution proposes the following modifications to address the above asserted problems:

Change the Interpretation of the value combination of po\_importance\_flag[ i ]=1 and po\_processing\_degree\_flag[ i ] =0 to the following:

“ The i-th SEI message type belongs to a sub-chain. when the decoding system cannot interpret or does not support the i-th SEI message type ,the derivation of PoSeiList associated with this sub-chain is terminated.”

No comments or conflicting opinions were voiced regarding the two cases which show different behavior of JVET-AI2006 and JVET-AJ2006.

It was commented that JVET-AK0165 also addresses the nested sub-chains aspect (similar to this proposal), but the proposed solution is different than that in JVET-AK0112.

It was decided to make the decision after discussing that proposal. See notes for that below.

[JVET-AK0156](https://jvet-experts.org/doc_end_user/current_document.php?id=15127) AHG9: On SEI processing order SEI message [C. H. Demarty, E. François, F. Aumont, O. Le Meur (InterDigital)]

This contribution proposes to enhance the SEI processing order (SPO) SEI message with information related to potential optimization for energy savings purposes.

When energy consideration comes to play, such information can for example be used by a receiver to decide whether or not to apply the processing chain specified by the SPO SEI message.

It was noted that the po\_reserved\_zero\_2bits should be in the range of 0 to 3, not 0 to 4.

One possible use case is that two SPO chains are present, with one SPO chain for normal operation and another SPO chain for energy efficient operation.

A question was raised about how the encoder would know what post-processing (applied after the SPO processing chain is applied) would be applied, which would be needed to indicate energy savings.

A question was raised about example use cases for bit number 3 of energy savings can be expected when using the content for machine analysis other than reducing the sample rate via spatial or temporal resampling. Cropping was suggested as an operation.

A question was raised as to what other processing chain that a chain marked as energy savings would be relative to. The proponent suggested that it would be relative to not applying the SPO at all.

It was suggested that it is not clear how an encoder can know if energy savings would occur for post-processing following the SPO operations, and that a receiver would know that it could trust that information.

If more than one SPO chain indicates energy savings, it would be unclear what the energy savings was relative to.

A participant noted that energy management will likely become important in the future.

Further study encouraged, including providing some use case examples involving SEI messages included in VSEI or the TuC.

[JVET-AK0165](https://jvet-experts.org/doc_end_user/current_document.php?id=15136) AHG9: On SPO sub-chain signaling [T. M. Borges, Y. Sanchez, R. Skupin, C. Hellge, T. Schierl (Fraunhofer HHI)]

Chaired by S. Deshpande on 14 January 2025 at 16:40.

This contribution proposes modifications on the signaling of sub-chains from the SPO. The following options are proposed:

1. Change the interpretation of state po\_sei\_importance\_flag[ i ] = 1, po\_sei\_processing\_degree\_flag[ i ] = 1, when within a sub-chain such that the concept of critical sub-chains is created, allowing to have dependency for further processes only.
2. Constrain the usage of state po\_sei\_importance\_flag[ i ] = 0, po\_sei\_processing\_degree\_flag[ i ] = 1, when outside of a sub-chain.
3. Use, as an alternative to the two separate flags, a single syntax element in the form of an idc, to simplify the derivation of PoSeiList and the interpretation of the processing states.

Item 3 is editorial and was discussed at previous meeting.

Regarding item 1 of this proposal: It was asked what is the behaviour when sub-chain is “closed” with flags 0,1. It was answered that the change relates to closing a sub-chain with flags 1,1.

It was commented that in JVET-AK0112 for example on page 4 of JVET-AK0165, the SEI with i=0 will be processed if SEI i=1 can not be processed, but the behavior is different in the JVET-AK0165 proposed solution. It was commented by proponents that the signaling should be different to support that example.

It was asked if the JVET-AK0165 proposal does not allow “critical” sub-chain of length 1. That seemed to be the case.

Regarding item 2 of this proposal, it proposes to add a constraint. It was commented that this can be decided after the decision regarding JVET-AK0112 and item 1 of this proposal.

Proponents of JVET-AK0112 commented that it is important to support the originally discussed use cases (e.g. broadcast use case). It was commented by proponents of JVET-AK0165 that the current VSEI draft does not support nesting and their proposed approach is trying to bring it back. It was asserted by JVET-AK0165 proponents that the proposal in JVET-AK0112 cannot support sub-chain with more than 2 important stages.

It was commented that we may need to decide which of these two things are important:

-to support sub-chains of length 1 or

-to support nested sub-chains with more than 2 important stages

It was commented that initial version of JVET-AJ0128 had retained the nesting functionality, but later revisions lost that functionality. So it was requested to also consider JVET-AJ0128 v1 and the proponent of that document volunteered to update the solution in it to show changes with respect to current VSEI text if the group was interested in.

There was a general agreement to support nested functionality.

An off-line activity was planned among JVET-AK0112, JVET-AK0165 and JVET-AJ0128-v1 proponents and other interested parties. The activity will be co-ordinated by M. M. Hannuksela.

Discussed item 1 again on 18 January 2025 at 16:37 PM. See notes under JVET-AK0333.

Regarding item 2: It is asserted that 0, 1 can be used to close the sub-chain currently irrespective of if sub-chain is opened or not. There was support for the proposed approach of adding a constraint from a non-proponent. But some more time to review was requested by a participant.

Discussed item 2 and item 3 again on 18 January 2025 at 16:37 PM. See notes under JVET-AK0333.

It was confirmed that item 3 is purely editorial and is delegated to editors’ for consideration.

[JVET-AK0333](https://jvet-experts.org/doc_end_user/current_document.php?id=15322) AHG9: On SPO sub-chain signaling [M. M. Hannuksela (Nokia), Y. Sanchez, T. M. Borges (Fraunhofer HHI), Y. Gao, P. Wu (ZTE)] [late]

Chaired by S. Deshpande on 18 January 2025 at 16:30.

This document results from the off-line activity related to JVET-AK0112, JVET-AK0165, and JVET-AJ0128-v1. The document proposes an SPO sub-chain design, which is asserted to support the nested sub-chain functionality.

In the JVET HLS review of JVET-AK0112 and JVET-AK0165, it was concluded that there is general agreement to support the nested sub-chain functionality, and the following specific features were raised for consideration:

* to support sub-chains of length 1
* to support nested sub-chains with more than 2 important stages

This contribution is asserted to support the nested sub-chain functionality through open-ended sub-chains (called "critical" sub-chains in JVET-AK0165).

In the proposed specification text section, not everything is new text. The changed part was explained.

Decision: Adopt

[JVET-AK0205](https://jvet-experts.org/doc_end_user/current_document.php?id=15194) AHG9: On SEI messages in spoPropertySeiList [H. Tan, J. Lee, J. Nam, C. Kim, J. Lim, S. Kim (LGE)]

It is asserted that the design of categorizing the SEI types into the list of SeiProcessingOrderSeiList, SpoProcessSeiList, and spoPropertySeiList is not clear on several aspects such as the meaning of having property SEI messages in a certain position (i.e., processing order) within a processing chain, why property SEI messages can be the first SEI message in a processing chain (i.e., what the different between that case with simply having the SEI message not included in an SPO SEI message), and finally, the process that may be associated with property SEI message (e.g., frame packing arrangement SEI message has associated unpacking process but the SEI is not included in SpoProcessSeiList).

This contribution includes several items to clarify the asserted unclear aspects above as follows:

1. To clarify the impact of SEI message in a processing chain that is part of spoPropertySeiList, do one of the following options:
   * Option 1:
     + Add a constraint such that SEI messages that are included in spoPropertySeiList shall not be the first SEI message in any processing chain.
     + When an SEI message that is included in spoPropertySeiList is the i-th SEI in a processing chain, it indicates property of the associated picture after the invocation of post-processing processes implied by up to the (i – 1)-th SEI message in the processing chain.
   * Option 2:
     + When an SEI message that is included in spoPropertySeiList is the first SEI in a processing chain, it indicates property of the associated picture prior to the application of any post-processing process in the processing chain.
     + Otherwise (i.e., an SEI message that is included in spoPropertySeiList is the i-th SEI in a processing chain), it indicates property of the associated picture after the invocation of post-processing processes implied by up to the (i – 1)-th SEI message in the processing chain.
2. Add text that describes when the i-th SEI message in a processing chain is included in spoPropertySeiList and the SEI message may be associated with a particular process, such process is not part of processing chain.

It was suggested that Item 1 Option 2 is an expression of our current intent, which could potentially be done with an editorial change to the semantics, rather than the proposed language in the processing chain.

Discussed Item 1 after language was provided to express the intent in the semantics. The discussion was chaired by S. Deshpande on 18 January 2025 at 16:30 PM.

The revised text is provided in version 2 of the document for item 1 option 2.

Decision: agreed item 1 as revised in v2 of the document.

For item 2, it was suggested by a participant that the suggested language conflicts with existing language.

It was suggested that an associated process is difficult to describe.

It seems difficult to write a proper note.

No action on item 2.

### NNPF modifications (13)

Contributions in this area were discussed at 1915–1950 on Tuesday 14 Jan. 2025, and 0900–1310 on Wednesday 15 Jan. 2025 (chaired by J. Boyce).

[JVET-AK0053](https://jvet-experts.org/doc_end_user/current_document.php?id=15024) AHG9: NNPF interface text in VVC [Y.-K. Wang, J. Xu, L. Zhang (Bytedance), K. Yang, Y. Li, Y. Xu (SJTU)]

It is asserted that, in the current NNPF interface text in the VVC spec, some aspects are not clearly specified, including:

1. The NNPF process may also be applied to an inserted (interpolated or extrapolated) picture, in addition to be applied to a cropped decoded output (CDO) picture or a processed version of a CDO picture. However, this aspect is not clearly specified. Furthermore, in this case, it is not clear what the current picture denoted by the variable currPic is.
2. For NNPFs, an input picture may be unavailable in the list of candidate input pictures but is padded based on pictures in that list or padded to be zero-valued. However, when the list of candidate input pictures is not a list of CDO pictures, this aspect is not clearly specified.
3. The phrases like “currPic is the last picture of the bitstream in output order”, “a picture prevPic that is the last cropped decoded output picture in output order among all cropped decoded output pictures”, and “SliceQpY of the first slice of inputPic[ i ]” are vague and unclear when the candidate input picture is a processed version of a CDO picture and more unclear when the candidate input picture is an inserted picture.

This contribution proposes changes to the NNPF interface text in the VVC spec to address the above asserted issues.

The detailed proposed text changes, marked relative to JVET-AJ2005-v3, are included in an attachment to this contribution.

It was suggested by a participant that the contribution is editorial in nature as it expresses the current intent.

It was noted that previous guidance had been to localize changes when possible.

It was noted that other SEI messages have a different way of wording handling of cropped decoded pictures vs coded pictures, but that NNPF is more complicated, so the other SEI messages probably do not require this type of change.

It was suggested to move the definitions from D.2.1 to D.12.11, to be local to NNPF interface text.

Decision: Adopt with moving the definitions from D.2.1 to D.12.11.

[JVET-AK0054](https://jvet-experts.org/doc_end_user/current_document.php?id=15025) AHG9: Semantics of the NNPFC and NNPFA SEI messages [Y.-K. Wang, J. Xu, L. Zhang (Bytedance), K. Yang, Y. Li, Y. Xu (SJTU)]

It is asserted that, in the current semantics of the NNPFC and NNPFA SEI messages, some aspects are not clearly specified, including:

1. The NNPF process may also be applied to an inserted (interpolated or extrapolated) picture, in addition to be applied to a cropped decoded output (CDO) picture or a processed version of a CDO picture. However, this aspect is not clearly specified. In this case, in particular, the current description of "input picture with index 0" is unclear.
2. The phrases like “an input picture that is not present in the bitstream”, "an input picture inputPicA that is not present in the bitstream", and "the input picture inputPicB that is the closest to inputPicA in output order and is present in the bitstream" are vague and unclear when the candidate input picture is a processed version of a CDO picture and more unclear or even incorrect when the candidate input picture is an inserted picture. Regarding the aspect of "incorrect", for example, the current semantics of nnpfc\_absent\_input\_pic\_zero\_flag would be interpreted as that an input picture that is an inserted picture (thus not present in the bitstream) would be padded while it should not be padded but was inserted.

This contribution proposes changes to semantics of the NNPFC and NNPFA SEI messages to address the above asserted issues.

The detailed proposed text changes, marked relative to JVET-AJ2006-v3, are included in an attachment to this contribution.

It was suggested by a participant that the proposed changes tie the NNPF semantics closely to the SPO SEI message, and may not be necessary given the interface text changes in JVET-AK0053.

It was noted that nnpfc\_absent\_info\_pic\_zero\_flag is already in VSEI v3.

Alternate wording was suggested to make NNPF work with SPO without being so tightly coupled to SPO, “is not present in the list of pictures that are available as input for the NNPF”.

It was suggested that the proposed language could be clarified with a note.

Some clarification in VSEI to support NNPF with SPO is needed, but there is discussion about the best way to express this in the spec.

It was suggested for a contribution to be submitted containing alterative text.

A number of purely editorial changes are all agreed upon.

Discussed after consideration of alternative text vs. this proposal on 18 January 2025 at 14:42 PM. Chaired by S. Deshpande.

A version 4 of the document is available. The provided text is the result of offline activity amongst interested parties. The offline discussion concluded that the issue is mostly editorial.

The proposed specification text changes in 4 were acceptable to the group.

Decision: Adopt version 4 text

[JVET-AK0070](https://jvet-experts.org/doc_end_user/current_document.php?id=15041) AHG9: On seed parameter for NNPF [T. Chujoh, Z. Fan, R. Ishimoto, T. Ikai (Sharp), L. Jin, H. Watanabe (Waseda Univ.)]

Chaired by H. Tan on 15 January 2025 at 1115.

The generative model, which the text prompt may control, has recently been used for image processing. At the Kamer meeting, it was agreed to expand NNPF SEI messages to send text prompts generally. In the so-called diffusion models, one of the well-known generative AIs, the random seed value is an important parameter for obtaining an always same-generation result. This contribution proposed to extend the semantics of nnpfc\_auxiliary\_inp\_idc to allow sending a seed value to set it on the input tensor.

Similar item (i.e., signaling of seed information) was also proposed in JVET-AK0074. Same motivation.

Both proponents aims at including this in V4.

It is asserted that the output of NNPF may depend on seed value (e.g., temporal extrapolation).

How the seed value is determined by encoder? It was responded that it similar to how encoder does the mode decision (e.g., encoder may try different seeds and choose the best one).

It was commented there could be values that encoder needs to test to find the best.

It was asserted that if GFV uses diffusion model, seed value should be useful for them too.

It was commented that one of the benefit of having seed is to avoid the bad result from the filtering / generation.

It was asked if seed value is used for training? It is not used for training.

Some support from non-proponents to having seed signaled.

Signalling:

* Both use auxiliary input: OK
* Do we need gating in NNPFC? OK
* Coding of the seed value (fixed or ue(v))? Fixed-length coding seems to be better to allow bigger range for initial selection of random value. Use u(16)
* Capability to update the seed in NNPFA? It is asserted that seed value can change over time. This value is more for the use of the filter (i.e., at the activation). It is common that the same seed can be used. The update is needed to avoid some bad result. OK for update in NNPFA.

Further discussion planned to confirm the final text.

Further discussed on 18 January 2025 at 1457 chaired by H. Tan to confirm the final text.

Updated implementation of the above guideline was provided in JVET-AK0326

Some minor editorial improvements can / should be handled by Editors. For example, adding the term ‘present’ in the gating flag in nnpfc, bracket in nnpfa, minus symbol.

Decision: Adopt JVET-AK0326 into VSEI v4 draft.

[JVET-AK0072](https://jvet-experts.org/doc_end_user/current_document.php?id=15043) AHG9: Comments on NNPF SEI messages [M. M. Hannuksela, F. Cricri (Nokia)]

Chaired by H. Tan on 15 January 2025 at 1210.

The following items were proposed in this document:

1. Since a temporal extrapolation NNPF may use only one picture as input and does not filter the input picture(s) unless the NNPF has other purposes too, it is proposed to change the inference "When nnpfc\_num\_input\_pics\_minus1 is equal to 0, nnpfc\_input\_pic\_filtering\_flag[ 0 ] is inferred to be equal to 1" as follows:
2. Infer nnpfc\_input\_pic\_filtering\_flag[ 0 ] to 1, when nnpfc\_purpose includes a filtering purpose other than temporal extrapolation.
3. Infer nnpfc\_input\_pic\_filtering\_flag[ 0 ] to 0, when nnpfc\_purpose indicates only temporal extrapolation.
4. Since it is not meaningful to require filtering of input pictures for a temporal extrapolation NNPF, it is proposed to update the constraint that requires at least one nnpfc\_input\_pic\_filtering\_flag[ i ] to be equal to 1 to apply when the purposes of the NNPF include neither picture rate upsampling nor temporal extrapolation.
5. In order to enable encoders to use dynamic text prompts without requiring a dummy non-null static text prompt, it is proposed to allow the NNPFC text prompt nnpfc\_prompt to be a null string, when the NNPFA SEI message includes a text prompt nnpfa\_prompt.
6. It is proposed to separate the creation of auxiliary input matrices into its own function, which is called from the process DeriveInputTensor. This item is asserted to be editorial and make the specification text shorter and easier to read.

For item 1: Some support from non-proponents. Agreed

For item 2: Some support from non-proponents. Agreed.

Do we need to apply similar change to the similar constraint in NNPFA as well? Yes

For item 3: Some support for relaxing the constraint. However, it was suggested to refine the suggested text. Agreed in principle. Further discussion request to review the text.

For item 4: Make some part of codes that are used repeatedly as a function and it can be called from different places. Agreed.

Item 3 was further discussed on 18 January 2025 at 1515 chaired by H. Tan to review the text.

The text for the proposed updated contstraint was provided and reviewed. The text is confirmed.

Decision: Adopt items 1, 2 (with the suggestion above), 3, 4.

[JVET-AK0073](https://jvet-experts.org/doc_end_user/current_document.php?id=15044) AHG9: Enabling multiple instances of NNPF extrapolated pictures [M. M. Hannuksela, F. Cricri (Nokia)]

For item 1: See notes on JVET-AK324

Item 2 was presented on 18 January 2025 at 1603 Chaired by H. Tan.

It was proposed to allow indicating text prompts for each NNPF inference process separately in the NNPFC and NNPFA SEI messages.

This is needed only when the same temporal extrapolation filter is invoked multiple times. Each invocation may need its own prompt / prompt update.

It was commented that the text from this contribution can be merged into the text implementation from JVET-AK0324.

Decision: Adopt into the TuC and implemented as suggested above.

[JVET-AK0074](https://jvet-experts.org/doc_end_user/current_document.php?id=15045) AHG9: Handling output variability of an NNPF [M. M. Hannuksela, F. Cricri (Nokia)]

See notes in JVET-AK0070.

[JVET-AK0147](https://jvet-experts.org/doc_end_user/current_document.php?id=15118) AHG9: On NNPF for Tone Mapping with Colour Volume Information [C.-H. Demarty, E. François, F. Aumont, O. Le Meur (InterDigital)]

This contribution proposes to add a new purpose to the neural-network post-filter characteristics (NNPFC) SEI message, called Tone Mapping. Metadata defining the target colour volume after applying the Tone Mapping operation are also proposed to be added to NNPFC syntax elements.

Assuming that the energy consumption of display devices is correlated to the amount of emitted light [1], the NNPF Tone Mapping purpose can be used to adapt the energy consumption of display devices. To this end, additional metadata related to the variation of luminance are proposed.

JVET-AK0136 proposes a new SEI message, which would be referenced in option #2.

Option #1 includes a number of additional syntax elements to describe a tone mapping with colour volume. Option #2 defines a new external SEI to carry those syntax elements.

Previous contribution JVET-AJ0147 is related. This contribution adds target volume information. A question was raised about including master colour volume.

A question was raised about if all of the presence flags in the proposed syntax are useful.

A question was raised about the usefulness of a ratio when other values are available.

Option #2 with an external SEI would allow the information to be used for other purposes than NNPF.

Decision: Adopt option #1 to the TuC.

[JVET-AK0152](https://jvet-experts.org/doc_end_user/current_document.php?id=15123) AHG9: On Spatial Extrapolation for NNPF [S. Deshpande (Sharp), M. M. Hannuksela (Nokia)]

Chaired by H. Tan on 15 January 2025 at 1237.

Corner cases related to patch-wise operation when using spatial extrapolation are identified and equations related to those are proposed to be modified. This aspect has been captured in the Editor’s note and it is asserted that the proposed modifications help to resolve the Editor’s note.

2 corner cases identified in VSEI v4 draft.

The proposed change is to check corner case first and then common case in the else part.

Decision: Adopt (bugfix).

[JVET-AK0167](https://jvet-experts.org/doc_end_user/current_document.php?id=15138) AHG9: On Enabling the usage of auxiliary layers for NNPF [T. M. Borges, Y. Sanchez, R. Skupin, C. Hellge, T. Schierl (Fraunhofer HHI)]

Chaired by H. Tan on 15 January 2025 at 1247.

This contribution proposed to allow the use of pictures from multiple layers as an input to NNPF, also known as cross-layer NNPF. It is proposed to add an additional value of nnpfc\_auxiliary\_inp\_idc to indicate such functionality.

The proposal suggests enabling the use of auxiliary layers in a straightforward manner as outlined below:

1. When nnpfc\_auxiliary\_inp\_idc equal to 4, the number of auxiliary layers is set by nnpfc\_num\_aux\_layers\_minus1, and the link between each selected layer is made using nnpfc\_layer\_target\_id[ i ].
2. Auxiliary layers are constrained to have picture size, bitdepth and order\_idc equal to the input picture for simplicity.
3. Different chroma format (ChromaFormatIdc) is permitted only when the auxiliary layers are monochrome. In all other cases, the input and auxiliary pictures must share the same chroma format for simplicity.

It was commented that pseudocode need to be updated to use bitmask.

It was commented that the idea of allowing cross-layer NNFP was proposed in AJ0130.

It was if cross-layer NNPF can be useful. One non-proponent agreed it can be useful.

Do we have NNPFC purpose that use this cross-layer? It is asserted that temporal extrapolation may use this.

It was asserted that the problem with AJ0130 (i.e., regarding scalable nesting SEI message) can be fixed. It was further commented that the problematic part with scalable nesting SEI is the extraction part.

It was commented this feature seems to be OK to be included in the TuC. Also AJ0130?

It was suggested the following 3 aspects are studied when considering the cross-layer NNPF:

1) what is possible already by existing features, including use of the scalable nesting SEI message;

2) what functionalities are good to support but are currently missing; and

3) what is the best way to add the supports of such desirable missing functionalities.

It was commented that the 2nd proposed item need to be made more flexible. This can be studied further if it is included in the TuC.

Decision: Include this and JVET-AJ0130 in to the TuC.

[JVET-AK0188](https://jvet-experts.org/doc_end_user/current_document.php?id=15159) AHG9: On NNPF filtering control strength value adjustment [J. Xu, Y.-K. Wang (Bytedance)]

In the current design of NNPF, filtering control strength values can be passed to the NNPF but cannot be adjusted. In this contribution, it is asserted that such a design has various limitations. A new syntax element is proposed to be added to the NNPFA SEI message so that adjustment of the filtering control strength values is enabled.

It was noted that NNPFA is signaled at most once per picture, so the proposed signaling of an adjustment would apply to an entire picture and not be adjusted per slice.

A -v3 bitstream would not include the proposed new syntax element, nnpfa\_delta\_strength\_control\_val.

A -v3 decoder would not make the signalled QP adjustment, but a -v4 decoder would make the adjustment.

The QpBdOffset variable is referred to in the proposed text, but is defined in VVC not in VSEI.

There were some questions regarding the usefulness of the proposal.

No action.

[JVET-AK0322](https://jvet-experts.org/doc_end_user/current_document.php?id=15311) AHG9: On the semantics of the NNPFC SEI message in relation to the SPO SEI message [M. M. Hannuksela (Nokia), H. Tan (LGE)] [late]

See notes under JVET-AK0054.

[JVET-AK0324](https://jvet-experts.org/doc_end_user/current_document.php?id=15313) AHG9: On Multiple NNPF Extrapolations [S. Deshpande (Sharp), E. François, C.-H. Demarty (Interdigital), M. M. Hannuksela, F. Cricri (Nokia)] [late]

Chaired by Hendry Tan on 18 January 2025 at 1522.

It was proposed to support multiple temporal extrapolations. For this, following is proposed related to the multiple extrapolations:

1. It is proposed to signal whether multiple extrapolations are reached through performing NNPF inference process more than once or through the NNPF generating multiple alternatives as subpictures within an NNPF-generated picture.
2. It is proposed to modify the TuC design of the multiple temporal extrapolations based on subpictures for asserted simplification. The proposed simplifications and modifications include the following: It is proposed to define subpictures along a single dimension. It is asserted that this simplifies the syntax, semantics, and derivation process. (Modification of the TuC design in JVET-AJ2032)
3. It is proposed to allow performing the NNPF inference process more than once for the same picture to generate multiple alternative sets of NNPF-generated pictures. (Aspect 1 of JVET-AK0073)

It was asked how this handle the multiple pictures outputted for the same time. It is expected that it will be handled thru external means.

The alternative pictures do not add the number of extrapolated pictures. Each of the alternative picture is like additional subpicture in the extrapolated picture. They will stacked horizontally.

There is a process to remove the alternative “subpictures” that are not selected (via external means).

The multiple alternative can also be achieved by invoking the same temporal extrapolation filter multiple times. The mechanism to choose the final output picture to be used is also via external means.

It was commented that the harmonized approach may be added into the TuC.

Decision: Adopt into the TuC. Replace the other two approaches with this.

[JVET-AK0326](https://jvet-experts.org/doc_end_user/current_document.php?id=15315) AHG9: Proposed specification of seed value for NNPF [T. Chujoh, Z. Fan, R. Ishimoto, T. Ikai (Sharp), L. Jin, H. Watanabe (Waseda Univ.), M. M. Hannuksela, F. Cricri (Nokia)] [late]

See notes on JVET-AK0070.

### Generative face video SEI messages (8)

Contributions in this area were discussed at 1430–1745 on Wednesday 15 Jan. 2025 (chaired by J. Boyce).

[JVET-AK0080](https://jvet-experts.org/doc_end_user/current_document.php?id=15051) AHG9: Comments on the GFV and GFVE SEI messages [M. M. Hannuksela (Nokia), J. Chen, B. Chen, Y. Ye (Alibaba)]

Chaired by H. Tan on 15 January 2025 at 1726

This input document proposed the following items:

1. Clarification of the base and driving picture interface variables for the GFV SEI message
2. Clarification of the base and generated picture interface variables for the GFVE SEI message
3. Addition of the VVC interface text for the usage of the GFV SEI message
4. Addition of the VVC interface text for the usage of the GFVE SEI message

Option 2 is asserted to be cleaner in codec side (for the interface). One participant commented it is better to have more complete text in the VSEI side, less on the interface. The following two reasons were asserted support this:

1. The interface would later needed for other codec as well
2. The interface is needed only some of the info needed to be provided

Decision: Adopt option 2 for all items.

[JVET-AK0124](https://jvet-experts.org/doc_end_user/current_document.php?id=15095) AHG9: On timing information and order of pictures in GFV SEI message [J. Lee, H. Tan, C. Kim, J. Nam, J. Lim, S. Kim (LGE)]

This contribution proposes modifications that are asserted to provide improvement in handling timing and order of pictures in GFV SEI message. The proposed modifications are as follows:

1. Clarify the order of a decoded output picture and GFV-generated pictures.
2. Clarify how to handle the timing information of GFV-generated pictures.

No action on Item 1, because it is not seen as necessary.

It was noted that at a previous meeting that timing information was not deemed necessary for GFV.

No action on Item 2.

[JVET-AK0127](https://jvet-experts.org/doc_end_user/current_document.php?id=15098) AHG9: On miscellaneous aspects of GFV and DSCI SEI messages [J. Lee, H. Tan, C. Kim, J. Nam, J. Lim, S. Kim (LGE)]

This contribution proposes miscellaneous modifications that are asserted to provide improvement to some aspects of GFV and DSCI SEI messages in VSEI v4 draft. The proposed modifications are as follows:

(Related to GFV SEI message)

1. Modify the condition where the matrix information is present

Simplifies the expression of the condition in the table.

Decision: Adopt Item 1.

[JVET-AK0128](https://jvet-experts.org/doc_end_user/current_document.php?id=15099) AHG9: Editorial updates for GFV SEI message [J. Lee, H. Tan, C. Kim, J. Nam, J. Lim, S. Kim (LGE)]

This contribution proposes several editorial modifications for GFV SEI message in VSEI v4 draft. The proposed modifications are as follows:

1. Modify constraint on a gfv\_cnt value.
2. Modify constraint on a gfv\_nn\_present\_flag.
3. Remove constraint in the semantics of gfv\_kps\_pred\_flag.
4. Remove constraint in the semantics of gfv\_matrix\_pred\_flag.

It was suggested that Item 1 disallows repetition of a GFV message, which is not consistent with the design intent.

No action on item 1.

Item 2 is purely editorial. It was suggested that the current text is easier to be understood as mandating a constraint on the presence of a translator NN that is later referenced.

No action on item 2.

Item 3 is asserted to be an editorial text simplification.

Decision: Adopt Item 3.

Item 4 is also asserted to be an editorial text simplification.

Decision: Adopt Item 4.

[JVET-AK0154](https://jvet-experts.org/doc_end_user/current_document.php?id=15125) AHG9: Comments on Generative Face Video SEI [A. C. Sidiya, S. Deshpande (Sharp)]

The current syntax for Generative Face Video SEI message signals each of the matrix (i.e. gfv\_matrix\_type\_idx[ i ) elements as decimal values. It is asserted that for some of the gfv\_matrix\_type\_idx[ i ] currently defined (e.g. head location matrix in 2D space) and application specific (gfv\_matrix\_type\_idx[ i ] in the range of 8 to 31, inclusive) and future Reserved matrices (gfv\_matrix\_type\_idx[ i ] in the range of 32 to 63 inclusive), the matrix elements may be signalled as integer values (instead of decimal values) for better coding efficiency. It is proposed to enable signaling that certain matrices contain integer values only.

For this following is proposed: a flag (gfv\_only\_integer\_matrix\_elements\_flag) is signalled to indicate if any of the signalled matrices contain integer values. If this flag is equal to 1, a flag is signalled for each of the signalled matrix types (gfv\_matrix\_only\_integer\_values\_flag[ i ]) to indicate if the particular matrix contains integer values only.

This proposal is asserting coding efficiency benefits. It was suggested that tests using the CTC should be performed to measure coding efficiency. However, the CTC software does not provide support head location. Currently, the CTC tests for FOMM currently signal only one type of matrix, and signals two types of matrices for the FV2V.

If the decimal portion is used, this proposal has a cost of 1 bit per SEI message.

A question was raised as to how often matrices would be expected to have integer values.

It was questioned if the decimal precision could be adjusted for each matrix type.

Further study to see bit rate savings using CTC-like conditions.

[JVET-AK0164](https://jvet-experts.org/doc_end_user/current_document.php?id=15135) AHG9: Supplementation of value range definition and editorial bugs fixes on the GFVE pupil position SEI messages [F. Ma, A. Trioux, F. Yang (Xidian Univ.), B. Li, F. Xing, Z. Wang (Hisense)]

In the 36th meeting, JVET-AJ0135[1] proposed to add pupil position information as part of the GFVE SEI messages. Additionally, during the same meeting, the inclusion of GFVC in VSEI v4 was agreed upon. In the GFVE SEI message of JVET-AJ2006 [2] (the VSEI draft 4), the syntax elements gfve\_pupil\_left\_eye\_dx\_ coordinate\_abs, gfve\_pupil\_left\_eye\_dy\_coordinate\_abs, gfve\_pupil\_right\_eye\_dx\_coordinate\_abs, and gfve\_pupil\_right\_eye\_dy\_coordinate\_abs were found to miss value ranges. In this contribution, we supplement the value range definitions for the above syntax elements, and provide some related editorial bugs fixes to the GFVE SEI messages. The detailed modification content can be found in the attachment.

JVET-AK0239 also proposes the same constraint on value range.

Agreed to apply the constraint.

Editorial bug fix to correct a syntax element name.

Editorial bug fix to add missing descriptor for a syntax element in the syntax table.

Decision (Bug fix): Adopt.

[JVET-AK0238](https://jvet-experts.org/doc_end_user/current_document.php?id=15227) AHG9: Semantics fixes for generative face video (GFV) SEI message [J. Chen, B. Chen, Y. Ye (Alibaba)] [late]

This contribution is to fix the semantics and syntax of GFV SEI message. The following items are proposed:

1. Fix the inference of gfv\_matrix\_width\_minus1[ i ] and gfv\_matrix\_height\_minus1[ i ].
2. Change the value range of gfv\_matrix\_delta\_element\_dec[ i ][ j ][ k ][ m ] to 0 to 2gfv\_matrix\_element\_precision\_factor\_minus1 + 1 – 1.
3. Fix the semantics of gfv\_base\_pic\_flag and gfve\_base\_pic\_flag.
4. Remove the redundant condition check in signaling of gfv\_num\_matrices\_equal\_to\_num\_kps\_flag[ i ].

For item 1, it was suggested that the existing semantics could alternatively be revised to remove “is present”.

Agreed for item 1.

Agreed for item 2, changing value range.

Agreed for item 3.

Agreed for item 4.

Decision: Adopt.

[JVET-AK0239](https://jvet-experts.org/doc_end_user/current_document.php?id=15228) AHG9: On generative face video enhancement (GFVE) SEI message [J. Chen, B. Chen, Y. Ye (Alibaba)] [late]

The generative face video enhancement SEI message was studied in this contribution and the following items are proposed.

1. Add a gating flag to control the signaling of the enhancement matrix
2. Separate the signaling of integer part and decimal part of matrix element by splitting the syntax element gfve\_matrix\_element[ i ][ j ][ k ] into gfve\_matrix\_element\_int[ i ][ j ][ k ] and gfve\_matrix\_element\_dec[ i ][ j ][ k ].
3. Apply prediction scheme for the signaling of matrix as what GFV SEI message does.
4. Align the matrix element precision of GFVE SEI message with that of GFV SEI message by replacing gfve\_matrix\_element\_precision\_factor with gfve\_matrix\_element\_precision\_factor\_minus1.
5. Add the value range for the gfve\_pupil\_left\_eye\_dx\_coordinate\_abs, gfve\_pupil\_left\_eye\_dy\_coordinate\_abs, gfve\_pupil\_right\_eye\_dx\_coordinate\_abs, gfve\_pupil\_right\_eye\_dy\_coordinate\_abs.
6. Clarify that the enhancement process EnhancerNN() and output process StoreOutputTensors( ) are only invoked when gfve\_base\_pic\_flag is equal to 0 to enhance the quality of the picture which is generated by the GFV SEI message and output enhanced picture.

Agreed for item 1.

Item 2 proposes to separate the matrix signalling into an integer part coded as ue(v) and a decimal part coded as u(v), which would be consistent with the signalling in the GFV SEI message. It is asserted to have better coding efficiency. For GFV, splitting the matrix into integer and decimal parts was shown to be more efficient in JVET-AI0195 and the proponents assert that it is expected to be similar for GFVE.

Item 3 proposes adding prediction functionality of matrix elements, which is asserted by the proponent to be more coding efficient. For GFV, performing prediction was shown to be beneficial, but no results are available for GFVE.

A participant suggested that it may be beneficial for GFVE to be consistent with GFV.

Consider asking AHG16 to consider if this functionality is desired.

Agreed for item 4, add \_minus1 to gfve\_matrix\_element\_precision\_factor.

Item 5 already addressed in JVET-AK0164.

Item 6 is asserted to be a clarification of the design intent.

Agreed for item 6.

It was suggested that there is benefit for design consistency between GFV and GFVE.

Decision: Adopt.

It is noted that the GFVE SW is not available for VTM.

### Digitally signed content SEI messages (8)

Contributions in this area were discussed at 1745–1940 on Wednesday 15 Jan. 2025 (chaired by J. Boyce) and 905 – 1100 on Thursday 16 Jan 2025 (chaired by S. Deshpande).

[JVET-AK0109](https://jvet-experts.org/doc_end_user/current_document.php?id=15080) AHG9: On signaling and constraint for RefDigest in digitally signed content SEI messages [C. Kim, H. Tan, J. Nam, J. Lee, J. Lim, S. Kim (LGE)]

Chaired by S. Deshpande on 16 January 2025 at 9:05 AM

The current digitally signed content SEI messages design references the RefDigest as the hash of the substream identified by dscv\_verification\_substream\_id − 1, always pointing to the immediately preceding substream. While this approach ensures temporal consistency in simple and continuous substream structures, it faces significant limitations in scenarios requiring scalability, such as non-contiguous substreams or hierarchical encoding based on temporal or layer IDs. These challenges hinder the ability to maintain authentication and scalability effectively, necessitating a more flexible and explicit referencing mechanism.

This contribution proposes the following items:

1. Addition of a syntax element to explicitly specify the ID of the substream to be used for RefDigest.

2. Addition of a constraint on RefDigest to ensure layer and temporal consistency

Regarding item 1:

Regarding the example 3, it was asked at time tn why substream 0, 1, 2 cannot be used again instead of using substreams 3, 4, 5.

It was commented that for only temporal sublayers or only spatial layers the problem may not exist but for an example which includes multilayer/ multiview, multi-temporal layer case – where the hierarchy is not clear, the current signaling may not be sufficient in covering it.

It was also suggested that the dscv\_ref\_info\_present\_flag could be put inside the if condition to save 1 bit for substream 0. This was agreed by proponents.

Regarding item 2:

It was commented that the proposed constraint is a bit unclear – especially regarding the wording.

Also, it was asked how the TemporalId of CurrDigest is known. Proponents asserted that it is known to the encoder side.

It was asked how the constraint will apply for AVC (regarding use of layer id, temporal id). It was agreed that some rethinking is needed to write the constraint in that case.

It was asked if single substream case may be harmed with this constraint (an example was provided).

To be discussed together with other related proposal JVET-AK0287.

Regarding item 1:

Regarding the example 3, it was asked at time tn why substream 0, 1, 2 cannot be used again instead of using substreams 3, 4, 5.

It was commented that for only temporal sublayers or only spatial layers the problem may not exist but for an example which includes multilayer/ multiview, multi-temporal layer case – where the hierarchy is not clear, the current signaling may not be sufficient in covering it.

It was also suggested that the dscv\_ref\_info\_present\_flag could be put inside the if condition to save 1 bit for substream 0. This was agreed by proponents.

Regarding item 2:

It was commented that the proposed constraint is a bit unclear – especially regarding the wording.

Also, it was asked how the TemporalId of CurrDigest is known. Proponents asserted that it is known to the encoder side.

It was asked how the constraint will apply for AVC (regarding use of layer id, temporal id). It was agreed that some rethinking is needed to write the constraint in that case.

It was asked if single substream case may be harmed with this constraint (an example was provided).

To be discussed together with other related proposal JVET-AK0287.

Discussed again on 19 January 2025, at 14:35 PM, chaired by S. Deshpande.

Regarding the constraint the first review pointed out that wording is too strict and does not work in some cases (related to substream 0).

After offline discussion the constraint is updated. V4 is available and was discussed.

Two options are proposed. One for VSEI and other for VVC interface.

It is asserted that VVC interface can more easily use some syntax elements from VVC directly.

Proponent of JVET-AK0287 suggested including this constraint text as a note to the sub-bitstream extraction part of JVET-AK0287 and that this conformance is stricter than required.

It was discussed that the sub-bitstream extraction process is more flexible compared to the constraint here which is stricter and more prescriptive about what is allowed.

A non-proponent expressed a preference for including constraint in VSEI (versus including in VVC interface).

It was commented that we could have both the approach from JVET-AK0287 sub-bitstream extraction of sub-bitstream and constraint from JVET-AK0109 included in specification text (not as a note).

It was asked by couple of non-proponents why the case of removal for part of bitstream and if the remaining sub-bitstream is conforming or not is a part of verification entity and why is such a text specified in this part.

It was commented that it does not make sense to sign something that is not decodable.

It was commented that the proposed constraint makes the verification process “scalable” which is desirable.

A non-proponent expressed support for VSEI constraint compared to sub-bitstream extraction.

Decision: Adopt VSEI constraint from v4 of the document (some wording modification would be needed during editing period).

[JVET-AK0110](https://jvet-experts.org/doc_end_user/current_document.php?id=15081) AHG9: On reference message digest for verification in digitally signed content SEI messages [C. Kim, H. Tan, J. Nam, J. Lee, J. Lim, S. Kim (LGE)]

Chaired by S. Deshpande on 16 January 2025 at 9:30 AM

It is asserted that the current design of digitally signed content (DSC) SEI messages allows for the possibility of using the RefDigest from a substream in the preceding CVS. However, referencing substreams from a preceding CVS violates the coding design principle of CVS independence and complicates random access scenarios. This contribution proposes the following item:

1. Modify the RefDigest determination process to ensure CVS independence while supporting random access.

It was commented that current design intent is to allow references across CVSs to have temporal consistency to e.g. avoid shuffling of temporal segments to tamper with the content.

Proponents asserted that such a design may break random access, splicing use cases. It was commented that such is not the case (except for first segment of random access) and that this has been described in a note.

It was commented that the trust of the bitstream will be broken if we don’t do this across CVS (as per the current design).

It was commented by a few participants to retain the current design intent.

For splicing case, it was suggested that adding end of sequence can make sure there is no problem (with the current design).

It was commented that flexible splicing is not desirable when doing digital signing.

It was agreed to keep the current design intent.

No action.

It is asserted that the current design of digitally signed content (DSC) SEI messages allows for the possibility of using the RefDigest from a substream in the preceding CVS. However, referencing substreams from a preceding CVS violates the coding design principle of CVS independence and complicates random access scenarios. This contribution proposes the following item:

1. Modify the RefDigest determination process to ensure CVS independence while supporting random access.

It was commented that current design intent is to allow references across CVSs to have temporal consistency to e.g. avoid shuffling of temporal segments to tamper with the content.

Proponents asserted that such a design may break random access, splicing use cases. It was commented that such is not the case (except for first segment of random access) and that this has been described in a note.

It was commented that the trust of the bitstream will be broken if we don’t do this across CVS (as per the current design).

It was commented by a few participants to retain the current design intent.

For splicing case, it was suggested that adding end of sequence can make sure there is no problem (with the current design).

It was commented that flexible splicing is not desirable when doing digital signing.

It was agreed to keep the current design intent.

No action.

[JVET-AK0125](https://jvet-experts.org/doc_end_user/current_document.php?id=15096) AHG9: On presence and persistency of digitally signed content SEI messages [J. Lee, H. Tan, C. Kim, J. Nam, J. Lim, S. Kim (LGE)]

This contribution proposes modifications that are asserted to provide improvement in handling presence and persistency of digitally signed content initialization (DSCI), digitally signed content selection (DSCS), and digitally signed content verification (DSCV) SEI messages. The proposed modifications are as follows:

(Related to DSCI SEI message)

1. Clarify the quantity number of DSCI SEI message(s) in an AU, and how to handle them.

(Related to DSCS SEI message)

2. Clarify the position of DSCS SEI message.

3. Extend persistence scope of DSCS SEI message

(Related to DSCV SEI message)

4. Add constraint on the dscv\_verification\_substream\_id value.

For Item 1, JVET-AK0206 proposes having an ID for verification, that would lead to multiple initialization messages.

Repetition of SEI messages is sometimes used for error resiliency purposes.

Further discuss Item 1 after review of JVET-AK0206.

Agree on Item 2.

For Item 3, it is most useful when the CVS contains several consecutive pictures for the same substream. It is noted that it is possible to infer the value of 0 of substream ID for the single substream use case.

No action on Item 3.

For Item 4, a constraint and a note are proposed.

It was suggested that the design intent of the DSC SEI messages are to be used with CRA, GDR, etc., and also to allow more frequent verification for large spacing. Initialization message content would change for the repetition.

Some support for Item 4, but to possibly refine the wording of the note.

Agreed for Item 4.

Item 1 further discussed after decision in JVET-AK0206 to include DSC ID. Option 2 of Item 1 was preferred.

The text should be revised to say that at most one DSC initialization SEI message with a particular ID value shall be in an AU.

Decision: Adopt Items 1, 2, 4. New version of the document to be uploaded containing the precise text of Item 1.

[JVET-AK0126](https://jvet-experts.org/doc_end_user/current_document.php?id=15097) AHG9: On signaling the mapping between verification substreams and layer and temporal sublayer for DSCI SEI message [J. Lee, H. Tan, C. Kim, J. Nam, J. Lim, S. Kim (LGE)]

Chaired by S. Deshpande on 16 January 2025 at 9:50 AM

This contribution proposes to include signaling the mapping between verification substreams and layer and temporal sublayer for digitally signed content initialization (DSCI) SEI message. It is asserted that the proposed signaling information enables digital signing functionality for each layer or temporal sublayer representation without receiving and parsing a bitstream including entire layers and/or temporal sublayers.

This can also be useful in the bitstream extraction process as extractor may use this information to ensure that the remaining data can still be verified.

It was asked how this signalled information will be used. It was commented that this is metadata information for receiver to know substream structure. It is asserted to be useful for the decoder in case there has been bitstream extraction done at decoder/ middle box.

It was commented that the current design assumes that for bitstream extraction case entire substreams are dropped and in that case there is no problem.

In current design substreams are hierarchical and they can be dropped from starting from the highest one.

It was commented that signalling layer id, temporal id is not codec agnostic. Also, some of the bitstream conformance constraint related to max layer id is unclear.

It was commented that this adds a lot of text which needs to be clearly checked if an action is taken.

To be discussed together after presentation of JVET-AK0287.

This contribution proposes to include signaling the mapping between verification substreams and layer and temporal sublayer for digitally signed content initialization (DSCI) SEI message. It is asserted that the proposed signaling information enables digital signing functionality for each layer or temporal sublayer representation without receiving and parsing a bitstream including entire layers and/or temporal sublayers.

This can also be useful in the bitstream extraction process as extractor may use this information to ensure that the remaining data can still be verified.

It was asked how this signalled information will be used. It was commented that this is metadata information for receiver to know substream structure. It is asserted to be useful for the decoder in case there has been bitstream extraction done at decoder/ middle box.

It was commented that the current design assumes that for bitstream extraction case entire substreams are dropped and in that case there is no problem.

In current design substreams are hierarchical and they can be dropped from starting from the highest one.

It was commented that signalling layer id, temporal id is not codec agnostic. Also, some of the bitstream conformance constraint related to max layer id is unclear.

It was commented that this adds a lot of text which needs to be clearly checked if an action is taken.

To be discussed together after presentation of JVET-AK0287.

[JVET-AK0127](https://jvet-experts.org/doc_end_user/current_document.php?id=15098) AHG9: On miscellaneous aspects of GFV and DSCI SEI messages [J. Lee, H. Tan, C. Kim, J. Nam, J. Lim, S. Kim (LGE)]

This contribution proposes miscellaneous modifications that are asserted to provide improvement to some aspects of GFV and DSCI SEI messages in VSEI v4 draft. The proposed modifications are as follows:

(Related to DSCI SEI message)

2. Clarify the value range of dsci\_num\_verification\_substreams\_minus1.

3. Clarify the value range of dsci\_key\_register\_idx.

For Item 2, option 1 is more bitrate efficient when the number of substreams is small.

Agreed on Item 2 option 1.

For Item 2, a range for dsci\_key\_register\_idx of 232 was seen as too large. Agreed that a range constraint should be applied.

Discussed again to select a suitable range on 19 January 2025, 15:15 PM chaired by S. Deshpande.

It was suggested that 10 bits may be sufficient and so the range can be 0 to 1023, inclusive.

Decision: Adopt Item 2 option 1 and to use the range as specified above

[JVET-AK0194](https://jvet-experts.org/doc_end_user/current_document.php?id=15165) AHG9: Digitally Signed Content SEI messages for AVC and HEVC [K. Sühring, T. Hinz, Y. Sanchez, J. Pfaff, H. Schwarz, D. Marpe, T. Wiegand (Fraunhofer HHI)]

Digitally signed content SEI messages were adopted at the 36th JVET meeting into VSEI with interface text for VVC. This contribution proposes interface text for AVC and HEVC, so that the digitally signed content SEI messages can also be used in AVC and HEVC bitstreams to allow adding digital content signatures for these widely deployed standards.

It was noted that there is a plan for HEVC CD at this meeting. There is no current plan for a new version of AVC.

The proposed text for HEVC does not include VPS in the list of nonVclDigitallySignedNalUnitsList. A participant suggested that VPS should be included.

Decision: Adopt HEVC interface to the HEVC CD.

Discussed again on 19 January 2025 at 15:05 PM chaired by S. Deshpande. to consider if VPS should be added to the list of nonVclDigitallySignedNalUnitsList.

It was commented by proponents that VPS does not need to be rewritten so it could be included or not included.

It was commented that in some multilayer cases VPS is needed doe decoding (for VVC, HEVC). Single layer HEVC, does not need VPS for decoding.

It was suggested to include VPS for HEVC and not for VVC.

Option 1: Don’t include VPS for HEVC and VVC (current WD)

Option 2: Include VPS for HEVC and VVC

Option 3: Include VPS for HEVC but not for VVC

Multiple participants preferred option 2.

Decision: Option 2 is agreed

Further discussed at 1625 on 19 Jan chaired by J. Boyce.

It is proposed to put the AVC interface to the AVC output document rather than the TuC. Agreed.

Decision: Adopt  the AVC interface to the AVC output document, JVET-AK1017.

[JVET-AK0206](https://jvet-experts.org/doc_end_user/current_document.php?id=15195) AHG9: On digitally signed content SEI messages [S. McCarthy, C. Fersch, I. Sodagar (Dolby)]

This contribution proposes additional syntax and semantics for the digitally signed content SEI messages that are currently included in JVET-AJ2006 (Additional SEI messages for VSEI version 4 (Draft 4)). Specifically, this contribution proposes 3 aspects:

1. Adding a syntax element to identify the specific combination of hash method, source URI, number of verification substreams, etc. used by the content producer for digitally signing the content and therefore to be used for verification by the receiver.
2. Indicating which non-VCL NAL units are included in digitally signed content.
3. Indicating which SEI messages are included in digitally signed content.

This contribution also proposes interface text for JVET-AJ2005 for aspects 2 and 3.

For aspect 1, it was suggested that additional verification messages can be signaled with different ID values for different hash methods or verified by different entities.

It was noted that if there is a single substream, it is still not necessary to signal a selection SEI for each PU, even with the addition of the dsc\_id. Agreed on Aspect 1 to be added to VSEI v4.

Decision: Adopt aspect 1 to VSEI v4.

It was suggested that some SEI messages, such as GFV or NNPF might be useful to protected.

A question was raised about repeated SEI messages. The current proposal doesn’t address repeated SEI messages.

It is noted that Aspects 2 and 3 require significant changes in the syntax of the DSC SEI messages.

It was noted that containing of SEI messages in SEI NAL unit and scalable nesting could further complicate the design.

It was suggested that aspect 3 could potentially be redesigned into a backwards compatible approach.

It was also suggested that additional data at the end of an SEI would impact the design.

No action on aspect 2.

Further study on aspect 3.

[JVET-AK0287](https://jvet-experts.org/doc_end_user/current_document.php?id=15276) AHG9: Multilayer digitally signed content authentication SEI messages [J. Boyce, M. M. Hannuksela (Nokia)] [late]

Chaired by S. Deshpande on 16 January 2025 at 10:10 AM

It is proposed to extend the digitally signed content (DSC) authentication SEI messages for more complete support of multilayer bitstreams. It is asserted that the current design does not fully support the flexibilities possible in multilayer bitstreams when dependencies are not hierarchical, including some combinations of spatial and temporal scalability.

The proposed changes can be summarized as follows:

* Modify the DSC initialization SEI:
  + Signal the number of substreams
  + Signal flags for each substream to indicate if it references previous substreams
  + Require that a DSCI SEI message is present in the first IDR, CRA, and GDR PU of an AU, rather than in all IDR, CRA, and GDR PUs as in the current design
* Modify the DSC selection SEI:
  + Require that the SEI message precede the relevant NAL units in each PU for which the SEI is present, rather than for each AU as in the current design

This part has already been addressed by actions taken on other proposals.

* Modify the DSC verification SEI:
  + Change calculation of RefDigest to not refer to previous CurrDigest of a substream
  + Two options provided to change calculation of CurrDigest to include reference substreams of the current substream

Regarding DSCV SEI changes it was commented that encoder will need to calculate multiple hashes. It was commented that option 1 proposed does this but the other option 2 does not. For option 1 it was asked what order is used for repeated hash calculations. Example of substream 0,1,0,1 was mentioned. In current design hashes are concatenated and asserted to also be somewhat equivalent to option 2. Also it was commented that the decoder may have more load when using option 1.

It was commented that option 2 where RefDigest is calculated based on reference information signalled in proposed modification to DSCI SEI could follow current design of concatenating hashes.

It was commented that there may be more concatenation of hashes with the proposed approach than done currently.

It was commented that for option 2 changing 0 to sId may not support current design intent of switching up substreams starting from 0. Proponents had specifically considered temporal upswitching and downswitching case and was asserted to work (in manner similar to current random access design with respect to the first segment).

It was commented that in JVET-AK0109 there is only one reference, but this design has multiple references and whether all references are used instead of direct reference. It was commented by proponents that this is done because there may not be a strict hierarchy of references.

It was commented that it may be ok to not cover some use cases and adding any flexibility should not enable manipulations that may allow tampering. It was agreed that the flexibility should not break such design goal.

It was commented that multiple of these proposals related to multilayer and scalability and it was then asked if scalable nesting SEI was considered to handle these use cases. I was commented that SEI messages are not included in signatures and so it was not clear how to use them. But the question was about nesting these 3 digital signing SEI messages. It was suggested to consider if such an approach can be solved using such scalable nesting mechanism. For example, we had agreed to do that for multilayer NNPF aspects at this meeting.

It was commented that we could study further if substream should always be considered to be an operating point.

It was agreed by proponents to focus on option 2 instead of option 1 for DSCV.

It was agreed to have offline discussion amongst interested parties.

Discussed again on January 19, 2025 at 14:10 PM chaired by S. Deshpande.

There was offline discussion as requested above. The v3 version is available as a result and was discussed.

For independent non-base layer, uuid is required to be included for verification to avoid tampering.

A mini sub-bitstream extraction process for substreams is specified.

It was to be double checked that the current substream is not removed.

It was commented that for this part JVET-AK0109 constraint is related and should be reviewed together.

The name “trustworthy” added to SEI is a leftover oversight.

It was asked why sub-bitstream extraction process needs to be defined.

It was answered that this is to show what part of the extracted stream can be verified.

Text change is needed related to replacing a flag with ID in the verification part.

It was commented that this proposal includes lot of changes and so it was questioned if it is appropriate for VSEI v4 or for TuC. The proponents commented that for supporting multilayer it would be better to include it in VSEI v4. It was commented that the proposed changes could be integrated in working software and studied during meeting cycle and if issues are found it can be taken out of vsei v4.

No objection was received and support was expressed to include this in VSEI v4 – with the part related to sub-bitstream extraction of substreams to be discussed together with JVET-AK0109. See notes about the relation between that part and that document under JVET-AK0109.

Decision: Adopt (except for sub-bitstream extraction part)

### Other (4)

Contributions in this area were discussed at 1105–1620 on Thursday 16 Jan. 2025 (chaired by J. Boyce).

[JVET-AK0075](https://jvet-experts.org/doc_end_user/current_document.php?id=15046) AHG9: Additional information on object-based optimization for the encoder optimization information SEI message [M. Pedzisz, M. M. Hannuksela, F. Cricri (Nokia)]

Chaired by S. Deshpande on 16 January 2025 at 11:05 AM

When the encoder has performed object-based optimization by encoding objects with a lower QP value than the background and indicated such optimization in the EOI SEI message, it is asserted that the decoding system cannot reliably conclude which QP values indicate objects, since QP values may vary spatially due to encoder's bitrate adaptation. Therefore, this input document proposes to indicate a QP threshold in the EOI SEI message to enable decoding systems to determine areas including objects from areas outside of objects. The following items are proposed in this input document:

1. When the EOI SEI message indicates object-based optimization with QP adaptation, it is proposed to add the following indications in the EOI SEI message:
   1. eoi\_quant\_threshold\_delta, which indicates the quantization parameter threshold relative to initial picture quantization parameter to classify areas either to be outside the detected objects or to include one or more detected objects. eoi\_quant\_threshold\_delta equal to 0 indicates that the quantization parameter threshold is unknown or unspecified.
   2. eoi\_pic\_quant\_object\_flag; when equal to 1, it indicates that the initial picture quantization parameter indicates an area including a detected object; and when equal to 0, it indicates that the initial picture quantization parameter indicates an area outside of detected objects.
2. Derivation of the initial picture quantization parameter interface variable in the VVC interface text of the EOI SEI message.

It was asked if using QP can be useful in this case for object detection. It was answered that bit 2 already mentions that encoder has done object-based optimization and eoi\_quant\_threshold\_delta can also be set to 0 if no information is available.

It was asked if the signalled information is sufficient for the listed use cases. The proponent thought that this information would be sufficient and needs fewer bits compared to using e.g. OMI SEI.

It was asked whether signaling a direct quantization parameter instead of delta compared to PicQuant is better.

It was asked if we have clear understanding if the interface invocation is done only once when an SEI is invoked or will PicQuant be updated at each picture even when the SEI (e.g. EOI SEI) persists for multiple pictures.

It was agreed that we need to clarify that aspect.

It was commented that we can invoke the codec interface at each picture to get the PicQuant value (and that this is similar to NNPF design).

Agreement in principle to invoke the interface on a per picture basis for this particular proposal. The text will be revised considering that.

Decision: Adopt (the revised version to be uploaded).

It was recommended to further study the aspect of per picture invocation for other SEIs.

It was asked if using QP can be useful in this case for object detection. It was answered that bit 2 already mentions that encoder has done object-based optimization and eoi\_quant\_threshold\_delta can also be set to 0 if no information is available.

It was asked if the signalled information is sufficient for the listed use cases. The proponent thought that this information would be sufficient and needs fewer bits compared to using e.g. OMI SEI.

It was asked whether signaling a direct quantization parameter instead of delta compared to PicQuant is better.

It was asked if we have clear understanding if the interface invocation is done only once when an SEI is invoked or will PicQuant be updated at each picture even when the SEI (e.g. EOI SEI) persists for multiple pictures.

It was agreed that we need to clarify that aspect.

It was commented that we can invoke the codec interface at each picture to get the PicQuant value (and that this is similar to NNPF design).

Agreement in principle to invoke the interface on a per picture basis for this particular proposal. The text will be revised considering that.

Decision: Adopt (the revised version to be uploaded).

It was recommended to further study the aspect of per picture invocation for other SEIs.

[JVET-AK0155](https://jvet-experts.org/doc_end_user/current_document.php?id=15126) AHG9: On Encoder optimization information SEI message [C. H. Demarty, E. François, F. Aumont, O. Le Meur (InterDigital)]

This contribution proposes to enhance the Encoder optimization information (EOI) SEI message with information related to potential optimization for energy savings purposes.

When energy consideration comes to play, such information can for example be used by a receiver to know the energy-aware optimization level associated to the received content. This information can be used by the receiver to decide how to proceed with the content, depending on e.g. its power capabilities, or on an energy profile defined by the end-user.

For example the receiver may drop the content when it is not compliant with the receiver eco mode; or the receiver may decide to apply energy-aware post-processings to reduce the energy for example when displaying the content or using it through machine analysis. Additionally, in an interactive use case, such information can be used by a sender system to inform the receiver about the energy-aware optimizations that were applied on a content, upon a request by a receiver.

The proposed 2-bit indicator uses reserved bits in the EOI SEI.

It was questioned why both a 2-bit indicator and a bit in the eoi\_type mask are necessary.

It would be possible to add syntax elements condition on the bit in the eoi\_type mask.

Some support was expressed for adding a bit mask for in eoi\_type.

It was suggested to replace the bitMask interpretation to be for luma range optimization, with an additional bit signaled for increasing or decreasing the level.

It was noted that the content light level SEI message can be used to indicate content light level, but does not indicate that a change was made by the encoder vs. the source content light level.

It is also proposed to add a bit mask to eoi\_object\_based\_idc.

It was suggested that the semantics for this mask be made more specific than just saying “energy aware method”.

It was suggested to provide some additional information justifying the use case for the object based processing for energy savings.

It would suggested to clarify that the bit mask for eoi\_type applies to the entire picture.

Further discussion requested after revised document is available.

Further discussed at 1653 on Saturday 18 Jan.

In version 3, the following changes are made:

* references are provided as examples of region based processing for energy savings,
* one of the 2 reserved bits is used to signal the proposed new syntax element eoi\_for\_energy\_saving\_idc,
* a “luma range adaptation” encoder optimization type is added at picture- and object-levels. Additional metadata specifying the luma range adaptation are added conditionally to the flag EoiLumaAdaptationFlag.

In version 3, one of the reserved bits was used to signal a new syntax element. At the initial review, that was not intended to be part of the new version to be reviewed.

In version 3, additional syntax elements were added not just for the new EoiLumaAdaptationFlag but also for other flag type (EoiObjectBasedFlag && eoi\_object\_based\_idc), for which a new value of eoi\_object\_based\_idc value was also added.

It was suggested to replace this: EoiLumaAdaptationFlag = ( ( ( eoi\_type & 0x40 ) > 0 ) | | ( ( ( eoi\_type & 0x01 ) > 0 ) && ( (eoi\_object\_based\_idc & 0x20 ) > 0 ) ) ) ? 1 : 0

With this: EoiLumaAdaptationFlag = ( ( ( eoi\_type & 0x40 ) > 0 )

It was noted that the EOI SEI is referenced by the TR for machine consumption. A flag for intended for human vision also exist. Adaption of luma may also benefit machine consumption.

The semantics for eoi\_ratio\_luma\_value contain the words “when required”, which are suggested to be removed.

The semantics for eoi\_ratio\_luma\_sign\_flag includes the following:

The value of eoi\_ratio\_luma\_value can be derived from the encoder optimization as ( ratioSign \* 100 \* ( lumaNew – lumaOld ) / lumaOld ), …

This wording refers to derive the value of a syntax element. It would need to be reworded to clarify that it is describing what the encoder had done.

It is unclear how lumaOld is determined.The EOI SEI has persistence, so presumably would apply to all pictures in the persistence range of the SEI message, e.g. the current picture and pictures that precede it in the CLVS following the EOI SEI message.

It was suggested that a decoder/receiver would benefit from additional information if it was desired to reverse the process done at the encoder to reconstruct something similar to the source picture, but that will not be required.

Some support for addition to the TuC of the bit mask field for eoi\_type, EoiLumaAdaptationFlag derived based on that mask only, and addition of the 3 syntax elements gated by it as highlighted in the syntax table in yellow.

Conditionally adopt to the TuC based on confirmation of the revised text.

Further discussed 1940 on 19 Jan 2025.

Reviewed the text in version -v4. Confirmed adoption.

Decision: Adopt to the TuC.

[JVET-AK0168](https://jvet-experts.org/doc_end_user/current_document.php?id=15139) AHG9: On SPTI SEI message: Robustness and source constant framerate [Y. Sanchez, T. M. Borges, R. Skupin, C. Hellge, T. Schierl (Fraunhofer HHI)]

This contribution proposes some modification to the SPTI SEI message as outlined below:

1. To allow defining the time interval relative to a previous or next output picture that is not a non-reference picture.
2. To allow defining the time interval relative to a previous or next output picture of a higher temporal sublayer.
3. To clearly state that the interval is defined relative to a previous or next output picture in the same CLVS.

Interface was provided for VVC.

It was noted that HEVC has a sub-layer non-reference flag, which is not addressed in this contribution. The semantics would need to be updated if this was to be supported.

It was suggested to provide interface text for the AVC and HEVC.

There is not any AVC interface text for SPTI SEI in the current TuC but is in JVET-AF2032. No AVC WD text currently exists.

JVET-AF2032 contains this text:

Use of the source picture timing SEI message in AVC

For purposes of interpretation of the source picture timing SEI message, the following variable is specified:

– If the bitstream conforms to any of the profiles defined in Annex G, H, I, or J, TemporalId is set equal to temporal\_id.

– Otherwise (the bitstream conforms to a profile defined in Annex A), TemporalId is set equal to ( nal\_ref\_idc = = 0 ? 1 : 0 ).

Further discussion requested item 1 after interface text is provided.

A spti\_reference\_mode flag is proposed

It is suggested that the spti\_source\_type bit mask could be extended to indicate that constant frame rate is likely. This spti\_source\_type bit mask was preferred vs. the approach in this contribution, and would not require any change to the semantics of spti\_sublayer\_interval\_scale\_factor[ i ].

Suggest wording for a new bitMask value is: Constant frame rate: The temporal distance between source pictures is likely to be constant.

Agreed add a bitMask value for spti\_source\_type.

Item 3 provides updated semantics for spti\_sublayer\_interval\_scale\_factor[ i ].

Agreed to Item 3.

Further discussion requested for Item 1.

Further discussed 1550 on 19 Jan.

The -v2 version of the document was discussed.

In -v2, text was provided to also handle sub-layer non-reference pictures.

Additionally, interface text for HEVC and AVC was provided, which had been lost in previous editing in JVET-AF2032.

A question was raised about if this amount of change is necessary to address sub-layer non-reference pictures.

The previous interface text for AVC available in JVET-AF2032 used nal\_ref\_idc to set the TemporalId value used in VSEI. A participant preferred the proposed approach vs. that approach shown in JVET-AF2032.

Decision: Adopt Item 1, a bit mask as described above, and Item 3, which includes the AVC interface in the AVC output document, in addition to the HEVC, VVC, and VSEI outputs.

Separately, it is noted that software for SPTI is requested for inclusion in the HM and JM software. S. McCarthy (Dolby) was the original proponent of SPTI, and has agreed to provide software.

[JVET-AK0321](https://jvet-experts.org/doc_end_user/current_document.php?id=15310) Current Issues of the Object Mask Information SEI message [D. Podborski, A.M. Tourapis (Apple)] [late]

The current draft (Draft 4) of the document on “Additional SEI messages for VSEI version 4” includes the Object Mask Information (OMI) SEI message, which could be used for applications that may wish to associate object masks with the encoded video data. However, after a preliminary study, the authors of this document feel that the current syntax and semantics of this SEI message may have several potential issues. This document identifies some of these potential issues. It is consequently requested that the (OMI) SEI message is removed from the VSEI draft document and instead further studied so as to at least ensure its correctness.

The current SEI message has a omi\_mask\_id[ i ][ j ] which is not optional. A presence flag could be added, with the ID inferred to be equal to the index when not present.

JVET-AK0203 is suggested to be related.

Item 1 prefers that the SEI message be associated with the auxiliary picture than with the primary picture. The current design uses primary picture because some of the characteristics described in the SEI message are associated with the primary picture, and not the map that is present in the aux pic.

Item 2 is a bug that is easily fixable.

Item 3 could possibly be addressed with a constraint on omi\_mask\_pic\_update\_flag or to initialize to a zero value.

Regarding Item 4, it was clarified that updating can be done for individual auxiliary picture layers. Could study to determine if some additional clarifying language such as a note can be added.

Regarding Item 5, the design intent is for omi\_mask\_cancel[ i ][ j ] to cancel an ID so that it can be reused for another object.

It was noted that the annotated regions SEI message defines a global array, with an ID used an index into the global array.

There seems to be a bug in the semantics of omi\_mask\_cancel[ i ][ j ], which should indicate that the cancels the persistence scope of the object with id equal to omi\_mask\_id[ i ][ j ].

Regarding Item 6, a constraint can be added that the omi\_mask\_id[ i ][[ j ] and omi\_aux\_sample\_value[ i ][ j ] are not the same for more than one value of i, j.

Item 7 describes a possible modification of a feature rather than a bug.

Item 8 can be addressed by fixing the variable name to start with a capital letter to indicate a global variable.

Item 9 expresses a preference for avoiding the need to support functionalities that are not needed for some use cases.

Regarding Item 10, perhaps a restriction could be imposed to avoid partial update chains across CRA and GDR boundaries. The solution suggested for Item 3 could also be applied for CRA, etc. AUs.

A number of issues were raised in the contribution. None seem to be so significant that they can’t be addressed in a timely manner that would preclude inclusion of the OMI SEI in VSEI v4.

It is requested that a contribution that implements the suggested fixes to the OMI SEI be provided and reviewed this meeting.

[JVET-AK0330](https://jvet-experts.org/doc_end_user/current_document.php?id=15319) AHG9: Fixes to object mask information SEI message [J. Chen, Y. Ye (Alibaba)] [late]

Reviewed 19 Jan 2025.

This contribution provides the fixed text of the OMI SEI message according to the meeting discussion of JVET-AK0321.

Most of the changes are editorial in nature. However, there is a syntax change in replacing the cancel flag with a new object flag that is used to indicate that an object id is assigned to a new object.

Two options were proposed for items 3 and 10. The first option requires that omi\_mask\_pic\_update\_flag[ ] be set to 1 for random access pictures. The second option removes the flag, effectively requiring that the removed flag would always be set to 1. The first option provides more flexibility to avoid updates for some layers.

The first option is preferred.

Decision: Adopt with the first option for items 3 and 10 as described above.

## AHG9: Aspects on SEI messages in TuC and related (23)

### Constituent rectangles SEI (4)

Contributions in this area were discussed at 1625–1820 on Thursday 16 Jan. 2025 (chaired by S. Deshpande).

[JVET-AK0115](https://jvet-experts.org/doc_end_user/current_document.php?id=15086) AHG9: On supporting 4:4:4 color format in constituent rectangles and enhanced colour format information SEI messages [C. Kim, H. Tan, J. Nam, J. Lee, J. Lim, S. Kim (LGE)]

Chaired by S. Deshpande on 16 January 2025 at 16:20 PM

It is asserted that the Constituent Rectangles (CR) SEI message and the Enhanced Colour Format Information (ECFI) SEI message, included in the TuC for the future extension of VSEI (JVET-AJ2032), while have different design, share functional similarities, such as delivering 4:4:4 source content to decoders or devices supporting lower chroma formats. This contribution provides analysis of the overlapping functionalities of the CR and ECFI SEI messages and evaluation of their distinct roles.

Based on the analysis and evaluation, the following items are proposed:

1. Modify the color description signalling in the CR SEI message

1. Add explicitly defined color description to ECFI SEI message.
2. Modify guard band signaling in ECFI SEI message.

Regarding item 1 CR SEI: It was commented that JVET-AJ0286 includes a simplified ECFI proposal. It was commented that CR SEI has more features that ECFI SEI. It was suggested to consider to include only one of the two SEIs in TuC. It was commented by another participant that keeping both messages in TuC is ok but if/ when time to consider moving something to WD arises we need careful thinking.

It was asked what VUI means in the context of VSEI. It was pointed out that VUI syntax structure is included in VSEI spec text but is not referred for HEVC.

It was pointed that CR SEI colour syntax elements have an inference to infer the values from an associated VUI. It was asked if the item 1 has any benefit compared to explicit colour information signaling already supported.

It was asked if colour\_description\_present\_flag[i] is needed as each of the colour description syntax elements allow unknown as possible indication.

The proposal would also need a change in the semantics of explicitly signalled 4 syntax elements (current proposal does not provide it).

Also it should be required that the referred layer should exist.

Item 1: agreed (revised text will be uploaded)

Regarding item 2 ECFI SEI:

It was commented that ECFI is not restricted to 4:4:4 it could support 4:2:2. It was commented that chroma location may also need to be signalled in the case VUI is not referred.

Item 2 was discussed after review of JVET-AJ0286.

There was support to add it as proposed.

Item 2: agreed

It was left for further study to decide how/ if chroma location information should be signalled.

Regarding item 3:

With the modification you can not have one guard band 0, which current syntax supports.

It was also asked why the flag is inside for loop and not just once outside.

No action on item 3.

It was asked what VUI means in the context of VSEI. It was pointed out that VUI syntax structure is included in VSEI spec text but is not referred for HEVC.

It was pointed that CR SEI colour syntax elements have an inference to infer the values from an associated VUI. It was asked if the item 1 has any benefit compared to explicit colour information signaling already supported.

It was asked if colour\_description\_present\_flag[i] is needed as each of the colour description syntax elements allow unknown as possible indication.

The proposal would also need a change in the semantics of explicitly signalled 4 syntax elements (current proposal does not provide it).

Also it should be required that the referred layer should exist.

Item 1: agreed (revised text will be uploaded)

Regarding item 2 ECFI SEI:

It was commented that ECFI is not restricted to 4:4:4 it could support 4:2:2. It was commented that chroma location may also need to be signalled in the case VUI is not referred.

Item 2 was discussed after review of JVET-AJ0286.

There was support to add it as proposed.

Item 2: agreed

It was left for further study to decide how/ if chroma location information should be signalled.

Regarding item 3:

With the modification you can not have one guard band 0, which current syntax supports.

It was also asked why the flag is inside for loop and not just once outside.

No action on item 3.

[JVET-AK0209](https://jvet-experts.org/doc_end_user/current_document.php?id=15198) AHG9: On constituent rectangles SEI message [H. Tan, C. Kim, J. Nam, J. Lee, J. Lim, S. Kim (LGE)]

Chaired by S. Deshpande on 16 January 2025 at 17:20 PM

This contribution proposes several changes to the specification of constituent rectangles SEI message.

Item 1:

The derivation for position and size have asserted problem such that the concept of CR within layers and CR in term of the whole set of CR are mixed. Hence, the calculation is asserted to be incorrect when there are more than one layer included in the CR SEI message.

Agreed (bug fix)

Item 2:

It is asserted that the equations for deriving the position and size of CRs that have same size in a particular layer can work only on the condition that the picture size from that particular layer is integer multiplication of the number of columns and the number of rows. Constraint to ensure that is asserted to be needed but it is missing from the current specification text.

It is proposed to add such a constraint.

It was commented that this constraint is unnecessarily limiting and the current calculation assertedly works and that CRs could have leftovers. The design is such that CRs will be same but there may be unused parts in the picture.

It was suggested to add a note to explain this. But it was not unclear what the note should contain.

No action on item 2.

Item 3:

When CRs from a particular layer is based on subpictures, the current constraints related to number of constituent and the number of subpictures are asserted to be not correct. The constraints should be for the number of subpictures related to the number of CRs from that particular layer only, however, the current constraint assertedly uses the variable that denotes the total number of CRs.

Agreed (bug fix)

Item 4:

It is asserted that : When CRs from a particular layer is based on subpictures, in the case that the number of CRs from that layer is less than the number of subpictures in that layer, then only the first set of subpictures (i.e., up to the number of CR from that layer) can be described / used in the CR SEI message. It is proposed to be desirable if the index of the subpicture for each CR from that layer can be specified.

This functionality improvement was commented to be useful.

Agreed.

Item 5:

Identifier for the CR is gated by 2 flags. It is asserted that, it is unnecessary to signal cr\_rect\_id\_present\_flag for each CR when cr\_rect\_id\_enabled\_flag is equal to 1. The use of identifier for CR is asserted to be more useful either all CRs have identifier, none of CR has identifier, or CRs in certain layers have identifier.

It was commented that the current design considers that for a set of consecutive CRs the ID can be inferred.

This may save some bits but may not be aligned with current design choice.

No action

Item 6:

The current design of constituent rectangles (CR) SEI message includes signalling of constituent rectangle type and type description. It is asserted that the design of signalling the CR type description is not effective and not intuitive. While combining the signalling of cr\_rect\_type\_description[ lIdx ][ I ] in just one loop at the end of the SEI message may save some byte-alignment bits, such signalling is asserted to be not intuitive. It is proposed to signal that syntax element right after the gating flag for it.

Current design saves byte alignment by doing it only once.

No action.

The derivation for position and size have asserted problem such that the concept of CR within layers and CR in term of the whole set of CR are mixed. Hence, the calculation is asserted to be incorrect when there are more than one layer included in the CR SEI message.

Agreed (bug fix)

Item 2:

It is asserted that the equations for deriving the position and size of CRs that have same size in a particular layer can work only on the condition that the picture size from that particular layer is integer multiplication of the number of columns and the number of rows. Constraint to ensure that is asserted to be needed but it is missing from the current specification text.

It is proposed to add such a constraint.

It was commented that this constraint is unnecessarily limiting and the current calculation assertedly works and that CRs could have leftovers. The design is such that CRs will be same but there may be unused parts in the picture.

It was suggested to add a note to explain this. But it was not unclear what the note should contain.

No action on item 2.

Item 3:

When CRs from a particular layer is based on subpictures, the current constraints related to number of constituent and the number of subpictures are asserted to be not correct. The constraints should be for the number of subpictures related to the number of CRs from that particular layer only, however, the current constraint assertedly uses the variable that denotes the total number of CRs.

Agreed (bug fix)

Item 4:

It is asserted that : When CRs from a particular layer is based on subpictures, in the case that the number of CRs from that layer is less than the number of subpictures in that layer, then only the first set of subpictures (i.e., up to the number of CR from that layer) can be described / used in the CR SEI message. It is proposed to be desirable if the index of the subpicture for each CR from that layer can be specified.

This functionality improvement was commented to be useful.

Agreed.

Item 5:

Identifier for the CR is gated by 2 flags. It is asserted that, it is unnecessary to signal cr\_rect\_id\_present\_flag for each CR when cr\_rect\_id\_enabled\_flag is equal to 1. The use of identifier for CR is asserted to be more useful either all CRs have identifier, none of CR has identifier, or CRs in certain layers have identifier.

It was commented that the current design considers that for a set of consecutive CRs the ID can be inferred.

This may save some bits but may not be aligned with current design choice.

No action

Item 6:

The current design of constituent rectangles (CR) SEI message includes signalling of constituent rectangle type and type description. It is asserted that the design of signalling the CR type description is not effective and not intuitive. While combining the signalling of cr\_rect\_type\_description[ lIdx ][ I ] in just one loop at the end of the SEI message may save some byte-alignment bits, such signalling is asserted to be not intuitive. It is proposed to signal that syntax element right after the gating flag for it.

Current design saves byte alignment by doing it only once.

No action.

[JVET-AK0210](https://jvet-experts.org/doc_end_user/current_document.php?id=15199) AHG9: On constituent rectangle grouping for 4:4:4 color format support in constituent rectangle SEI message [H. Tan, J. Nam, C. Kim, J. Lee, J. Lim, S. Kim (LGE)]

Chaired by S. Deshpande on 16 January 2025 at 17:50 PM

It is asserted that the text in the semantics of constituent rectangles (CR) SEI message related to the support for 4:4:4 color format has the following problems:

1. The existing constraint such that “there shall be one value y in the range of 0 to cr\_num\_rects\_minus1, for which cr\_rect\_group\_id[ y ] equals j and cr\_rect\_idc[ y ] equals 0 or 4” is asserted to be unclear. It is asserted that it only ensures that such CR is present but fails to limit the number of such CR. In the case that there are more than one CR like that in the group, it is not clear how the decoding system can decide which of them to be used. The same problem applies to CR for Y and U components as well.
2. Related to the constraint (i.e., “there shall be one value y in the range of 0 to cr\_num\_rects\_minus1, for which cr\_rect\_group\_id[ y ] equals j and cr\_rect\_idc\_ y ] equals 0 or 4”), the value range for the CR index should be 0 to CrNumRects – 1, inclusive, instead of 0 to cr\_num\_rects\_minus1.
3. A CR can be from partition within a picture in a particular layer (e.g., subpicture of a picture) or the whole picture itself. Thus, for reconstructing the target picture when 4:4:4 color format is enabled, instead of referring to syntax element such as cr\_rect\_width\_in\_units\_minus1[ y ] and cr\_rect\_height\_in\_units\_minus1[ y ], it should refer to the variables such as crRectWidth[ y ] and crRectHeight[ y ].

In this contribution, the following changes are proposed to overcome the asserted problems:

1. Update the constraints related to the CR grouping when cr\_group\_444\_flag[ j ] equal to 1, there shall be exactly one CR in the group with type for Y component, exactly one CR in the group with type for U component and exactly one CR in the group with type for V component.
2. For those constraints, the CR shall be one of CR in the range of 0 to CrNumRects – 1, inclusive.

For reconstructing target picture from constituent rectangles when 4:4:4 colour format is enabled, use the picture size based on the derivation of the picture that consider whether the constituent rectangles are from partition or the whole picture. In other words, use the size from derived variables instead of the from syntax elements.

Regarding cyan highlighted text, the range should include -1 when using the derived variables.

Agreed (the text needs to be updated to add a -1 for the rectangle derived variables.)

Regarding cyan highlighted text, the range should include -1 when using the derived variables.

Agreed (the text needs to be updated to add a -1 for the rectangle derived variables.)

[JVET-AK0278](https://jvet-experts.org/doc_end_user/current_document.php?id=15267) AHG9: Constituent Rectangles / Enhanced Colour Format Information SEI extension to support 4:2:2 source content coded in bitstream with 4:2:0 chroma format [S. Keating, M. Ikeda (Sony)] [late]

Chaired by S. Deshpande on 16 January 2025 at 18:00 PM

It is proposed to extend the Constituent Rectangles (CR) SEI message and/or the Enhanced Colour Format Information (ECFI) SEI message to support coding of 4:2:2 chroma format source content in a bitstream coded with 4:2:0 or monochrome chroma format.

It was asked if chroma location information is signalled.

It is agreed that currently it is not signalled in the proposal but needs to be signalled.

For ECFI change is not needed as it is already supported.

Also, for CR only the 2.1.1 unequal rectangle sizes needs to be considered.

There currently exists a 4:4:4 flag in CR which would need some change. Some size requirement adjustment would be needed.

The idea is supported by the group for CR based on 2.1.1 section of the proposal.

Further discussion needed to review the exact text which needs to be developed.

See notes for JVET-AK0332.

[JVET-AK0332](https://jvet-experts.org/doc_end_user/current_document.php?id=15321) AHG9: Syntax for constituent Rectangles SEI extension to support 4:2:2 source content coded in bitstream with 4:2:0 chroma format [S. Keating (Sony), J. Boyce (Nokia)] [late]

Chaired by S. Deshpande on 18 January 2025 at 16:50 PM

JVET-AK0278 proposed to extend the Constituent Rectangles (CR) SEI message and/or the Enhanced Colour Format Information (ECFI) SEI message to support coding of 4:2:2 chroma format source content in a bitstream coded with 4:2:0 or monochrome chroma format. It was requested to provide SEI syntax and semantics based on the initial review of the contribution, that is provided here.

It was commented that cr\_group\_enhanced\_chroma\_format\_idc[ i ] could be a flag to support 4:4:4 and 4:2:2 but is an indicator for possibility of supporting monochrome.

The text needs to be modified to specify that the height needs to be the same. This will be fixed in a revision.

Decision: Adopt (the revised version).

### Display overlay SEI (6)

Contributions in this area were discussed at 1840–1945 on Thursday 16 Jan. 2025 (chaired by J. Boyce).

[JVET-AK0099](https://jvet-experts.org/doc_end_user/current_document.php?id=15070) AHG9: On bit depth alignment between target and coded pictures in DOI SEI [T. Biatek, J. Boyce, M. M. Hannuksela (Nokia)]

Chaired by S. Deshpande on 16 January 2025 at 18:40 PM

The current DOI SEI design does not indicate the bit depth of the target picture, and the target picture formation process does not consider that DOI components may be coded with different bit depths. This contribution proposes to signal the bit depth of the target picture and updates the target picture formation process to consider the bit depths of the DOI components and the target picture.

There was support for this proposal from a non-proponent.

Decision: Adopt

[JVET-AK0100](https://jvet-experts.org/doc_end_user/current_document.php?id=15071) AHG9: DOI SEI support for target picture with alpha channel [T. Biatek, J. Boyce, M. M. Hannuksela (Nokia)]

Chaired by S. Deshpande on 16 January 2025 at 18:45 PM

The DOI SEI message supports use of alpha channel pictures in the formation of a target picture. However, the target picture itself represents only texture. It is proposed to add support in the DOI SEI message for formation of a target picture alpha channel that is associated with the primary target picture. Formation of a target picture alpha channel enables usage of the DOI SEI message for generation of a standalone overlay, without the background video contained in the bitstream. Support for alpha target pictures also enables cascading of DOI SEI messages.

The target picture has texture component but additionally the proposal can provide alpha channel for the target picture.

There was support for this proposal from a non-proponent.

Decision: Adopt

[JVET-AK0157](https://jvet-experts.org/doc_end_user/current_document.php?id=15128) AHG9: On the number of DOI SEI message in an AU [J. Nam, H. Tan, J. Lee, C. Kim, J. Lim, S. Kim (LGE)]

This contribution proposes to handle multiple DOI SEI messages associated with an AU.

Option 1: specify that when multiple DOI SEI messages are associated with a picture, the order of application / execution of those DOI SEI messages is specified by SEI processing order (SPO) SEI message.

Option 2: specify that when multiple DOI SEI messages are associated with a picture, decoder shall choose one DOI SEI message to be applied to the picture.

It was suggested that the proposals may be overly restrictive.

It was suggested that a note could be added to indicate that SPO could be used to determine the order.

Another option is when multiple DOI SEIs are in the SPO, to indicate the order is defined by the SPO. If not in the SPO, the order is determined via decoder means. It would be good to study exact language for this to make sure it isn’t too restrictive.

It was commented that the list of SEIs which may be described by the SPO SEI needs to be updated to include DOI.

Discussed again on 19 January 2025 at 18:50 PM chaired by S. Deshpande.

V2 is available based on offline discussion. This version was discussed. There was support for the new text.

Decision: adopt (text in section 2.1 of v2)

[JVET-AK0158](https://jvet-experts.org/doc_end_user/current_document.php?id=15129) AHG9: On reconstruction of target display picture in DOI SEI message [J. Nam, H. Tan, J. Lee, C. Kim, J. Lim, S. Kim (LGE)]

It is asserted that the signalling of syntax elements and their semantics that are related to the chroma format and bit depth for target display picture and each overlay in the display overlays information (DOI) SEI message have problem.

This contribution proposes the following modifications to address the asserted problems:

1. Signal the chroma format idc for target display picture explicitly. This may be named doi\_target\_pic\_chroma\_format\_idc.
   1. With addition of doi\_target\_pic\_chroma\_format\_idc, the syntax element doi\_target\_init\_three\_comp\_flag is not needed since the number of initial samples can be determined based on the value of doi\_target\_pic\_chroma\_format\_idc.
2. Use the appropriate chroma format idc for reconstructing the target display picture and each of each overlay
   1. For target display picture, its chroma format is doi\_target\_pic\_chroma\_format\_idc.
   2. For each overlay, their chroma format is the chroma format of its layer that is provided by the codec.
3. Specify list of SubWidthC[ i ] and SubHeightC[ i ] to be provided by codec from the interface. In addition to that, use the appropriate SubWidthC and SubHeightC for target display picture and each overlay.
4. Correct the layer id and layer index in the interface of VVC for the use of DOI SEI message.
5. For the BitDepth value to be assigned for BitDepthY and BitDepthC, use BitDepth value from the first layer.

Item address in JVET-AK0099.

Decision: Adopt Items 1, 2, 3, and 4.

[JVET-AK0190](https://jvet-experts.org/doc_end_user/current_document.php?id=15161) AHG9: On the display overlays information SEI message [J. Xu, Y.-K. Wang (Bytedance), [H. Tan (LGE)](mailto:dr.hendry@lge.com)]

This contribution proposes the following changes for the display overlays information (DOI) SEI message:

1. Remove a constraint on doi\_partition\_idx.
2. Disallow referencing non-existing constituent rectangles when doi\_partition\_type\_idc is equal to 1.

Item 1 fixes an error during the editing period.

Agree to Item 1.

Item 2 address a problem where the semantics may refer to a syntax element that is not present in the CR SEI message.

Agree in principle to Item 2, but the exact wording to be checked.

Discussed again on 19 January 2025 at 19:00 PM chaired by S. Deshpande..

After offline discussion v2 is available and was discussed.

It was commented that this change is similar in spirit to JVET-AK0189 proposal related to BRI, which was agreed.

The revised text is agreed for item 2.

Decision: Adopt (v2 version)

[JVET-AK0265](https://jvet-experts.org/doc_end_user/current_document.php?id=15254) AHG9: Display overlay sets for DOI SEI [J. Boyce, T. Biatek, M. M. Hannuksela (Nokia)] [late]

Chaired by S. Deshpande on 16 January 2025 at 18:55 PM

It is proposed to modify the Display overlays info SEI to enable signaling of sets of display overlays, to enable formation of target pictures that include a subset of the display overlays described in the SEI message. Sets may be used, for example, to include display overlays for multiple languages in the same DOI SEI message. For this use case, a separate set may be defined for each language, with each set containing language-specific text overlays as well as non-text overlays that are common to multiple sets.

It was asked how external means would know which set to use for the process. It was suggested that some information about sets or overlays may be helpful for that. Further study recommended about that.

The signalled information is to help simplify the process.

There was support for this proposal from a non-proponent.

Decision: Adopt

### Quality metrics SEI (3)

Contributions in this area were discussed at 0900–XXXX on Saturday 18 Jan. 2025 (chaired by J. Boyce).

[JVET-AK0159](https://jvet-experts.org/doc_end_user/current_document.php?id=15130) AHG9: On miscellaneous aspects of quality metrics SEI message [J. Nam, H. Tan, J. Lee, C. Kim, J. Lim, S. Kim (LGE)]

This contribution proposes modifications that are asserted to provide improvement to some aspects of quality metrics (QM) SEI messages in JVET-AJ2032. The proposed modifications are as follows,

1. Modify the signalling of qm\_metric\_description[ i ] syntax element.

2. Specify a constraint on the values of qm\_metric\_definitions\_present\_flag, qm\_clvs\_values\_present\_flag, and qm\_pic\_values\_present\_flag syntax elements.

3. Move general texts on usage of the SEI message from VVC to VSEI.

4. Move the signaling of qm\_three\_component flag[ i ] syntax element.

5. Specify a constraint on the value of qm\_num\_metrics\_minus1 syntax element.

Agree Item 1, Option 1.

Agree Item 2.

Item 3 proposes moving text from the VVC interface text to within the SEI message description.

Agree Item 3.

Agree Item 4.

Item 5 proposes a constraint that the value of qm\_num\_metrics\_minus1 be the same for all QM SEI messages. It is asserted that there is an issue with the current design if the value changes. Imposing this constraint means that the same number of metrics must be signaled for all pictures in the CLVS.

Agree Item 5.

Further study to consider allowing flexibility in signaling different number of metrics per picture.

Decision: Adopt Item 1 Option 1, and Items 2, 3, 4, and 5.

[JVET-AK0207](https://jvet-experts.org/doc_end_user/current_document.php?id=15196) AHG9: On signalling of average quality in quality metrics SEI message [H. Tan, J. Nam, J. Lee, C. Kim, J. Lim, S. Kim (LGE)]

The current design of quality metrics (QM) SEI message includes signalling of average quality for CLVS. It is asserted that the design has the following problems:

1. The constraint to ensure that when a QM SEI message includes signalling of average quality for CLVS, the first of such SEI shall be present in the first picture unit.
2. When signalling of average quality for CLVS is present and one or more pictures is removed from the bitstream (e.g., due to bitstream extraction based on temporal sublayer), all the QM SEI messages in the CLVS need to be updated to ensure the correctness of the information signalled in the SEI messages.

This contribution proposes the following changes for handling the asserted problems above:

1. Specify a constraint such that when average quality value for CLVS is present in any QM SEI message, the first QM SEI message for that CLVS shall be present in the first picture unit of the CLVS and the average quality value(s) signalled in these two SEI messages shall be the same.
2. Modify the signalling of average quality metric values to be based on temporal id rather than just one value for the all pictures in the CLVS.

Item 1 aligns a constraint between when the SEI message is carried within an SPO or not.

Agree to Item 1.

For item 2, the qm\_num\_pic\_in\_tid[ k ] seems unnecessary.

It is noted that this change disallows signalling of a quality metric for the entire CLVS.

The proposed signaling of the quality metric per temporal layer applies to all temporal layers up to the temporal ID. It might be useful to know the metric per individual temporal layer.

There is some support for some type of signalling of quality metric for temporal layers. Further study encouraged to consider the utility of signaling quality for cumulative layers vs for individual layers.

Decision: Adopt Item 1.

[JVET-AK0208](https://jvet-experts.org/doc_end_user/current_document.php?id=15197) AHG9: On the inclusion of quality metric SEI message in SEI processing order SEI message [H. Tan, C. Kim, J. Nam, J. Lee, J. Lim, S. Kim (LGE), M. M. Hannuksela (Nokia)]

This contribution proposes the several changes to the semantics of quality metric (QM) SEI message. It is asserted that the proposed changes provide improvement to solve the asserted problems related to the inclusion of QM SEI message in SEI processing order (SPO) SEI message and when it is wrapped as PON-nested SEI message.

The proposed changes are as follows:

1. Clarify that QM SEI message may be present in the bitstream as non-PON-nested SEI message, PON-nested SEI message, or both.
2. Specify that PON-nested QM SEI messages that are part of the same processing chain (i.e., included in a particular sei processing order SEI message) have different scope / domain from the PON-nested QM SEI message that are part of another processing chain.
3. Clarify the association between QM SEI message that is included in an SPO SEI message to its presence in PON SEI message based on its processing order and prefix information.
4. Add a constraint such that when a QM SEI message is included in an SPO SEI message, the QM SEI message shall not be the first SEI message included in the SPO SEI message.
5. Update the current text such that the quality metric information that is signalled in the QM SEI message that is the i–th SEI message in an SPO SEI message indicates picture quality resulting from post-processing after invocation of all other SEI messages in the SPO SEI message that have po\_sei\_processing\_order[ j ] equal to or lower than the value of po\_sei\_processing\_order[ i ].

Regarding Item 1, it is noted that the current design already allows the QM SEI to be present in a PON and/or not nested in a PON. It seems unnecessary to add the proposed wording.

Item 2 is a clarification. Agreed to Item 2, with some potential editing.

It was suggested that the language proposed in Item 3 is already handled in the language for the SPO SEI.

Item 4 is withdrawn.

Item 5 is seen to already be the design intent and may already be addressed in AK0205.

It was noted that the original design intent was to be able to signal quality differences associated with processing of SEI messages. Care would need to be taken to avoid affecting the design intent.

Discussed again for item 5 follow-up on 19 January 2025 at 19:05 PM chaired by S. Deshpande.

After offline discussion v2 is available and was discussed.

Definition from general section of interface text needs to be moved for handling the proposed text correctly. This is delegated to the editors.

The text for item 5 in v2 was agreed.

Decision: Adopt Item 2, Item 5 (from v2).

### Other (10)

Contributions in this area were discussed at 1010–1300 on Saturday 18 Jan. 2025 (chaired by J. Boyce) and 1900 to 2000 on Sunday 19 Jan. 2025.

[JVET-AK0140](https://jvet-experts.org/doc_end_user/current_document.php?id=15111) AHG9/AHG8: Showcase for Packed regions information SEI [J. Boyce, H. Zhang, M. M. Hannuksela (Nokia)]

TBP in plenary

[JVET-AK0189](https://jvet-experts.org/doc_end_user/current_document.php?id=15160) AHG9: On the bitdepth range information SEI message [J. Xu, Y.-K. Wang (Bytedance)]

This contribution proposes the following changes for the bitdepth range information (BRI) SEI message:

1) Disallow referencing non-existing constituent rectangles when bri\_partition\_type\_idc is equal to 1.

2) Refine a constraint on bri\_partition\_idx so that it applies only when the syntax is present.

3) Disallow signalling partition parameters that lead to a partition out of picture boundaries when bri\_partition\_type\_idc is equal to 2.

4) Disallow a sample belonging to different ranges.

Agreed Item 1.

Agreed Item 2.

Regarding Item 3, it was noted that the DOI SEI should be checked to see if there is a similar issue.

Agreed Item 3.

Regarding Item 4, it is noted that overlap between partitions is not required when partitioning is used, but could be a problem for explicit signaling.

Agreed Item 4.

Decision: Adopt Items 1, 2, 3, and 4.

[JVET-AK0114](https://jvet-experts.org/doc_end_user/current_document.php?id=15085) AHG9: Updates and suggestion on AI usage restrictions SEI message [C. Kim, H. Tan, J. Nam, J. Lee, J. Lim, S. Kim (LGE)]

This document proposes the following items:

1. Addition of the "Do not use for AI-inference" restriction to aur\_restriction.

2. Addition of the AI usage restrictions (AUR) SEI message into a working draft of a VSEI amendment.

Aside from the proposed additional restrictions, it is also asserted that the design of AUR SEI message is stable and has enough maturity. Therefore, it is recommended to promote this SEI message to the VSEI version 4 draft.

Agreed to Item 1.

Regarding Item 2, it is noted that AUR SEI message was added to the TuC in July 2024 at the AI meeting and that the design has been stable, with just addition of values of au\_restriction[ i ].

It is noted that thus far this meeting, no SEI messages have been decided to be promoted to VSEI -v4.

Decision: Adopt Item 1.

[JVET-AK0179](https://jvet-experts.org/doc_end_user/current_document.php?id=15150) AHG9: On Lens Optical Correction SEI message [J. Samuelsson-Allendes, S. Deshpande (Sharp)]

This input contribution relates to the Lens Optical Correction (LOC) SEI message which is included in the technologies under consideration document for future versions of VSEI. The proposals in this document are intended to broaden the applicability of the LOC SEI message and are asserted to provide clarifications of the usage of the LOC SEI message. Specifically, the following items are proposed:

1. Separate the case of “no information” from “no distortion”
2. Remove two sentences that include the term “by default”
3. Express focal location as ratio of picture width and height, respectively

The existing semantics for the presence flag currently indicate no expected distortion, when in general they simply indicate that no information is provided.

It was questioned if there are use cases where parameters in some category would be available when other categories are not available.

A comment was that most real lenses have some distortion in all three categories.

A question was raised about if computer generated content such as generative AI may wish to emulate lens distortion in some but not all categories.

Agreed to Item 1.

Agreed to Item 2.

Regarding Item 3, it would be useful to have information on how focal parameters have been signaled in other standards by JVET or MPEG.

It is noted that the semantics language for loc\_focal\_center\_x and \_y use “default”, when it would be more appropriate to have inference language.

It is suggested to check if constraints on value ranges are necessary.

Further discussion requested on Item 3.

Decision: Adopt Items 1 and Item 2 (editorial).

Further discussed 1910 on 19 Jan.

The -v2 version adds the following to its abstract:

4. Add missing inferred values and value ranges for several syntax elements

The second version of the proposal adds an option B for item 3, where the focal location is expressed in luma samples. A fourth item is also added to define inferred values and value ranges for several syntax elements.

Decision: Adopt Item 3 option A and Item 4.

Follow-up discussion Monday 20 Jan. 1005 (chaired by JRO). It was commented that there is an ambiguity in the definition of the syntax elements loc\_rd\_poly\*, some of which are u(25), some (i25), some unclarity about correctness and necessity of the ranges, and due to the fact that a division occurs, it is unclear what the result would be, in which floating point format it comes.

It was agreed to modify the decisions above by removing the ranges in item 4, and conduct further study about correctness of syntax and semantics, and the necessity to impose range constraints.

[JVET-AK0204](https://jvet-experts.org/doc_end_user/current_document.php?id=15193) AHG9: Proposed Lens Optical Correction SEI message for version 5 of VSEI [S. Wenger, G. Teniou, A. T. Hinds (Tencent)]

TBP in plenary.

[JVET-AK0153](https://jvet-experts.org/doc_end_user/current_document.php?id=15124) AHG9: On Photosensitive Content Information Signaling [S. Deshpande (Sharp)]

Following modifications are proposed related to the photosensitive content information SEI message in the TuC:

* Proposal 1: It is proposed to allow to optionally signal more detailed information regarding the content with respect to specific categories/ guidelines specified in the photosensitive content standard.
* Proposal 2: It is proposed to optionally signal text string(s) specifying the text which can describe the information regarding the photosensitivity content at the discretion of the content provider.

The proposed approach seems consistent with the design philosophy to refer to 3rd party specs.

Agreed to Proposal 1.

Proposes u(8) signaling for the number of strings. It was suggested to reduce that number and include some reserved bits.

A question was raised if there is a constraint on the values of the individual standard indicators.

Agreed to Proposal 2 with using 4 bits for pci\_num\_descr\_strings and adding 4 reserved bits. A revised version of the contribution with these changes is to be uploaded.

Decision: Adopt Proposals 1 and Proposal 2 with modifications indicated above.

[JVET-AK0058](https://jvet-experts.org/doc_end_user/current_document.php?id=15029) [AHG9/AHG13] Comments on Film grain regions characteristics SEI message [S. Xie, P. Wu, Y. Gao, Y. Bai, C. Huang (ZTE)]

While the model adaptation method proposed in JVET-AH0166 [1] was adopted into the VSEI TuC doc, i.e., JVET-AJ2032 [2], some model-adaptive modifications to the related syntax and semantics in the Film Grain Regions Characteristics (FGR) SEI message are proposed in this contribution to ensure the standard usage of adaptive film grain technology.

The proposed revised semantics of fgr\_single\_model\_flag have editorial issues and should be clarified. It was suggested to replace with the following:

When fgr\_single\_model\_flag is equal to 0, the fgr\_regions\_model\_id[ i ] syntax element is present. When fgr\_single\_model\_flag is equal to 1, the fgr\_model\_id syntax element is present.

The use of a fgrModelidMax variable in the for loop has an issue with extensibility. It is suggested to just use the value 2 instead.

This contribution is improving the text for a problem identified at the previous meeting.

Conditionally Adopt as modified by the above suggestions, subject to review of a revised contribution with these changes. Further discussion after revised contribution is available.

Further discussed 1930 on 19 Jan.

Confirmed adoption.

Decision: Adopt.

It was suggested by a participant that the single model design from AG2032 be also included as an option in the TuC, in addition to the multiple model design described in this contribution and included in AJ2032.

This seems unnecessary because there would be no changes from the design in AG2032, and if a decision was made to move the single model to a WD for VSEI, the AG2032 design could be used. If further modifications to the single model design were expected, it may be appropriate to add it as an additional option in the TuC.

[JVET-AK0211](https://jvet-experts.org/doc_end_user/current_document.php?id=15200) AHG9/AHG13: Proposed Film Grain Region SEI message for version 5 of VSEI [S. Wenger, G. Teniou, A. T. Hinds (Tencent)]

TBP in plenary.

[JVET-AK0281](https://jvet-experts.org/doc_end_user/current_document.php?id=15270) AHG9: On AI usage restrictions SEI message [M. M. Hannuksela, F. Cricri (Nokia)] [late]

Chaired by S. Deshpande on 18 January 2025 at 12:25 PM.

This document proposes the following items related to the AI usage restrictions SEI message:

1. It is asserted that AI usage restrictions of the cropped decoded pictures may differ from those of the post-processed video as described by a processing chain of an SEI processing order (SPO) SEI message. Thus, it is proposed to allow the AI usage restrictions SEI message in an SPO SEI message and provide an informative NOTE that describes the potential use of the AI usage restrictions SEI message for post-processed video.

It was asked if when AUR SEI is included in SPO, does it always needs to be in a PON? It was answered that it does not need to be in PON.

It was clarified that if AUR SEI is included but is not in SPO then that applies to the cropped decode picture. The only way to specify how AUR applies to post processed pictures is by including it in SPO.

The AUR in SPO can only be more restrictive that AUR outside SPO since the restrictions are of the type “Don’t do something”.

It was commented we need to address the following case: AUR is not in SPO and there is no AUR in SPO then the AUR from outside SPO should apply to output from processing chain. This should be added not as a note but as a statement of behavior.

Regarding combination of AUR with SPO it was asked if following case is allowed:

AUR SEI message says do not use for generative AI and this is followed by a NNPF which may do generative AI or GFV. It was commented that this should be disallowed. However, if those two SEIs in reverse order (i.e. AUR is later) then it is ok. It was commented it may be better to specify this in one common place, e.g. in AUR SEI – saying what it can be combined with.

Another place to describe this can be decided by checking the discussion at previous meeting regarding combinations of SEIs being allowed/ not.

Another comment was: generally the semantics of AUR SEI should be clarified as currently they give the impression that you are required to obey what the SEI message says should be done. But you cannot have a requirement on a decoder (especially a previous version decoder) controlled by a new SEI such as this. Thus, semantics should clarify what encoder is requesting and not what decoder has to do. Because of this using imperative wording such as “do not use it for xxx” should not be done. There may be other SEIs messages in TuC or even in VSEI draft text to which this comment may apply.

Agree to item 1 with revised text to be uploaded to handle the case to be addressed mentioned above.

1. It is asserted that there are use cases where AI usage restrictions apply to regions rather than complete pictures. Thus, it is proposed to indicate if AI usage restrictions apply to complete pictures or on the basis of bounding boxes indicated by the annotated regions SEI message. Two syntax options are proposed:
2. Add region indication type (aur\_region\_type[ i ]) for each AI usage restriction indicated in a AI usage restrictions SEI message
3. Add a single region indication type (aur\_region\_type) syntax element and allow multiple AI usage restrictions SEI messages to persist, differentiated by their region indication types.

It was commented that for described use case 1, 2 in the document, encoder could pre-process the pictures and then EOI may be used to signal this. It was answered that the proposed use case may be more for social media.

It was commented that adding region wise information to AUR may complicate the AUR envisioned usage.

It was commented that encoder can always just indicate that this applies to the entire picture and then there is no added complexity. But it was commented that the decoder needs to support receiving region wise information which has added complexity. It was commented that the decoder could still choose to apply the “do not xxx” restriction to complete picture instead of just to part of the picture.

Regarding use case 3 from the document, it was commented that the training algorithm could decide the quality of the video and decide not to use it. It was answered that for encoder to signal this information makes it less complex for the decoder/ training algorithm.

Further study encouraged.

[JVET-AK0286](https://jvet-experts.org/doc_end_user/current_document.php?id=15275) AHG9: Simplified enhanced colour format information SEI [J. Boyce, M. M. Hannuksela (Nokia)] [late]

Chaired by S. Deshpande on 16 January 2025 at 18:00 PM.

Two alternative approaches to support carriage of 4:4:4 content in a 4:2:0 or monochrome bitstream were added to the TuC in JVET-AJ2032. It is proposed to simplify the enhanced colour format information SEI originally proposed in JVET-AJ0215 to restrict the flexibility to only support carriage of chroma colour planes to form a 4:4:4 target picture. The proposed modified SEI message uses one or two auxiliary picture layers to carry the full resolution Cb and Cr components. The association between the auxiliary picture layer(s) and the primary picture layer is proposed to be determined from the parameters in the HEVC VPS or VVC/VSEI Scalability dimension information (SDI) SEI message, with new AuxId types defined.

The proponents of ECFI suggested removing alpha and depth (from semantics). ECFI design was done to align with HEIF. And the ECFI SEI syntax is one to one mapping to HEIF.

But alpha and depth may not be needed and can be removed by semantics (because we have aux id)

But the other functionality should be retained to be aligned with HEIF.

It was commented that the proposed removals don’t allow: RGB support and 4:2:2 support.

It was asked why ECFI has flexibility to put each colour component in any random place.

It was asked if in ECFI the regions are all of same size. The answer is yes. But it is asserted that the rectangle location may not be clearly specified. It was asked how 4:2:2 is supported if rectangles are same size. It was commented that luma an chroma need to be put in different layers.

4:2:2 can be supported using two layers with packing (Top-Bottom or side-by-side) or using three layers.

The group agrees to keep ECFI syntax as it is. Two aspects are left for revision of ECFI:

* Remove alpha and depth support from semantics
* Provide text for clearly determining the location of “rectangles” and mapping of sample locations.

Also describe in more details how 4:4:4 (and 4:2:2) formats are supported i.e. which channel if has which value.

Also, some constraints are needed to make sure each colour component occurs only once and the resulting representation is valid (Cr has such language).

The proponents of ECFI are planning to provide text for the above updates possibly at this meeting in a separate contribution. Since ECFI is TuC this is not extremely urgent but needs to be addressed.

## AHG9: Other SEI topics (6)

Contributions in this area were discussed at XXXX–XXXX on XXday XX Jan. 2025 (chaired by XXX).

[JVET-AK0107](https://jvet-experts.org/doc_end_user/current_document.php?id=15078) AHG9: Modality Information SEI for HEVC [J. Gao, H.-B. Teo, C.-S. Lim, K. Abe, V. Drugeon (Panasonic)]

The **modality information SEI message** is included in VSEI version 4 (Draft 4). The **modality information SEI message** provides information about the source of optical radiation used to generate pictures. Following the JVET-AJ meeting, WG5 has requested a new amendment to HEVC to include a new profile and SEI messages. Given the wide range of applications for infrared and ultraviolet imaging, incorporating such information would be beneficial to the legacy codec. It is proposed to add the **modality information SEI message** to the new edition of HEVC, with the interface text for integrating this SEI message included in the TuC for future VSEI extensions.

It is noted that the corresponding interface text for AVC is also included in the TuC. Additionally, it is suggested to move the AVC interface text to JVET-AK1017.

The proponents say that software is ready for submission for HM, but hasn’t been submitted yet.

It is noted that if support for the SEI message was added to AVC, software support in the JM would also be required.

Decision: Move HEVC interface text from TuC into the HEVC output document, JVET-AK1006, and adopt AVC interface text into AVC output document JVET-AK1017.

[JVET-AK0136](https://jvet-experts.org/doc_end_user/current_document.php?id=15107) AHG9: Target Colour Volume SEI message [C. H. Demarty, E. François, F. Aumont, O. Le Meur (InterDigital)]

See notes under JVET-AK0147.

[JVET-AK0142](https://jvet-experts.org/doc_end_user/current_document.php?id=15113) AHG9: Display rectangles SEI [J. Boyce, T. Biatek, M. M. Hannuksela (Nokia)]

A new SEI message is proposed for VSEI to describe multiple display rectangles, similar to the pan-scan rectangle SEI message defined in AVC and HEVC, but not included in VVC or VSEI.

The proposed display rectangles SEI message targets repurposing a coded video for a variety of video streaming platforms and devices that use different default display aspect ratios. The proposed SEI message includes signaling of the display aspect ratio of each display rectangle. Display rectangles may contain regions outside of the cropped decoded picture and a filling method may be indicated for those regions.

Discussed Mon. 20 Jan. 1020 (chaired by JRO).

It has similar functionality and could be interpreted as an extension of pan-scan, further signalling aspect ratio to have a simpler identification of the best recommended output window, and additionally allowing filling samples outside of the window by a value. Multiple sizes of output are definable, the most appropriate one to be selected at the display.

It was asked how this would be operated with SPO, e.g. if spatial extension is used by NNPF. Would it always be operated immediately before the display?

Decision: Adopt in TuC.

[JVET-AK0191](https://jvet-experts.org/doc_end_user/current_document.php?id=15162) AHG9: Examples of danmaku applications [J. Xu, Y.-K. Wang, L. Zhang (Bytedance)]

This contribution presents exemplar danmaku applications in video streaming and video archiving. To support these applications, a danmaku information SEI message, which is essentially the same as presented in JVET-AJ0072, is proposed.

For the archiving use case, it was suggested that the Danmuku display could alternatively be stored in a separate overlay video.

It was suggested that if text was available, a system could auto-translate to another language, but it would be unlikely to archive auto-translated content.

For a large number of Danmaku messages, the bitrate could be large.

The proposed semantics restrict di\_num\_danmaku\_minus1 shall be in the range of 0 to 232 – 2, which is very large.

A question was raised as to why this information should be at the video coding layer as it doesn’t describe the video content. Are there examples of other SEI messages that are similarly separate?

It was noted that T.35 is used to carry closed caption text in systems standards defined by CTA.

It may be useful to discuss with a larger audience about if this is within our scope.

The semantics for di\_start\_delta\_poc[ i ] and di\_end\_delta\_poc[ i ] seem inadequate because the POC value of the current picture is not known.

It was questioned if any other SEI message defined in VSEI require the use of POC.

Interface text is not provided in the contribution.

Do any of the VUI parameters apply to the colour signaled by di\_colour\_rgb[ i ] and the table for di\_colour\_idc[ i ]?

The semantics of di\_position[ i ] seem unclear. If multiple danmaku messages for the same POC value all indicate top, it is unclear where the decoder should display the individual messages.

In a follow-up discussion on Tuesday 21 Jan., it was commented that there is some overlap with general text description information SEI message, however positioning, timing and colour information is included in danmaku.

It was further commented that captioning mechanisms are often implemented at systems layer rather than using SEI messages (also using T.35).

It was further commented that in danmaku (Japanese word, danmu in Chinese) applications the text is behaving more dynamically (e.g. flying in, moving around) than in typical captioning systems.

As this kind of dynamic captioning may be closely related to the timing of video pictures, an SEI message could be justified beyond systems layer captioning. It was however commented that dynamic objects (including text) can also be overlaid at systems layer.

Several experts supported inclusion in TuC, but objection was also raised.

Further study on design aspects mentioned above, and seek for communication with WG 3 on an appropriate place, or similar technology would already be available.

[JVET-AK0203](https://jvet-experts.org/doc_end_user/current_document.php?id=15192) AHG9: On signalling segmentation and object tracking information [E. Thomas, E. Potetsianakis, E. Alexiou, M.-L. Champel (Xiaomi)]

Two known vision tasks are image segmentation and object tracking. In the image segmentation case, the task aims at labelling each pixel of an image as belonging to a category. In object tracking case, the task aims at localising in the image the presence of one or more objects.

With the relative affordability of running neural networks, there is currently more and more media devices on the market that are capable of processing images and video sequences to perform vision tasks. At the same time, the mobile ecosystems offer native APIs to application developer to perform those tasks. Example generated sequences are provided as attachment.

From the provided API examples for image segmentation and object tracking, it appears that there is a clear distinction between those two in terms of use cases and APIs. It is argued that there is no dependency on one and the other.

Therefore, it is argued that supporting both image segmentation and object tracking information is valuable. Additionally, it is also argued that having two distinct solutions for those two use cases is preferable.

The version 2 of this contribution adds a reference to AK0325-v1 [3] from another proponent which provides a corresponding signalling, i.e. based on a new auxiliary picture type and the associated SEI message for image segmentation.

This contribution seems to be providing requirements input for a use case, but does not include a technical solution. It recommends the SEI message proposed in JVET-AK0325.

It would be straightforward to generate masks for test content using available tools.

A segmentation mask could be carried in an auxiliary picture, with an SEI. For multiple masks for the same primary picture, multiple auxiliary pictures can be associated with the primary picture, with each auxiliary picture having a separate SEI message.

[JVET-AK0325](https://jvet-experts.org/doc_end_user/current_document.php?id=15314) New SEI message for Segmentation Information [D. Podborski, A.M. Tourapis (Apple)] [late]

During the 36th JVET meeting JVET-AJ0263 proposed introducing a new SEI message intended for segmentation applications. However, it was pointed out during the meeting that an alternative proposal, i.e. the Object Mask Information (OMI) SEI message, which was intended for object tracking applications, was already under consideration and could be reused for this purpose. The group recommended to the proponents to study the OMI SEI message and see how that could be modified or improved for both applications.

After careful study of the OMI SEI message, the proponents believe that the current syntax of the OMI SEI message is too complicated and may have other issues that make it less desirable for use in the applications envisioned for the SEI message presented in JVET-AJ0263. The OMI SEI also is more Object centric, where information is first indicated for particular objects that may be present in a scene. The corresponding mapping of the pixel values to the object ID is only a property of the object. However, the primary scope for the segmentation application is to identify the pixel mapping information of any objects in a scene, while the object information is not as relevant or could be provided through external means (e.g. at the system layer). Therefore, the proponents request considering including an improved version of the SEI message presented in JVET-AJ0263 for any inclusion in any appropriate standards completely separately, if possible from what is decided to do for the OMI SEI message.

2nd version:

* editorial corrections
* added a column for the purpose in Table F.14.3.13
* added a psi\_single\_component\_flag to be able to also use 3 components of the auxiliary picture if needed

There are editorial issues with the syntax table, where a syntax element is assigned a value.

The proponents assert that this SEI message is simpler than the OMI SEI message, which would be beneficial for some implementations.

Three values of psi\_type\_idc are proposed, which are used to map a pixel value to an ID value, full, implicit, and sparse.

A three colour component flag allows use of chroma values. 4:2:0 chroma format is allowed, with chroma upsampling needed at the decoder.

There is a problem in the syntax table for psi\_single\_component\_flag. A number of additional issues were identified.

It was noted that existing auxiliary picture types limit to single colour component, but a new type could be defined that does not have that limitation.

The proposed SEI message would be included in the auxiliary picture layer, rather than in the primary picture layer as in the OMI SEI message.

A question was raised about how the primary layer associated with the auxiliary picture layer is determined. The SDI SEI could be used for VVC.

In the FULL case, an ID could be mapped to more than one sample (or pixel) value. This could be used, for example, to combine segments as a post-process of the result from the segmentation algorithm.

A suggestion was made to consider splitting the OMI SEI message into separate messages for segmentation and for tracking. The proponents of this contribution and of the OMI SEI message do not support that.

Decision: Adopt to the TuC

## Non-SEI HLS aspects (0)

Kept as template for future use.

# Plenary meetings, joint meetings, BoG reports, and liaison communications

## General

The following topics were discussed in JVET plenary Sunday 19 Jan. 0900–1345:

* Scheduling for remaining week (further detail on scheduling is recorded in section 2.12).
* Report from EE/HLS tracks and follow-up discussion on documents XXXX.
* Review of sections 4.1 and 4.7 (add links)

Specific discussion on some general issues related to HLS was as follows:

It was asked what the process should be in supporting additional SEI messages by interfaces to AVC and HEVC. It is understood that for some recent cases (e.g. NNPF), it was not requested to contribute software. A request should be made for that, and for new SEI messages (e.g. digitally signed content) it should be required to be delivered at latest when DAM/DIS stage is reached.

For HEVC, new SEI interfaces should be put into JVET-AK1006

For AVC, a new JVET output document should be issued to collect interface text related to new VSEI messages. It is reasonable to have such a document to prevent “text losses”, as it occurred e.g. for SPTI where the interface text was lost when it was moved into VSEIv4.

Some SEI messages had been requested to be moved from TuC to VSEIv4.

A follow-up session on some open HLS issues was conducted Mon. 20 Jan. 0900-1200 (chaired by JRO). The general issues were discussed as follows:

**Promotion of SEI messages from TuC to VSEIv4 draft**

Criteria for moving SEI messages from TuC to VSEIv4 draft were agreed as follows:

* Maturity and stability in terms of syntax and semantics
* Software status
* Urgency of needing the message by end of this year
* Usefulness demonstrated by a showcase, implementability

SEI messages requested to be moved from TuC to VSEIv4 draft:

* AK0140 Packed regions SEI – not modified since the 36th meeting, software beyond parsing available in showcase, mature according to other experts’ opinion, could be referenced by the TR on machine analysis – decision to include in JVET-AK2006
* AK0114 AI usage restrictions SEI message – mature according to other experts’ opinion, complementary with copyright text information – decision to include in JVET-AK2006
* AK0178 Image Format Metadata – not modified since the 36th meeting, however the problem of potentially large payload was not resolved so far; issue of referencing CIPA spec should be resolved by October, otherwise it could not be included in final spec
* AK0204 Lens optical correction SEI – more clarification about syntax/semantics necessary, see discussion under JVET-AK0179
* AK0211 Film grain regions SEI – is stable for case of single model per picture, but dual model modifications made at the current meeting require more study; in recent meetings, concern had been raised about hardware implementability, that might require further investigation. It was also commented that moving regions in subsequent frames could be problematic after cropping.

Follow-up discussion on Tuesday 21 Jan.:

It is very simple and straightforward to resolve the size problem of Image Format Metadata by explicitly stating that the payload shall not be larger than 64 kbyte. Further study might be useful if this is appropriate for the purpose of intended applications.

Decision: Include IFM in JVET-AK2006.

It was agreed that by the next meeting, only minor necessary changes should be done on included SEI messages, or SEI messages might be removed that don’t have sufficient minimum software support (see below), no additional SEI messages to be included.

**SEI message software implementation status**

(Tables in this section to be updated based on input by K. Sühring)

VSEIv3

|  |  |  |  |
| --- | --- | --- | --- |
| **Message** | **Status** | | |
|  | **VTM** | **HM** | **JM** |
| NNPF | SE writing & parsing | not implemented\*) | not implemented\*) |
| Phase indication | SE writing & parsing | SE writing & parsing | not implemented\*) |
| Annotated regions | SE writing & parsing | SE writing & parsing | not implemented\*\*) |

\*) Proposed in JVET-AE0101 \*\*) Proposed in JVET-AxXXXX

When VSEIv3 was developed, interfaces of the above messages with AVC and HEVC were also defined without taking care that also minimum support for testing and implementing (by adding corresponding elements for writing and parsing those messages in the corresponding reference software) was provided by proponents.

As a possibility to enforce development of such software elements, it might be considered to remove support for those SEI messages from the corresponding specification texts unless such elements would be provided within a reasonable timeframe.

VSEIv4

|  |  |  |  |
| --- | --- | --- | --- |
| **Message** | **Status** | | |
|  | **VTM** | **HM** | **JM** |
| NNPF extension | SE writing & parsing | not implemented\*) | not implemented\*) |
| SPO | SE writing & parsing | N/A | N/A |
| PON | SE writing & parsing | N/A | N/A |
| EOI | SE writing & parsing | N/A | N/A |
| SPTI | SE writing & parsing | adopted at this meeting | adopted at this meeting |
| OMI | SE writing & parsing and output mask to text/file | N/A | N/A |
| Modality | SE writing & parsing | adopted at this meeting | adopted at this meeting |
| Text description | SE writing & parsing | N/A | N/A |
| GFV | SE writing & parsing | N/A | N/A |
| GFVE | missing | N/A | N/A |
| DSCI | generation & verification | adopted at this meeting | adopted at this meeting |
| DSCS | generation & verification | adopted at this meeting | adopted at this meeting |
| DSCV | generation & verification | adopted at this meeting | adopted at this meeting |
| PRI | adopted at this meeting | N/A | N/A |
| IFM | adopted at this meeting | N/A | N/A |
| AUR | adopted at this meeting | N/A | N/A |

\*) For HEVC and AVC, it could be assumed that the extensions would operate with a similar interface that was defined by adopting JVET-AE0101. However, no software was ever developed for this.

To avoid a situation as for v3, the following principles were discussed:

For any SEI message that would be promoted to VSEIv4, proponents are requested to issue a merge request for inclusion of software in VTM (and HM/JM in cases where it is also requested to add an interface with HEVC/AVC). This should at minimum be a software that allows including the syntax of the SEI message (including validity check e.g. for range) into a bitstream at the encoder and parsing it at the decoder.

For SEI messages marked as “N/A” for HEVC and AVC in the tables above, proponents are asked to think about usefulness and if deemed useful, to propose interface text and software implementation by the next meeting, which could then potentially be included in HEVC extension draft (JVET-AL1006) and AVC extension draft (JVET-AL1017).

Any SEI message for which no sufficiently mature software would be available by the time of ISO/IEC DAM/DIS approval, this will not be included in the DAM/DIS.

Proposing companies shall provide names of the persons (by sending an email to Karsten Sühring) responsible for providing a software that matches the most recent text until the end of the meeting (revisit).

|  |  |  |  |
| --- | --- | --- | --- |
| **Message** | **Responsible person** | | |
|  | **VVC** | **HEVC** | **AVC** |
| NNPF extensions |  |  |  |
| SPO |  | N/A | N/A |
| PON |  | N/A | N/A |
| EOI | Chul Keun Kim (LGE)) | N/A | N/A |
| SPTI | Sean McCarthy (Dolby) | Sean McCarthy (Dolby) | Sean McCarthy (Dolby) |
| OMI | Jie Chen (Alibaba) | N/A | N/A |
| Modality | Jingying Gao (Panasonic) | Jingying Gao (Panasonic) | Jingying Gao (Panasonic) |
| Text description |  | N/A | N/A |
| GFV |  | N/A | N/A |
| GFVE | Jie Chen (Alibaba) | N/A | N/A |
| DSCI | Karsten Sühring (HHI) | Karsten Sühring (HHI) | Karsten Sühring (HHI) |
| DSCS | Karsten Sühring (HHI) | Karsten Sühring (HHI) | Karsten Sühring (HHI) |
| DSCV |  |  |  |
| PRI | Jill Boyce (Nokia) | N/A | N/A |
| IFM |  | N/A | N/A |
| AUR | Chul Keun Kim (LGE)) | N/A | N/A |

It is pointed out that not all functionalities of an SEI message can be checked by just creating and parsing a valid syntax. If it is important for an SEI message to be operated in different modes, it would be desirable to have some mechanisms to check that as well.

For some SEI messages additional software elements exist that at the encoder automatically generate the syntax dynamically, send it in the bitstream and drive some processing after decoding. Such pieces of software are typically not included in the reference software.

**TuC of previous meeting**

It was reported that contribution JVET-AH0167 was forgotten to be included in JVET-AH2032. To simplify processing, the JVET chair had suggested the editors to generate a new version of JVET-AJ2032 (v3). This is agreed to be used as a basis for JVET-AK2032.

Joint meetings involving JVET were scheduled as follows:

* …

Further detail about joint sessions with other groups is provided in the subsection 7.3.

Break-out group reports are discussed in section 7.4.

General plenary wrap-up discussions are recorded under sections 8, 9, and 10.

## Information sharing meetings

(kept for future use in meetings under SC 29 auspices)

Information sharing sessions with other WGs and AGs of the MPEG community were held on Monday 15 July 0900–1200, Wednesday 17 July 0900–1000, and Friday 19 July 1400–1600.

The status and plans for the work in the MPEG WGs and AGs was reviewed at these information sharing sessions.

## Joint meetings

### Joint session XXXX-XXXX XXday X Jan. on XXXX: MPEG WG X / XXXX, MPEG AG 5 / Quality Assessment, MPEG WG 5 / JVET and VCEG (ITU-T Q6/21)

(These notes were recorded by XXX)

Chaired by XXX (XXX Convenor/Rapporteur), … .

## BoGs (0)

The following break-out groups were established at this meeting to conduct discussion and develop recommendations on particular subjects (retained for future use).

## Liaison communications (X) (update)

[m70439](https://dms.mpeg.expert/doc_end_user/current_document.php?id=96222&id_meeting=200) Liaison statement from SC 29/WG 1 to WG 5 on JPEG AI [WG 1 via SC 29 Secretariat]

The liaison response WG 5 N 324 was reviewed in JVET on Thursday 7 Nov. at 1445-1500. The draft reply was also presented in the MPEG AG 3 Communication meeting Thursday 7 Nov. at 1500-1800.

# Project planning

## Software timeline (update)

ECM 15.0 software (including all adoptions) was planned to be available 3 weeks after the meeting (29 November).

The NNVC 11.0 codebase software was planned to be available 3 weeks after the meeting (29 November).

Extensions on top of VTM23.5 software will be released as appropriate (e.g., integration and updates of SEI messages incuded in JVET-AJ2006 by the current meeting).

Updates on top of HM18.0 and HTM16.3 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 (update)

An EE on neural network-based video coding was established, as recorded in output document JVET-AJ2023.

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-AJ2024.

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 X Jan. 2025 at XXXX–XXXX.

Chairs of AHGs are asked to send draft mandates to JRO before 2000 on 21 Jan, preferably copy from the table below and sending in a word file with changemarks.

|  |  |  |
| --- | --- | --- |
| **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-AJ2005 and JVET-AJ2006). * Collect reports of errata for the VVC, VSEI, HEVC, AVC, CICP, and the published related technical reports and produce the JVET-AJ1004 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 (for both VTM and HM), 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 tests for VVC multi-layer features, and update the test plan accordingly. * 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. * Coordinate with AHG17 on investigating sequences outside of CTC, and making arrangements for viewing at the next meeting. * 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) | N |
| **Conformance testing (AHG5)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Prepare for a new edition of VVC conformance with new multi-layer streams (JVET-AI2028) and corrections of the existing streams. * Study the draft conformance bitstreams for new HEVC multiview profiles in JVET-AI1008, and further develop related conformance bitstreams. * Coordinate with AHG3 on implementation of the new HEVC multiview profiles. * 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, T. Ikai, S. Iwamura, H.-J. Jhu, K. Kawamura, P. de Lagrange, 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-16.0 software version (and potential updates), corresponding VTM anchor, 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, such as the aspects of memory access and bandwidth, number of maximum processing cycles, block decoding dependencies, number of context coded bins, pipeline and parallelization. * 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. * Coordinate with AHG17 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. * Develop methodology of more reliable runtime measurement | X. Li (chair), L.-F. Chen, Z. Deng, J. Gan, E. François, R. Ishimoto, H.-J. Jhu, J. Lainema, X. Li, J. Pardo, A. Stein, H. Wang (vice‑chairs) | Y (tel., 2 weeks notice) |
| **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. * Discuss improvements on the evaluation framework, including evaluation procedures and methodologies. * Coordinate software development and experiments on optimization of encoders and receiving systems for machine analysis of coded video content, including combinations of proposed technologies. * Maintain the software implementation examples and develop tool combination examples 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 JVET-AK2030 on optimization of encoders and receiving systems for machine analysis of coded video content. * Study the potential of using SEI messages for the purpose of machine analysis in coordination with AHG9. * Investigate the impact of using different machine task models in the evaluation of the compression performance of tools optimized for machine analysis tasks. * Coordinate with WG 4 VCM AHG on aspects such as unified common test conditions, evaluation metrics, test and training materials, usage of SEI messages, and on studying characteristics and requirements of machine analysis tasks, etc. Provide WG 4 VCM with AHG software and scripts to experiment non-normative tools and tool combination examples, and provide bitstreams generated by running a dense set of QP points for each sequence by uploading to the mpeg.expert repository. | S. Liu, J. Ström, S. Wang, M. Zhou (AHG chairs) | Y (tel., 2 weeks notice, joint with WG 4 AHG) |
| **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. * Study JVET-AK1006, JVET-AK2005, and JVET-AK2006, identify any issues and propose solutions as appropriate. * Study JVET-AK2032 and JVET-AK1017, and propose improvements. * Collect software and showcase information for SEI messages, including encoder and decoder implementations and bitstreams for demonstration and testing. * 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 for JM, HM, and VTM. | S. McCarthy, Y.-K. Wang (co-chairs), J. Boyce, T. Chujoh, S. Deshpande, C. Fogg, M. M. Hannuksela, P. de Lagrange, G. J. Sullivan, H. Tan, 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, and coordinate with AHG4 on improving encoders and test settings for multi-layer verification testing. * Study optimized encoding and suitable test settings for noisy materials, such as sequences containing film grain. * Study optimized encoding and tool combinations for low latency and for 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, higher objective quality, or lower [encoding ?] complexity, 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) | Y (tel., 2 weeks notice) |
| **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. * Update the test conditions for NN-based video coding in JVET-AJ2016, including list, location and md5sums for training set. 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. * Discuss and propose improved metrics to perform complexity analysis of NN architectures, in particular also considering the impact of sparse models. * Finalize and discuss the EE on neural network-based video coding. * Promote the call for training materials, distribute it, and actively communicate with content owners. * Coordinate with other groups, including SC 29/AG 5 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 and studying the impact of training set extension. | 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 Nov. 28, second on Dec. 20 |
| **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 ECM15 algorithm description JVET-AJ2025. * 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-AI2017. * Analyse the results of exploration experiments described in JVET-AJ2024 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. * Discuss and propose refinements to the draft of the TR 2nd ed. JVET-AJ2020. * Study alternative film grain models and their associated documentation. * Discuss and enumerate updates, improvements, and additions for the second edition of the technical report. * 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 plan for subjective quality testing of the FGC SEI message JVET-AJ2022, 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-11.0 software version (and potential updates), and provide reference configuration encodings according to the NNVC common test conditions as described in JVET-AJ2016. Study the impact of the addition of new dataset on the already integrated models. * Investigate and bridge coding performance of NNVC VTM anchor compared to the latest VTM version and/or VTM anchor in ECM software. * 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) and Very Low Operation Point (HOP/LOP/VLOP) 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-AJ2019) 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), R. Chang, Y. Li, Y. Li, M. Santamaria, J. N. Shingala, Z. Xie (vice chairs) | Y (tel., 2 weeks notice), first on Nov. 28, second on Dec. 20 |
| **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. * Produce VTM and ECM anchor encodings according to CTC JVET-AJ2027, and provide test results at the next meeting. * Develop and maintain interfaces for supporting use cases of camera parameters and depth maps in gaming applications, including mechanisms for efficient transporting these elements in the coded video bitstream. * Evaluate JVET test models (such as ECM, VTM, NNVC, etc.) under the proposed test conditions. * Investigate possibilities to enhance compression capability for gaming content. * Study conversion of depth maps using integer representation, and identifying efficient bit-depth resolution of depth maps to support identified use-cases that will be an input to compression. * Solicit contributions from industry on typical bitrate/quality/resolution used for gaming content compression. | S. Puri, J. Sauer (co-chairs), R. Chernyak, A. Duenas, L. Wang, V. Zakharchenko (vice chairs) | Y (tel., 2 weeks notice) |
| **Generative face video compression (AHG16)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Maintain GFVC software tools, associated configuration files, and software usage documentation. * Study the extension of GFVC software capability to handle video content with higher resolutions. * Identify and study additional test content, including content with higher resolutions, suitable for use in GFVC performance evaluation. * Study GFVC performance under test conditions defined in JVET-AJ2035, as well as performance on additional test content, and wider bitrate ranges. * Coordinate with AHG9 on further development of the GFV and GFVE SEI messages in JVET-AJ2006. | Y. Ye (chair), H.-B. Teo, Z. Lyu, S. McCarthy, S. Wang (vice chairs) | N |
| **Testing of video coding technology beyond CTC (AHG17)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Identify test sequences extending the represented range of applications (i.e. beyond those defined in common test conditions documents JVET-AB2010, JVET-AC2011 and JVET-AI2017) and investigate their suitability for use in visual quality assessments. * Prepare encoded bitstreams from the identified test sequences using VTM23.5 and ECM14.1 * Make preparations for viewing at the 37th JVET meeting in coordination with AHG4 and AG 5. * Develop the output document JVET-AJ2026, starting from the BoG report JVET-AJ0378. * Investigate potential need of arranging a face-to-face AHG meeting (with remote participation for group discussions) between the 37th and 38th JVET meetings. | M. Wien (chair), Y. Ye, V. Baroncini, E. Alshina (vice chairs) | Y (tel., 2 weeks notice, first on Nov. 20, second on Dec. 18) |
| **Ultra low latency and packet loss resilience (AHG18)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Investigate and identify test conditions, evaluation criteria and evaluation methodology, supporting a set of end-to-end latency targets in a range of 25-100ms. * Investigate creation of practical simulation software based on VTM, including network transmission aspects, and conduct performance evaluation * Identify potential requirements and feasibility of standard based technologies to support ultra-low delay requirements, including packet loss resilient decoding * Investigate technologies beyond VVC towards ultra-low delay coding for interactive and live broadcasting scenarios | S. Ikonin, S. Wenger, V. Zakharchenko (co-chairs), S. Deshpande, J. Ström, X. Ma, N. Song, C. Kim, S. Puri, S. Fößel (vice-chairs) | Y (tel., 2 weeks notice), first Feb 20 (?) |

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 339) 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 339, as noted in section 9.

[JVET-AK1000](https://jvet-experts.org/doc_end_user/current_document.php?id=14988) Meeting Report of the 37th JVET Meeting [J.-R. Ohm] [WG 5 N 333] (2025-02-19)

Initial versions of the meeting notes (d0 … dX) 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-AH1003](https://jvet-experts.org/doc_end_user/current_document.php?id=14259) Coding-independent code points for video signal type identification (Draft 3) [G. J. Sullivan, A. Tourapis]

Primary editor: G. J. Sullivan.

[JVET-AK1004](https://jvet-experts.org/doc_end_user/current_document.php?id=14989) Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP [Y.-K. Wang, B. Bross, I. Moccagatta, C. Rosewarne, G. J. Sullivan] (2025-03-14, near next meeting)

Primary editor: Y.-K. Wang.

This includes changes from new bug tickets.

Remains valid – not updated: [JVET-AH1005](https://jvet-experts.org/doc_end_user/current_document.php?id=14261) Technology under consideration for future editions of CICP [E. Thomas, A. Tourapis] [WG 5 N 289)]

From JVET-AH0154 and JVET-AH0217.

[JVET-AK1006](https://jvet-experts.org/doc_end_user/current_document.php?id=14990) HEVC extensions and corrections (draft 2) [Y.-K. Wang, B. Bross, S. Deshpande, G. J. Sullivan, A. Tourapis ] [WG 5 CDAM N xxx)] (2025-01-29)

Primary editor: Y.-K. Wang.

Remains valid – not updated: [JCTVC-V1007](https://mpeg.expert/jct/files/JCTVC-V1007-v1.zip) 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]

Remains valid: [JVET-AI1008](https://jvet-experts.org/doc_end_user/current_document.php?id=14609) Conformance testing for HEVC multiview extended and monochrome profiles [I. Moccagatta, S. Paluri, A. Tourapis, Y.-K. Wang]

Remains valid – not updated: [JVET-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](https://mpeg.expert/jct/files/JCTVC-O1010-v1.zip) Guidelines for Conformance Testing Bitstream Preparation [T. Suzuki, W. Wan]

Remains valid – not updated: [JVET-AJ1011](https://jvet-experts.org/doc_end_user/current_document.php?id=14991) White paper on HEVC [B. Bross, J.-R. Ohm, G. J. Sullivan, Y.-K. Wang] [AG 3 N 174]

Remains valid – not updated: JVET-[AJ1012](https://jvet-experts.org/doc_end_user/current_document.php?id=14992) Overview of IT systems used in JVET [J.-R. Ohm, I. Moccagatta, K. Sühring, M. Wien]

Requires update for new test sequence download server, and new JCT-VC/-3V sites

Remains valid – not updated: [JCT3V-G1003](https://mpeg.expert/jct3v/files/JCT3V-G1003-v2.zip) 3D-AVC Test Model 9 [D. Rusanovskyy, F. C. Chen, L. Zhang, T. Suzuki] [WG 11 N 14239]

Remains valid – not updated: [JCT3V-K1003](https://mpeg.expert/jct3v/files/JCT3V-K1003-v1.zip) 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](https://mpeg.expert/jct/files/JCTVC-V1014-v1.zip) 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-AH1016](https://jvet-experts.org/doc_end_user/current_document.php?id=14264) AVC with extensions and corrections (draft 3) [B. Bross, T. Ikai, G. J. Sullivan, A. Tourapis, Y.-K. Wang]

Primary editor: B. Bross.

[JVET-AK1017](https://jvet-experts.org/doc_end_user/current_document.php?id=14264) Support for additional VSEI messages in AVC (draft 1) [J. Boyce, B. Bross, Y.-K. Wang] (2025-02-14)

Primary editor: B. Bross

No output: JVET-Axx1018 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-AH2002](https://jvet-experts.org/doc_end_user/current_document.php?id=14265) Algorithm description for Versatile Video Coding and Test Model 22 (VTM 22) [Y. Ye, A. Browne, S. Kim] [WG 5 N 284]

Primary editor: Y. Ye.

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-AJ2003](https://jvet-experts.org/doc_end_user/current_document.php?id=14993) Guidelines for VTM-based software development [F. Bossen, X. Li, K. Sühring]

Update from JVET-AJ0187

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]

[JVET-AK2005](https://jvet-experts.org/doc_end_user/current_document.php?id=14994) Additions and corrections for VVC version 4 (Draft 11) [G. J. Sullivan, B. Bross, M. M. Hannuksela, Y.-K. Wang] [WG 5 preliminary DAM update N 337)] (2025-03-14)

Also issue Draft DoC WG 5 N 341 (with editing period 2025-03-14)

Agreed changes are … Not yet to be sent for DAM ballot.

Primary editor: G. J. Sullivan.

[JVET-AK2006](https://jvet-experts.org/doc_end_user/current_document.php?id=14995) Additional SEI messages for VSEI version 4 (Draft 5) [J. Boyce, J. Chen, S. Deshpande, M. M. Hannuksela, S. McCarthy, G. J. Sullivan, H. Tan, Y.-K. Wang] [WG 5 preliminary DAM N 335)] (2025-03-14)

Also issue Draft DoC WG 5 N 340 (with editing period 2025-03-14)

New elements from notes elsewhere in this report:

* General editorial improvements
  + …
* SEI Processing Order SEI
  + …
* Neural Network Post Filter
  + …
* Encoder optimization information SEI
  + …
* Text description information SEI
  + …
* Source picture timing SEI
  + …
* Generative Face video SEI
  + …
* Digitally signed content SEIs
  + …

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.

Primary editor: J. Boyce.

Remains valid – not updated: [JVET-AJ2007](https://jvet-experts.org/doc_end_user/current_document.php?id=14996) 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]

Number to be re-used for 3rd edition.

Remains valid – not updated: [JVET-AJ2009](https://jvet-experts.org/doc_end_user/current_document.php?id=14997) Reference software for versatile video coding 2nd edition (Draft 2) [F. Bossen, K. Sühring, X. Li] [WG 5 DIS N 322)]

Software can be found at <https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/-/tree/2nd-edition>.

Primary editor: F. Bossen.

[JVET-AK2010](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] (2025-xx-xx)

Add alternative configurations (see discussion under JVET-AK0181)

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-AJ2016](https://jvet-experts.org/doc_end_user/current_document.php?id=14998) Common test conditions and evaluation procedures for neural network-based video coding technology [E. Alshina, F. Galpin, R.-L. Liao, S. Liu, A. Segall]

Remains valid – not updated: [JVET-AI2017](https://jvet-experts.org/doc_end_user/current_document.php?id=14615) 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-AK2019](https://jvet-experts.org/doc_end_user/current_document.php?id=14999) Description of algorithms version 10 and software version 12 in neural network-based video coding (NNVC) [F. Galpin, Y. Li, D. Rusanovskyy, J. Ström, L. Wang] (2025-03-14)

New elements in text and software 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.

[JVET-AK2020](https://jvet-experts.org/doc_end_user/current_document.php?id=15000) Film grain synthesis technology for video applications ed. 2 (Draft 2) [W. Husak, P. de Lagrange, A. Norkin, A. Tourapis] (2025-03-14)

first thoughts on other content to be added (see from JVET-AK0013), new: JVET-AK0169

Remains valid – not updated: [JVET-AJ2021](https://jvet-experts.org/doc_end_user/current_document.php?id=15001) Verification test plan for VVC multilayer coding (update 5) [O. Chubach, P. de Lagrange, M. Wien]

Remains valid – not updated: [JVET-AJ2022](https://jvet-experts.org/doc_end_user/current_document.php?id=15002) Plan for subjective quality testing of the FGC SEI message (update 4) [P. de Lagrange, W. Husak, M. Radosavljević, M. Wien]

[JVET-AK2023](https://jvet-experts.org/doc_end_user/current_document.php?id=14986) 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 (EE coordinators)] (2025-02-xx)

An initial draft of this document was reviewed and approved at XXXX-XXXX on Wednesday 22 Jan.

This round of EE1 tests includes:

* **…**

[JVET-AK2024](https://jvet-experts.org/doc_end_user/current_document.php?id=14987) 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 (EE coordinators)] (2025-02-xx)

An initial draft of this document was reviewed and approved at XXXX-XXXX on Wednesday 22 Jan.

This round of EE2 tests will include:

* **…**

[JVET-AK2025](https://jvet-experts.org/doc_end_user/current_document.php?id=15003) Algorithm description of Enhanced Compression Model 16 (ECM 16) [M. Coban, R.-L. Liao, K. Naser, J. Ström, L. Zhang] (2025-03-14)

New elements from notes elsewhere in this report:

* Intra prediction:
  + …
* Inter prediction:
  + …
* Transforms and coefficient coding:
  + …
* In-loop filtering:
  + …
* Software/CTC:
  + …

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-AK2026](https://jvet-experts.org/doc_end_user/current_document.php?id=15004) Testing of video coding technology beyond conditions of exploration experiments [M. Wien, Y. Ye, V. Baroncini, E. Alshina] (2025-02-14)**

to be finalzed in AHG telco

Remains valid – not updated: [JVET-AJ2027](https://jvet-experts.org/doc_end_user/current_document.php?id=15005) Common test conditions for gaming applications [J. Sauer, R. Chernyak, S. Puri, S. Thiebaud]

Remains valid – not updated: [JVET-AI2028](https://jvet-experts.org/doc_end_user/current_document.php?id=14621) Additional conformance bitstreams for VVC multilayer configurations [S. Iwamura, P. de Lagrange, I. Moccagatta]

3rd edition should become JVET-AL2008 by next meeting, the JVET-Ax2028 can be re-used

Request for 3rd edition WG 5 N 338 was reviewed Wednesday 22 Jan. XXXX, with delivery date 2025-01-24.

Remains valid – not updated: [JVET-AH2029](https://jvet-experts.org/doc_end_user/current_document.php?id=14274) Visual quality comparison of ECM/VTM encoding [V. Baroncini, J.-R. Ohm, M. Wien] [AG 5 N 118]

[JVET-AK2030](https://jvet-experts.org/doc_end_user/current_document.php?id=15006) Optimization of encoders and receiving systems for machine analysis of coded video content (Draft 8, future updates) [S. Liu, J. Chen, J. Ström] (2025-03-14)

Primary editor: S. Liu. 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-AI2031](https://jvet-experts.org/doc_end_user/current_document.php?id=14623) Common test conditions for optimization of encoders and receiving systems for machine analysis of coded video content [S. Liu, C. Hollmann]

[JVET-AK2032](https://jvet-experts.org/doc_end_user/current_document.php?id=15007) Technologies under consideration for future extensions of VSEI (version 7) [S. McCarthy, J. Boyce, J. Chen, S. Deshpande, M. M. Hannuksela, H. Tan, Y.-K. Wang] (2025-03-14)

New elements from notes elsewhere in this report:

* General editorial improvements
  + …
* Constituent rectangles SEI
  + …
* Display overlays SEI
  + …
* Image Format Metadata SEI
  + …
* Packed regions information SEI
  + …
* Bitdepth range information SEI
  + …
* AI usage restrictions SEI
  + …
* Lens optical correction SEI
  + …
* Quality metrics SEI
  + …
* Photosenstive content information SEI
  + …
* Graphics rendering information SEI
  + …
* Neural Network Post Filter
  + …
* Derived chroma formats
  + …

It was agreed during approval of this document that the following removals are also to be implemented to the TuC:

* …

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 agreed that MPI is not further changed, and will be removed upon completion of a corresponding MIV profile.

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)]

Remains valid – not updated: [JVET-AI2034](https://jvet-experts.org/doc_end_user/current_document.php?id=14624) Call for new HDR materials for future video coding development [E. François, W. Husak, S. Iwamura, D. Rusanovskyy, A. Segall, M. Wien] [WG 5 N 312)]

Remains valid – not updated: [JVET-AJ2035](https://jvet-experts.org/doc_end_user/current_document.php?id=15008) Test conditions and evaluation procedures for generative face video coding [S. McCarthy, B. Chen]

Remains valid – not updated: [JVET-AG2036](https://jvet-experts.org/doc_end_user/current_document.php?id=13921) Call for training materials for neural network-based video coding tool development [E. Alshina, F. Galpin, S. Liu, M. Wien] [WG 5 N 266)]

Remains valid – not updated: [JVET-AJ2037](https://jvet-experts.org/doc_end_user/current_document.php?id=15009) Report on subjective quality testing of the FGC SEI message (AG 5 N 140) [P. de Lagrange, W. Husak, M. Wien] [AG 5 N 140)] (2024-12-31)

[JVET-AK2038](https://jvet-experts.org/doc_end_user/current_document.php?id=15006) White paper on VSEI [J. Boyce, S. McCarthy, Y.-K. Wang] (2025-03-21)

JVET-AK2039 Announcement of JVET AHG17 Meeting, Aachen, Germany, March 11-13 [J.-R. Ohm, M. Wien] (2025-02-07)

Date to be confirmed, alternative 10-12

The draft of WG 5 recommendations was reviewed and approved in JVET at XXXX-XXXX on Wednesday 22 January.

# Future meeting plans, expressions of thanks, a.o.b., and closing of the meeting

Future meeting plans were established according to the following guidelines (assuming face-to-face meetings):

* Meeting under ITU-T SG21 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 SG21 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. 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 7 – 15 October 2025, 40th meeting under ITU-T SG21 auspices in Geneva, CH,
* During 14 – 23 January 2026, 41st meeting under ISO/IEC JTC 1/‌SC 29 auspices, to be conducted as teleconference meeting,
* During 24 April – 1 May 2026, 42nd meeting under ISO/IEC JTC 1/‌SC 29 auspices in Santa Eulària, ES,
* During 7 – 15 July 2026, 43rd meeting under ITU-T SG21 auspices in Geneva, CH.
* During October 2026, 44th meeting under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.
* During January 2027, 45th meeting under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.

A face-to-face meeting of AHG17 (with best-effort online access to discussions) may be conducted during March 10-14, 2025 in Aachen, Germany (estimated duration of 3 days), with the purpose of identifying non-CTC test cases for subjective investigation. This is to be decided by the 37th JVET meeting.

The agreed document deadline for the 38th JVET meeting was planned to be Wednesday 19 March 2025.

(update thanks)

TCL was thanked for offering new HDR test material, and Bytedance was thanked for offering test material showing user generated content.

TCL and Vestel were thanked for providing 4K displays, Fraunhofer HHI was thanked for providing play-out equipment used in the experts viewing in Kemer. Adam Wieckowski was thanked for helping with the test setup. The experts who volunteered to participate in the viewing were also thanked.

Marius Preda was thanked for maintaining the document site jvet-experts.org, and for re-installing access to the documents of JCT-VC and JCT-3V. Institut Mines-Télécom was thanked for hosting the sites.

Fraunhofer HHI and RWTH Aachen University were thanked for hosting the new JVET content sites.

Ali Begen was thanked for planning, preparing and hosting the 36th JVET meeting in Kemer. Mustafa Bay, Alev Yavuz, Basak Erel and Merve Dağlı of Dekon Congress and Tourism, and the staff of Mirage Park Resort Hotel were thanked for the excellent support during the meeting. Further thanks were expressed to the silver sponsors Ofinno, Perculus and Unified Streaming, bronze sponsors Dolby, Ericsson, Nokia, and Qualcomm, and supporters Kuru Kahveci Mehmet Efendi, Ozyegin University, TCL and Vestel.

The 37th JVET meeting was closed at approximately XXX hours CET on Wednesday 22 January 2025.

# Annex A to JVET report: List of documents

(Dates and times in the table below are in Paris/Geneva time (1 hr. ahead of UTC).

# Annex B1 to JVET report: List of meeting participants attending in person

The participants who were personally present at the meeting site of the thirty-seventh meeting of the JVET, according to a sign-in sheet circulated in the JVET meeting rooms (approximately XXX people in total), were as follows:

1. …

# Annex B2 to JVET report: List of meeting participants attending remotely

The remote participants of the thirty-sieventh 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 had attended the meeting in person at least part-time (see annex B1), and not including those who attended only the joint sessions with other groups), were as follows:

1. …

# Annex C to JVET report: Recommendations of the 18th meeting of ISO/IEC JTC 1/SC 29/WG 5 MPEG Joint Video Experts Team with ITU-T SG 21

**ISO/IEC JTC 1/SC 29/WG 5 N XXX**