|  |  |
| --- | --- |
| **Joint Video Experts Team (JVET)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29**  31st Meeting, Geneva, CH, 11–19 July 2023 | Document: JVET-AE\_notes\_d4 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Meeting Report of the 31st Meeting of the Joint Video Experts Team (JVET), Geneva, CH, 11–19 July 2023** | | |
| *Status:* | Report document from the chair of JVET | | |
| *Purpose:* | Report | | |
| *Author(s) or Contact(s):* | **Jens-Rainer Ohm** Institute of Communication Engineering RWTH Aachen Melatener Straße 23 D-52074 Aachen | Tel: Email: | +49 241 80 27671 [ohm@ient.rwth-aachen.de](mailto:ohm@ient.rwth-aachen.de) |
| *Source:* | Chair of JVET | | |

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

# Summary

The Joint Video Experts Team (JVET) of ITU-T WP3/16 and ISO/IEC JTC 1/‌SC 29 held its thirty-first meeting during 11–19 July 2023 at the ITU premises in Geneva, Switzerland. The meeting was held as a face-to-face meeting, but 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 twelfth 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 CEST (UTC+2) on Tuesday 11 July 2023. Meeting sessions were held on all days, including the weekend days of Saturday and Sunday 15 and 16 July 2023, until the meeting was closed at approximately XXXX hours CEST on Wednesday 19 July 2023. 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), 14 AHG reports, 2 EE summary reports, X BoG reports, and X incoming liaison documents were discussed. The meeting took place in a collocated fashion with a meeting of SG16 – 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 with partial temporal overlap – where WG 5 is representing the Joint Video Coding Team(s) and their 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 thirtieth JVET meeting in producing the following documents:

* [JVET-AD1003](https://jvet-experts.org/doc_end_user/current_document.php?id=12566) Coding-independent code points for video signal type identification (Draft 2 of 3rd edition), also issued as WG 5 preliminary FDIS N 206
* [JVET-AD1004](https://jvet-experts.org/doc_end_user/current_document.php?id=12567) Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP
* [JVET-AD1008](https://jvet-experts.org/doc_end_user/current_document.php?id=12568) Additional colour type identifiers for AVC and HEVC (Draft 4), also issued as WG 5 preliminary WD N 200
* [JVET-AD1012](https://jvet-experts.org/doc_end_user/current_document.php?id=12569) Overview of IT systems used in JVET
* [JVET-AD2002](https://jvet-experts.org/doc_end_user/current_document.php?id=12572) Algorithm description for Versatile Video Coding and Test Model 20 (VTM 20)
* [JVET-AD2005](https://jvet-experts.org/doc_end_user/current_document.php?id=12574) New level and systems-related supplemental enhancement information for VVC (Draft 5)
* [JVET-AD2006](https://jvet-experts.org/doc_end_user/current_document.php?id=12574) Additional SEI messages for VSEI (Draft 4)
* [JVET-AD2007](https://jvet-experts.org/doc_end_user/current_document.php?id=12575) Guidelines for NNVC software development
* [JVET-AD2016](https://jvet-experts.org/doc_end_user/current_document.php?id=12576) Common test conditions and evaluation procedures for neural network-based video coding technology
* [JVET-AD2017](https://jvet-experts.org/doc_end_user/current_document.php?id=12576) Common test conditions and evaluation procedures for enhanced compression tool testing
* [JVET-AD2019](https://jvet-experts.org/doc_end_user/current_document.php?id=12563) Description of algorithms and software in neural network-based video coding (NNVC) version 3
* [JVET-AD2021](https://jvet-experts.org/doc_end_user/current_document.php?id=12578) Verification test plan for VVC multilayer coding
* [JVET-AD2022](https://jvet-experts.org/doc_end_user/current_document.php?id=12579) Draft plan for subjective quality testing of FGC SEI message
* [JVET-AD2023](https://jvet-experts.org/doc_end_user/current_document.php?id=12560) Exploration experiment on neural network-based video coding (EE1)
* [JVET-AD2024](https://jvet-experts.org/doc_end_user/current_document.php?id=12562) Exploration experiment on enhanced compression beyond VVC capability (EE2)
* [JVET-AD2025](https://jvet-experts.org/doc_end_user/current_document.php?id=12580) Algorithm description of Enhanced Compression Model 9 (ECM 9)
* [JVET-AD2027](https://jvet-experts.org/doc_end_user/current_document.php?id=12582) SEI processing order SEI message in VVC (Draft 4), also issued as WG 5 preliminary WD N 203
* [JVET-AD2028](https://jvet-experts.org/doc_end_user/current_document.php?id=12583) Additional conformance bitstreams for VVC multilayer configurations
* [JVET-AD2030](https://jvet-experts.org/doc_end_user/current_document.php?id=12564) Optimization of encoders and receiving systems for machine analysis of coded video content (draft 2), also issued as WG 5 preliminary WD N 199
* [JVET-AD2031](https://jvet-experts.org/doc_end_user/current_document.php?id=12584) Common test conditions for optimization of encoders and receiving systems for machine analysis of coded video content

As main results, the JVET produced XX output documents from the current meeting (to be updated):

* [JVET-AD1003](https://jvet-experts.org/doc_end_user/current_document.php?id=12566) Coding-independent code points for video signal type identification (Draft 2 of 3rd edition), also issued as WG 5 preliminary FDIS N 206
* [JVET-AD1004](https://jvet-experts.org/doc_end_user/current_document.php?id=12567) Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP
* [JVET-AD1008](https://jvet-experts.org/doc_end_user/current_document.php?id=12568) Additional colour type identifiers for AVC and HEVC (Draft 4), also issued as WG 5 preliminary WD N 200
* [JVET-AD1012](https://jvet-experts.org/doc_end_user/current_document.php?id=12569) Overview of IT systems used in JVET
* [JVET-AD2002](https://jvet-experts.org/doc_end_user/current_document.php?id=12572) Algorithm description for Versatile Video Coding and Test Model 20 (VTM 20)
* [JVET-AC2003](https://jvet-experts.org/doc_end_user/current_document.php?id=12573) Guidelines for VTM-based software development
* [JVET-AD2005](https://jvet-experts.org/doc_end_user/current_document.php?id=12574) New level and systems-related supplemental enhancement information for VVC (Draft 5)
* [JVET-AD2006](https://jvet-experts.org/doc_end_user/current_document.php?id=12574) Additional SEI messages for VSEI (Draft 4)
* [JVET-AD2007](https://jvet-experts.org/doc_end_user/current_document.php?id=12575) Guidelines for NNVC software development
* [JVET-AD2016](https://jvet-experts.org/doc_end_user/current_document.php?id=12576) Common test conditions and evaluation procedures for neural network-based video coding technology
* [JVET-AD2017](https://jvet-experts.org/doc_end_user/current_document.php?id=12576) Common test conditions and evaluation procedures for enhanced compression tool testing
* [JVET-AD2019](https://jvet-experts.org/doc_end_user/current_document.php?id=12563) Description of algorithms and software in neural network-based video coding (NNVC) version 3
* [JVET-AD2021](https://jvet-experts.org/doc_end_user/current_document.php?id=12578) Verification test plan for VVC multilayer coding
* [JVET-AD2022](https://jvet-experts.org/doc_end_user/current_document.php?id=12579) Draft plan for subjective quality testing of FGC SEI message
* [JVET-AD2023](https://jvet-experts.org/doc_end_user/current_document.php?id=12560) Exploration experiment on neural network-based video coding (EE1)
* [JVET-AD2024](https://jvet-experts.org/doc_end_user/current_document.php?id=12562) Exploration experiment on enhanced compression beyond VVC capability (EE2)
* [JVET-AD2025](https://jvet-experts.org/doc_end_user/current_document.php?id=12580) Algorithm description of Enhanced Compression Model 9 (ECM 9)
* [JVET-AD2027](https://jvet-experts.org/doc_end_user/current_document.php?id=12582) SEI processing order SEI message in VVC (Draft 4), also issued as WG 5 preliminary WD N 203
* [JVET-AD2028](https://jvet-experts.org/doc_end_user/current_document.php?id=12583) Additional conformance bitstreams for VVC multilayer configurations
* [JVET-AD2030](https://jvet-experts.org/doc_end_user/current_document.php?id=12564) Optimization of encoders and receiving systems for machine analysis of coded video content (draft 2), also issued as WG 5 preliminary WD N 199
* [JVET-AD2031](https://jvet-experts.org/doc_end_user/current_document.php?id=12584) Common test conditions for optimization of encoders and receiving systems for machine analysis of coded video content

The following documents were produced as WG 5 documents only:

* WG 5 N 197 Draft disposition of comments received on ISO/IEC 23002-7:2022 DAM 1
* WG 5 N 198 Preliminary text of ISO/IEC FDIS 23002-7:202x Versatile supplemental enhancement information messages for coded video bitstreams (3rd edition)
* WG 5 N 201 Draft disposition of comments received on ISO/IEC 23090-3:2022 DAM 1
* WG 5 N 202 Preliminary text of ISO/IEC FDIS 23090-3:202x Versatile video coding (3rd edition)
* WG 5 N 205 Draft disposition of comments received on ISO/IEC DIS 23091-3:202X (Video CICP 3rd edition)

For the organization and planning of its future work, the JVET established XX “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 13 – 20 October 2023 under ISO/IEC JTC 1/‌SC 29 auspices in Hannover, DE; during 17 – 26 January 2024 under ISO/IEC JTC 1/‌SC 29 auspices, to be conducted as teleconference meeting; during April 2024 under ITU-T SG16 auspices, date and location t.b.d.; during 12 – 19 July 2024 under ISO/IEC JTC 1/‌SC 29 auspices in Sapporo, JP; during October 2024 under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.; during January 2025 under ITU-T SG16 auspices, date and location t.b.d.; during April 2025 under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.; and during 26 June – 4 July 2025 under ISO/IEC JTC 1/‌SC 29 auspices in Daejeon, KR.

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

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

# Administrative topics

## Organization

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

The Joint Video Experts Team (JVET) of ITU-T WP3/16 and ISO/IEC JTC 1/‌SC 29 held its thirty-first meeting during 11–19 July 2023 at the ITU premises in Geneva, Switzerland. The meeting was held as a face-to-face meeting, but 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 twelfth meeting as WG 5. The JVET meeting was held under the chairmanship of Dr Jens-Rainer Ohm (RWTH Aachen/Germany).

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

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

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

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

This report contains three important annexes, as follows:

* Annex A contains a list of the documents of the JVET meeting
* Annex B contains a list of the meeting participants, 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/2023_07_AE_Geneva/>.

## Primary goals

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

* [JVET-AD1003](https://jvet-experts.org/doc_end_user/current_document.php?id=12566) Coding-independent code points for video signal type identification (Draft 2 of 3rd edition), also issued as WG 5 preliminary FDIS N 206
* [JVET-AD1004](https://jvet-experts.org/doc_end_user/current_document.php?id=12567) Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP
* [JVET-AD1008](https://jvet-experts.org/doc_end_user/current_document.php?id=12568) Additional colour type identifiers for AVC and HEVC (Draft 4), also issued as WG 5 preliminary WD N 200
* [JVET-AD1012](https://jvet-experts.org/doc_end_user/current_document.php?id=12569) Overview of IT systems used in JVET
* [JVET-AD2002](https://jvet-experts.org/doc_end_user/current_document.php?id=12572) Algorithm description for Versatile Video Coding and Test Model 20 (VTM 20)
* [JVET-AD2005](https://jvet-experts.org/doc_end_user/current_document.php?id=12574) New level and systems-related supplemental enhancement information for VVC (Draft 5)
* [JVET-AD2006](https://jvet-experts.org/doc_end_user/current_document.php?id=12574) Additional SEI messages for VSEI (Draft 4)
* [JVET-AD2007](https://jvet-experts.org/doc_end_user/current_document.php?id=12575) Guidelines for NNVC software development
* [JVET-AD2016](https://jvet-experts.org/doc_end_user/current_document.php?id=12576) Common test conditions and evaluation procedures for neural network-based video coding technology
* [JVET-AD2017](https://jvet-experts.org/doc_end_user/current_document.php?id=12576) Common test conditions and evaluation procedures for enhanced compression tool testing
* [JVET-AD2019](https://jvet-experts.org/doc_end_user/current_document.php?id=12563) Description of algorithms and software in neural network-based video coding (NNVC) version 3
* [JVET-AD2021](https://jvet-experts.org/doc_end_user/current_document.php?id=12578) Verification test plan for VVC multilayer coding
* [JVET-AD2022](https://jvet-experts.org/doc_end_user/current_document.php?id=12579) Draft plan for subjective quality testing of FGC SEI message
* [JVET-AD2023](https://jvet-experts.org/doc_end_user/current_document.php?id=12560) Exploration experiment on neural network-based video coding (EE1)
* [JVET-AD2024](https://jvet-experts.org/doc_end_user/current_document.php?id=12562) Exploration experiment on enhanced compression beyond VVC capability (EE2)
* [JVET-AD2025](https://jvet-experts.org/doc_end_user/current_document.php?id=12580) Algorithm description of Enhanced Compression Model 9 (ECM 9)
* [JVET-AD2027](https://jvet-experts.org/doc_end_user/current_document.php?id=12582) SEI processing order SEI message in VVC (Draft 4), also issued as WG 5 preliminary WD N 203
* [JVET-AD2028](https://jvet-experts.org/doc_end_user/current_document.php?id=12583) Additional conformance bitstreams for VVC multilayer configurations
* [JVET-AD2030](https://jvet-experts.org/doc_end_user/current_document.php?id=12564) Optimization of encoders and receiving systems for machine analysis of coded video content (draft 2), also issued as WG 5 preliminary WD N 199
* [JVET-AD2031](https://jvet-experts.org/doc_end_user/current_document.php?id=12584) Common test conditions for optimization of encoders and receiving systems for machine analysis of coded video content

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

## Documents and document handling considerations

### General

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

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

The document registration and upload times and dates listed in Annex A and in headings for documents in this report are in Paris/Geneva time. Dates mentioned for purposes of describing events at the meeting follow the TRT timezone (local time in Antalya), 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.
* Some decisions are recorded with the word “agreed” rather than “Decision:”, especially for 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, 4 July 2023. Any documents uploaded after 1159 hours Paris/Geneva time on Wednesday 5 July 2023 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-AE0183 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 meetings, 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-AE0XXX (a proposal on …), uploaded 07-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-AE0XXX (a document discussing …), uploaded 07-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-AE0045, … .

The cross-verification reports JVET-AE0XXX and … had still been missing by the end of the meeting on Wednesday 19 July, but became available one or two days later. The following cross-verification reports were still missing three weeks after the end of the meeting: JVET-AE0XXX. These were thus marked as withdrawn by the JVET chair (this case did not happen at the current meeting, sentence kept for future use).

“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 with documents JVET-AE0XXX and … , which were categorized as late in the list above, based on the time of the first reasonable document upload.

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

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.

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.

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-AD1000, the Coding-independent code points for video signal type identification (Draft 2 of 3rd edition) JVET-AD1003, the Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP JVET-AD1004, the Additional colour type identifiers for AVC, HEVC and Video CICP (Draft 4) JVET-AD1008, the Overview of IT systems used in JVET [JVET-AD1012](https://jvet-experts.org/doc_end_user/current_document.php?id=12569), the Algorithm description for Versatile Video Coding and Test Model 20 (VTM 20) JVET-AD2002, the New level and systems-related supplemental enhancement information for VVC (Draft 5) JVET-AD2005, the Additional SEI messages for VSEI (Draft 4) JVET-AD2006, the Guidelines for NNVC software development JVET-AD2007. the Common test conditions and evaluation procedures for neural network-based video coding technology JVET-AD2016, the Common test conditions and evaluation procedures for enhanced compression tool testing JVET-AD2017, the Description of algorithms and software in neural network-based video coding (NNVC) version 3 JVET-AD2019, the Verification test plan for VVC multilayer coding JVET-AD2021, the Draft plan for subjective quality testing of FGC SEI message JVET-AD2022, the Description of the EE on Neural Network-based Video Coding JVET-AD2023, the Description of the EE on Enhanced Compression beyond VVC capability JVET-AD2024, the Algorithm description of Enhanced Compression Model 9 (ECM 9) JVET-AD2025, the SEI processing order SEI message in VVC (Draft 4) JVET-AD2027, the Additional conformance bitstreams for VVC multilayer configurations JVET-AD2028, the Optimization of encoders and receiving systems for machine analysis of coded video content (Draft 2) [JVET-AD2030](https://jvet-experts.org/doc_end_user/current_document.php?id=12564), and the Common test conditions for optimization of encoders and receiving systems for machine analysis of coded video content [JVET-AD2031](https://jvet-experts.org/doc_end_user/current_document.php?id=12584), had been completed and were approved. The software implementations of VTM versions 20.1, 20.2, and 21.0, ECM (versions 9.0 and 9.1), and NNVC (versions 5.0 and 5.1) were also approved.

Only minor editorial issues were found in the meeting report JVET-AD1000; no need to produce an update was identified (see section 2.14 for details).

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

## Attendance

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

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

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

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

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

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

## 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 0830-1900 CEST 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.iecapc.jp/F/IEC_Code_of_Conduct.pdf>

These include points relating to:

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

## IPR policy reminder

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

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

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

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

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

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

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

## Software copyright disclaimer header reminder

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

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

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

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

## Communication practices

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

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

It was emphasized that reflector subscriptions and email sent to the reflector must use real names when subscribing and sending messages and subscribers must respond to inquiries regarding the nature of their interest in the work. The current number of subscribers on the JVET email list was 1215 (as of 10 July 2023). All discussions (including those on AVC, HEVC, VVC, CICP, etc.) shall be conducted on the JVET reflector rather than any of the old reflectors (including JVT, JCT-VC, and JCT-3V) which are retained for archiving purposes.

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

It is further emphasized that the document JVET-AD1012 gives valuable hints about communication practices as well as other IT resources used in JVET, such as software, conformance, and test materials.

## Terminology

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

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

* MPEG-2 | H.262 (coding specification is common text)
  + ITU-T H.262 V3 was approved in 2012-02; Amd.1 was approved in 2013-03 and was not published separately; it was instead incorporated directly into the V3 text and published 2013-09
  + ISO/IEC 13818-2:2013 V3 FDIS ballot closed 2012-05-08; FDAM 1 ballot closed 2013-04-12 and was not published separately; it was instead incorporated directly into the V3 text and published 2013-10
  + Conformance testing (not joint with ITU-T)
    - ISO/IEC 13818-4:2004 V2 FDIS closed 2004-08-22, published 2004-12-12; it specifies conformance testing for Part 1 (Systems), Part 2 (Video), Part 3 (Audio), and Part 7 (AAC)
    - ISO/IEC 13818-4:2004/Amd 3:2009 Level for 1080@50p/60p conformance testing
    - Cor 1:2007, Cor 2:2009, Cor 3:2012, Cor 4:2011 may also have video relevance
  + Reference software (not joint with ITU-T)
    - ISO/IEC TR 13818-5:2005 V2 FDIS closed 2005-07-24, published 2005-10; it specifies reference software for Part 1 (Systems), Part 2 (Video), Part 3 (Audio), Part 7 (AAC) and Part 11 (IPMP)
* AVC (twin text)
  + ITU-T H.264 V14 was Consented at 22nd meeting on 2021-04-30 (with annotated regions, shutter interval, and miscellaneous corrections), approved 2021-08-22, published 2021-10-13
  + ISO/IEC 14496-10:2020 (Ed. 9) FDIS ballot closed 2020-11-27, published 2020-12-15
  + ISO/IEC 14496-10:2022 (Ed. 10), had been forwarded from DIS directly for publication 2022-01-21 (with annotated regions, shutter interval, and miscellaneous corrections) with an editing period, submitted to ITTF in 2022-05 after consultation with ISO staff on format of graphics files, upgraded to “DIS approved for registration” in ISO Project system 2022-07-04, published 2022-11-07
  + Preliminary draft text for YCgCo-Re and YCgCo-Ro issued at 26th meeting, second draft including SMPTE ST 2128 issued at 28th meeting 2022-10, third draft issued at 29th meeting 2023-01, fourth draft issued at 30th meeting 2023-04 (not yet formally requested as a project)
  + Conformance testing (twin text)
    - ITU-T H.264.1 V6 Approved 2016-02-13, published 2016-06-17
    - Various amendments of ISO/IEC 14496-4:2004, including:
      * ISO/IEC 14496-4:2004/AMD 6:2005 Advanced Video Coding conformance
      * ISO/IEC 14496-4:2004/AMD 9:2006 AVC fidelity range extensions conformance
      * ISO/IEC 14496-4:2004/AMD 30:2009 Conformance testing for new profiles for professional applications
      * ISO/IEC 14496-4:2004/AMD 31:2009 Conformance testing for SVC profiles
      * ISO/IEC 14496-4:2004/AMD 38:2010 Conformance testing for Multiview Video Coding
      * ISO/IEC 14496-4:2004/AMD 41:2014 Conformance testing of MVC plus depth extension of AVC
      * ISO/IEC 14496-4:2004/AMD 42:2014 Conformance testing of Multi-Resolution Frame Compatible Stereo Coding extension of AVC
      * ISO/IEC 14496-4:2004/AMD 43:2015 3D-AVC conformance testing
      * ISO/IEC 14496-4:2004/AMD 45:2016 Conformance Testing for the Multi-resolution Frame Compatible Stereo Coding with Depth Maps Extension of AVC
  + Reference software (twin text)
    - ITU-T H.264.2 V7 Approved 2016-02-13, published 2016-05-30
    - Various amendments of ISO/IEC 14496-5:2001 have been published, including:
      * ISO/IEC 14496-5:2001/AMD 6:2005 Advanced Video Coding (AVC) and High Efficiency Advanced Audio Coding (HE AAC) reference software
      * ISO/IEC 14496-5:2001/AMD 8:2006 AVC fidelity range extensions reference software
      * ISO/IEC 14496-5:2001/AMD 15:2010 Reference software for Multiview Video Coding
      * ISO/IEC 14496-5:2001/AMD 18:2008 Reference software for new profiles for professional applications
      * ISO/IEC 14496-5:2001/AMD 19:2009 Reference software for Scalable Video Coding
      * ISO/IEC 14496-5:2001/AMD 33:2015 Reference software for MVC plus depth extension of AVC
      * ISO/IEC 14496-5:2001/AMD 34:2014 Reference software of the multi-resolution frame compatible stereo coding of AVC
      * ISO/IEC 14496-5:2001/AMD 35:2015 3D-AVC Reference software
      * ISO/IEC 14496-5:2001/AMD 39:2016 Reference software for the Multi-resolution Frame Compatible Stereo Coding with Depth Maps of AVC
      * ISO/IEC 14496-5:2001/AMD 42:2017 Reference software for the alternative depth information SEI message extension of AVC
* HEVC (twin text)
  + ITU-T H.265 V7 approved 2019-11-29, published 2020-01-10
  + ISO/IEC 23008-2:2020 (Ed. 4) FDIS ballot closed 2020-07-16, published 2020-08-27
  + ITU-T H.265 V8 Consented at the 22nd meeting (shutter interval information SEI message and miscellaneous corrections), published 2020-10-13
  + ISO/IEC 23008-2:2020/AMD 1:2021 (shutter interval information SEI message) published 2021-07-12
  + ISO/IEC 23008-2:202x (Ed. 5) began as CDAM 2 High-range levels output of 25th meeting of January 2022, CDAM ballot closed 2022-04-15, conversion to 5th edition with miscellaneous corrections planned at 26th meeting of 2022-04, text submitted for DIS ballot 2022-07-10, DIS ballot closed 2023-01-10, FDIS issued 29th meeting of 2023-01, waiting for FDIS ballot (ready for Consent at this meeting)
  + Preliminary draft 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 (not yet formally requested as a project)
  + Conformance testing (twin text)
    - ITU-T H.265.1 V3 approved 2018-10-14, published 2019-01-15
    - ISO/IEC 23008-8:2018 (Ed. 2) Conformance specification for HEVC, published 2018-08-06
    - ISO/IEC 23008-8:2018/AMD 1:2019 Conformance testing for HEVC screen content coding (SCC) extensions and non-intra high throughput profiles, published 2019-10-15
  + Reference software (twin text)
    - ITU-T H.265.2 V4 approved 2016-12-22, published 2017-04-10
    - ISO/IEC 23008-5:2017 (Ed. 2) Reference software for high efficiency video coding, published 2017-03-01
    - ISO/IEC 23008-5:2017/AMD 1:2017 Reference software for screen content coding extensions, published 2017-11-09
* VVC (twin text)
  + ITU-T H.266 V1 approved 2020-08-29, published 2020-11-10
  + ISO/IEC 23090-3:2021 (Ed. 1) published 2021-02-16
  + ITU-T H.266 V2 with operation range extensions, Consented 2022-01-28, Last Call began 2022-04-01, Approved 2022-04-29, pre-published 2022-06-06, published 2022-07-12
  + ISO/IEC 23090-3:2022 (Ed. 2) with operation range extensions, approval at WG level to proceed to FDIS 2022-01-21, published 2022-09-25
  + ISO/IEC 23090-3:202x (Ed. 2) / Amd.1 New level and systems-related supplemental enhancement information, CDAM 1 issued from 26th meeting, ballot closed 2022-07-14, DAM 1 issued from 27th meeting, ballot closed 2023-01-03 (ready for FDIS and Consent at this 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 V1 approval at WG level to proceed to FDIS 2022-10-15, upgraded to “DIS approved for registration” in ISO Projects system 2021-10-24, upgraded to “FDIS registered for formal approval” 2022-07-11, FDIS ballot closed 2022-11-04, published 2022-11-24
    - ISO/IEC 23090-15:202x Amd.1 Operation range extensions – DAM 1 issued from 25th meeting 2022-01-21, upgraded to “CD approved for registration as DIS” status in ISO Projects system 2022-05-31, upgraded to “DIS registered” 2022-06-22, DAM ballot closed 2022-11-15, consolidated into FDIS issued as an output of the 29th meeting in January 2023 (remaining under preparation at the time of this meeting, ready for Consent at this meeting)
  + Reference software (twin text)
    - ITU-T H.266.2 V1 Consented 2022-01-28, Last Call began 2022-04-01, Approved 2022-04-29, pre-published 2022-05-17, published 2022-07-12
    - ISO/IEC 23090-16:2022 V1 approval at WG level to proceed to FDIS 2022-01-21, upgraded to “DIS approved for registration” status in ISO Projects system 2022-04-21, upgraded to “FDIS registered for formal approval” 2022-04-22, FDIS ballot initiated 2022-07-24, FDIS ballot closed 2022-09-19, published 2022-10-23
* VSEI (twin text)
  + ITU-T H.274 V1 approved 2020-08-29, published 2020-11-10
  + ISO/IEC 23002-7:2021 (Ed. 1) published 2021-01-28
  + ITU-T H.274 V2 Consented 2022-01-28, Last Call began 2022-04-01, Approved 2022-05-22 (after 1 Last Call comment and Additional Review), pre-published 2022-06-17, published 2022-07-25
  + ISO/IEC 23002-7:2022 (Ed. 2) approval at WG level to proceed to FDIS 2022-01-21, upgraded to “DIS approved for registration” status in ISO Projects system 2022-05-05 and “FDIS registered for formal approval” 2022-05-08, FDIS ballot closed 2022-09-27, published 2022-10-30
  + ISO/IEC 23002-7:202x (2nd Ed.) Amd.1 Request for new edition and CD for additional SEI messages issued at 27th meeting, ballot closed 2022-10-10, DAM registered 2022-11-13, DAM ballot closed 2022-04-06: comments in [m62571](https://dms.mpeg.expert/doc_end_user/current_document.php?id=86620&id_meeting=194) from Finland, Japan, US, ISO CS (minor) (ready for FDIS and Consent at this 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 (ready for FDIS and Consent at this meeting)
* Conversion and coding practices for HDR/WCG Y′CbCr 4:2:0 video with PQ transfer characteristics (twin text)
  + H.Sup15 V1, approved 2017-01-27, published 2017-04-12
  + ISO/IEC TR 23008-14:2018 published 2018-08
* Signalling, backward compatibility and display adaptation for HDR/WCG video coding (twin text)
  + H.Sup18 V1, approved 2017-10-27, published 2018-01-18
  + ISO/IEC TR 23008-15:2018 published 2018-08
* Usage of video signal type code points (twin text)
  + H.Sup19 V3 approved 2021-04-30, published 2021-06-04
  + ISO/IEC TR 23091-4 (Ed. 3) published 2021-05-23
* Working practices using objective metrics for evaluation of video coding efficiency experiments (twin text)
  + HSTP-VID-WPOM V1: approved 2020-07-03, published 2020-11
  + ISO/IEC TR 23002-8 (Ed. 1) published 2021-05-20
* Film grain synthesis technologies for video applications (twin text)
  + ISO/IEC TR 23002-9 Request for subdivision and WD 1 issued at 25th meeting 2022-01-21, WD 2 issued at 27th meeting, WD 3 issued at 28th meeting, CDTR issued at 29th meeting 2023-07, consultation period ended 2023-07-09 (ready for DTR at this meeting)
* The following freely available standards are published here in ISO/IEC:  
  <https://standards.iso.org/ittf/PubliclyAvailableStandards/index.html> as of 2023-05-26.
  + ISO/IEC 13818-4:2004 Conformance for MPEG-2
  + ISO/IEC 13818-4:2004/Amd 3:2009 Level for 1080@50p/60p conformance testing
  + ISO/IEC TR 13818-5:2005 Software simulation for MPEG-2
  + Various amendments of ISO/IEC 14496-4:2004 Conformance for AVC
  + Various amendments of ISO/IEC 14496-5:2001 Reference software for AVC
  + ISO/IEC 14496-10:2022 (Ed. 10) AVC
  + ISO/IEC 23002-7:2022 (Ed. 2) – VSEI
  + ISO/IEC 23008-2:2020 (Ed. 4) HEVC
  + ISO/IEC 23090-3:2022 (Ed. 2) VVC
  + ISO/IEC 23090-15:2022 (Ed. 1) Conformance for VVC
  + ISO/IEC 23090-16:2022 (Ed. 1) Reference software for VVC
  + ISO/IEC 23091-2:2021 (Ed. 2) Video CICP
* The following standards that have been intended by JVET to be publicly available were not available at <https://standards.iso.org/ittf/PubliclyAvailableStandards/index.html> as of 2023-05-26. (These should be checked for previously issued requests for free availability.)
  + ISO/IEC 23008-2:2020 (Ed. 4) Amd.1:2021: Shutter interval information SEI message, published 2021-07-12 (has not been requested)
  + ISO/IEC 23008-5:2017 (Ed. 2) Reference software for high efficiency video coding, published 2017-03-01
  + ISO/IEC 23008-5:2017/AMD 1:2017 Reference software for screen content coding extensions, published 2017-11-09
  + ISO/IEC 23008-8:2018 (Ed. 2) Conformance specification for HEVC, published 2018-08, published 2018-08-06
  + ISO/IEC 23008-8:2018/AMD 1:2019 Conformance testing for HEVC screen content coding (SCC) extensions and non-intra high throughput profiles, published 2019-10-15
* It appears necessary to check if all older software and conformance packages are publicly available – it might be that it was never requested, e.g. for those that were produced by JCT-3V. This topic was left TBD until the next meeting – perhaps it would be best to compile a list of all relevant software and conformance parts of AVC, HEVC, MPEG-2 aka H.262, CICP, and request these in bulk.

## Draft standards progression status

* AVC and HEVC colour type indicators for YCgCo-Re, YCgCo-Ro, and SMPTE ST 2128 (IPT-PQ-C2) are drafted and pending formal action.
* HEVC new levels (from JVET-Z1005) – ISO/IEC 23008-2 DIS of new edition of HEVC was issued from the April 2022 26th meeting, incorporating Amd.1 and corrigenda items (ballot closed 2023-01-10, ballot comments in the Summary of Voting document [m61834](https://dms.mpeg.expert/doc_end_user/current_document.php?id=85619&id_meeting=193)); note that Amd.1 = shutter interval SEI is already included in latest ITU-T edition of H.265. It is noted that there are potential additional items (corrigenda+tickets, YCgCo-Re and YCgCo-Ro draft, SMPTE ST 2128, multiview profiles draft) where only corrigenda items were included in the FDIS text based on ballot comments, ballot had not been started yet. ITU-T consent for a new edition is planned for July 2023. It was noted that the referencing of VSEI is also somewhat different in the ITU-T and ISO/IEC versions of HEVC and/or AVC, which might be aligned at the next convenient time (basically editorial – e.g., the ITU version of AVC specifies the annotated regions SEI message without referencing VSEI, whereas the ISO/IEC version references VSEI for the syntax and semantics of that SEI message). However, there is currently no other need for HEVC to reference the VSEI standard. An FDIS for HEVC was issued as an output of the 29th meeting in January 2023 (and it does not reference VSEI). Ready for Consent at this meeting.
* VVC new level and systems-related supplemental enhancement information (from JVET-AA2005) – VVC DAM was issued from 27th meeting, ballot closed 2023-01-03, ballot comments in the Summary of Voting document [m61833](https://dms.mpeg.expert/doc_end_user/current_document.php?id=85618&id_meeting=193). This was converted into a preliminary FDIS of VVC 3rd edition ([WG 5 N 183](https://dms.mpeg.expert/doc_end_user/current_document.php?id=86365&id_meeting=193)) at the 29th meeting of January 2023, anticipating that some alignment will be necessary with the ongoing VSEI amendment. ITU-T consent is planned for July 2023 when a new edition of VSEI is also planned to be consented, to keep VVC and VSEI aligned.
* VVC Conformance testing for operation range extensions – (from JVET-Y2026) – the DAM ballot closed 2022-11-15 (ballot comments in the Summary of Voting document [m61832](https://dms.mpeg.expert/doc_end_user/current_document.php?id=85617&id_meeting=193)), and this was consolidated into an FDIS at the 29th meeting, but the ballot had not been started yet. ITU-T consent was planned for July 2023.
* VSEI additional SEI messages (from JVET-AB2006) – VSEI DAM (JVET draft 3) was issued from the 28th meeting and a DAM ballot was issued, Summary of Voting document is available as [m62571](https://dms.mpeg.expert/doc_end_user/current_document.php?id=86620&id_meeting=194). It was planned to be consolidated into FDIS of new edition in April (if possible) or July, with ITU-T consent in July 2023. There were numerous input contributions in particular on NNPF, so the FDIS was deferred to July 2023.
* Film grain synthesis technology for video applications – JVET draft 4 was issued at the 29th meeting (JVET-AC2020), and the ISO/IEC 23002-9 CDTR was issued (a request to start work on the TR had been made at the 25th meeting), and the CDTR consultation period ended 2023-07-09 (ready for DTR at this meeting; it was noted that a second DTR could become necessary in case of comments). The publication limit date was reportedly 2023-08-09, so action to extend that date may be needed. No action in ITU-T at this meeting.
* Video CICP new edition with for YCgCo-Re and YCgCo-Ro (from JVET-Z1003), an ISO/IEC 23091-2 preliminary FDIS was issued from the 30th meeting and the Summary of Voting document was available as [m62572](https://dms.mpeg.expert/doc_end_user/current_document.php?id=86621&id_meeting=194), and the draft DoC in XXXX. A delay might be considered depending on status of SMPTE ST 2128, which was tentatively included in the preliminary FDIS, based on an NB comment.Video CICP colour type indicator for SMPTE ST 2128 is drafted and incorporated into the preliminary FDIS issued at the 30th meeting of 2023-04. It was reported that it is expected to become finalized in the SMPTE meeting in March. FDIS and ITU-T consent are planned for July 2023.
* 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 11 July at 0930 CEST were as follows.

* Timing and organization of the meeting and online access, calendar posting of session plans
  + The initial number of documents was lower than for the previous meeting (200->160) – parallel sessions might nevertheless be necessary, considering the preparation of various deliverables.
  + Scheduling of NNVC discussions – might be alignement with NNPF discussions, and starting from July 15 with JPEG-AI (JPEG meeting starts only next week and finishes after JVET)
* Plans for subsequent F2F meeting in October (Hannover), April (??), and July 2024 (Sapporo).
* The meeting logistics, agenda, working practices, policies, and document allocation considerations were reviewed.
  + Remote access to the meeting was provided using Zoom. This required discipline in the meeting room (no microphone to be switched on, podium and room microphones to be under central control).
  + Having text and software available is crucial (and not just arriving at the end of the meeting).
  + There were no objections voiced in the opening plenary to the consideration of late contributions.
* The results of the previous meeting and the meeting report JVET-AD1000 were reviewed. The following small issues in the meeting report were noted and were not considered sufficient to warrant issuing a revision. These are obviously left over from a previous report, and the correct information can be found in other places of the report:
  + The title of the document designates the meeting as teleconference meeting, whereas it had been held in face-to-face mode in Antalya, TR.
  + In the summary (section 1), the meeting of April 2025 was expressed to be held under ITU-T SG16 auspices, whereas regulary it should be expected to be held under ISO/IEC SC 29 auspices.
  + In the summary (section 1), JVET-AC2003 is wrongly listed as output from the meeting. The corresponding line should have been removed, and the total number of outputs was 20 instead of 21.
  + In 2.4.3 outputs of preceding meeting, document JVET-AD2027 (SEI processing order SEI message) was designated to be draft 2, whereas actually it had already been the third draft. Further below, a reference to section 2.12 about minor errors in the previous meeting report should have been referencing section 2.14 (opening remarks).
  + In section 8.2 (experiment planning), documents JVET-AC2023 and JVET-AC2024 are referred to instead of JVET-AD2023 and JVET-AD2024, respectively.
* There was somewhat less of a problem of late non-cross-check documents.
* There were only few documents registered where authors’ given names were not abbreviated, and/or company affiliation was missing in the authors’ list. Participants were reminded to stick to JVET’s conventions.
* Experts are 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.
* The primary goals of the meeting were:
  + 2nd edition of VVC confomance (H.266.1) is originally planned to be delivered for ITU consent at this meeting, but the corresponding JVET document JVET-AC2026 and WG 5 FDIS 23090-15 N 185 have not been produced yet. To be clarified from AHG5 report.
  + Any action on a new version of VVC software as standards part – target April 2024, for inclusion of NNPF SEI, software from TRs, layered coding, bug fixes, etc.
  + VVC and VSEI third editions FDIS, and consent on next ITU editions
    - VVC DAM ballot response in [m61833](https://dms.mpeg.expert/doc_end_user/current_document.php?id=85618&id_meeting=193), draft DoCR (from last meeting) in [MDS22705](https://dms.mpeg.expert/doc_end_user/current_document.php?id=87851&id_meeting=194), and preliminary FDIS text in [MDS22706](https://dms.mpeg.expert/doc_end_user/current_document.php?id=87852&id_meeting=194).
    - VSEI DAM ballot response in [m62571](https://dms.mpeg.expert/doc_end_user/current_document.php?id=86620&id_meeting=194), draft DoCR (from last meeting) in [MDS22702](https://dms.mpeg.expert/doc_end_user/current_document.php?id=87848&id_meeting=194), and preliminary FDIS text in [MDS22703](https://dms.mpeg.expert/doc_end_user/current_document.php?id=87849&id_meeting=194).
  + New edition video CICP FDIS (DIS ballot response in [m62572](https://dms.mpeg.expert/doc_end_user/current_document.php?id=86621&id_meeting=194) draft DoCR in [MDS22710](https://dms.mpeg.expert/doc_end_user/current_document.php?id=87856&id_meeting=194), and preliminary FDIS text in [MDS22711](https://dms.mpeg.expert/doc_end_user/current_document.php?id=87857&id_meeting=194)), and consent for next ITU next edition) – inclusion of ST 2128 descriptor needs clarification of the status, this may impose delay
  + TR on film grain synthesis technology for video applications – CD TR 23002-9 ballot response in [m63723](https://dms.mpeg.expert/doc_end_user/current_document.php?id=88067&id_meeting=195), plan to issue DTR in ISO, target ITU consent for April 2024
  + Optimization of encoders and receiving systems for machine analysis of coded video content – new WD of TR to be issued
  + Preparation of verification tests film grain and multilayer – expert viewing?
  + Any action items on reference software JM/HM/VTM? Status of MV-HEVC software and test conditions (refer to resolution of last meeting for the latter)?
  + Plan for new edition of AVC (both ISO and ITU)
    - Additional colour type identifiers for AVC and HEVC (Draft 3 in JVET-AD1008 was issued at the last meeting) – clarify for status of SMPTE ST 2128 (same as for CICP
    - Errata items from JVET-AD1004
    - Anything else? SEI messages?
    - Would be premature for ITU consent (could be targeted for April 2024), but ISO CD (or DIS) and request could be issued
  + Plan for new edition of HEVC (only ITU version; FDIS of ISO version was issued recently)
    - Additional colour type identifiers for AVC and HEVC (Draft 3 in JVET-AD1008 was issued at the last meeting) – clarify for status of SMPTE ST 2128 (may not be included)
    - Errata items from JVET-AD1004
    - Anything else? SEI messages? 10-bit Multiview profile (depends on SW and CT status)? Another H.265 edition could already follow in April 2024, aligned with AVC timeline.
  + Exploration Experiments
    - Neural network-based video coding
    - Enhanced compression beyond VVC
* Liaison communication:
  + Any incoming liaison statements? Any outgoing? JPEG?
* Joint meetings were expected with AG 5 (on Monday 17 July)
* Principles of standards development were discussed.

## Scheduling of discussions

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 in the range of 0900-1900 CEST 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 JVET calendar and the ITU posting system in advance as far as possible, 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. 11 July, 1st day
  + Morning session:
    - 0930–1025 Opening remarks, review of practices, agenda, IPR policy reminder
    - 1025–1300 Reports of AHGs 1-11
  + Afternoon session2:
    - 1430–1515 Reports of AHGs 12-14
    - 1515–1815 EE2 summary report
    - 1515–1800 BoG (G. Sullivan) on VSEI/VVC v3 SEI messages
* Wed. 12 July, 2nd day
  + Morning sessions:
    - 0830–1300 Review of 5.2.1 EE2 summary, 5.2.3 EE2 related, and 5.2.4 non-EE2
    - 0830–1300 BoG (G. Sullivan) on VSEI/VVC v3 SEI messages
  + Afternoon sessions:
    - 1430–1540 Review of 5.1.1 EE1 summary,
    - 1540-1940 Review of 5.2.4 non-EE2
    - 1430–1800 BoG (G. Sullivan) on VSEI/VVC v3 SEI messages
* Thu. 13 July, 3rd day
  + Morning session:
    - 0830–1300 BoG report, coordination/planning; review EE1 (5.1.2)
  + Afternoon session:
    - 1500–2000 Review non-EE2 (5.2.4)
    - 1500–1830 BoG on NNVC / EE1 preparation (E. Alshina, F. Galpin)
* Fri. 14 July, 4th day
  + Morning session:
    - 0830–1300 SEI messages beyond v3 (6.1)
    - 0900–1300 BoG on NNVC / EE1 preparation (E. Alshina, F. Galpin)
  + Afternoon session:
    - 1430–1530 NNVC BoG report
    - 1530–1800 SEI messages beyond v3 (6.1/6.2); optimization for machine analysis (AHG8/4.10)
    - 1615–1815 BoG on film grain (W. Husak)
    - 1800–2000 Review non-EE2 (5.2.4)
* Sat. 15 July, 5th day
  + Morning session:
    - 0830–1300 SEI messages beyond v3 (6.2); optimization for machine analysis (AHG8/4.10)
  + Afternoon session:
    - 1430–2000 Optimization for machine analysis (AHG8/4.10); film grain synthesis incl. BoG report (4.12); remaining EE2 (5.2.4/5.2.5); ECM tool assessment (4.7); encoding algorithm optimization (4.11)
* Sun t.bd.
* Mon. 17 July, 7th day
  + 0900–1130 MPEG information sharing session (out of JVET)
  + …
* Tue. t.b.d.
* Wed. 19 July, 9th day
  + 0900–1030 MPEG information sharing session (out of JVET)
  + XXXX–XXXX and XXXX–XXXX Plenary:
    - BoG reports
    - Review and approval of output docs
    - Establishment of AHGs
    - Approval of meeting recommendations
    - Future planning, a.o.b.
* Fri. 28 April
  + 1400–1600 MPEG information sharing session (out of JVET)
  + 1645–1655 WG 5 presentation of meeting recommendations

## Contribution topic overview

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

* AHG reports (14) (section 3)
* Project development (section 4)
  + AHG1: Deployment and advertisement of standards (1)
  + AHG2: Text development and errata reporting (1)
  + AHG3: Test conditions (0)
  + AHG3: Software development (1)
  + AHG4: Subjective quality testing and verification testing (3)
  + AHG4: Test Material (1)
  + AHG4: Quality assessment methodology (1)
  + AHG5: Conformance test development (1)
  + AHG7: ECM tool assessment (2)
  + AHG8: Optimization of encoders and receiving systems for machine analysis of coded video content (5)
  + AHG10: Encoding algorithm optimization (5)
  + AHG13: Film grain synthesis (2)
  + Implementation studies (1)
  + Profile/tier/level specification (0)
  + General aspects of standards development and applications of standards (0)
* 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 (17) (section 5.1)
  + AHG6/AHG12 and EE2: Enhanced compression beyond VVC capability (76) (section 5.2)
* AHG9: High-level syntax (HLS) proposals (section 6) with subtopics
  + SEI messages on neural-network post filter (28) (section 6.1)
  + SEI messages on topics other than NNPF (13) (section 6.2)
  + Non-SEI HLS aspects (0) (section 6.3)
* Joint meetings, plenary discussions, BoG reports (X), liaison (X), 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 (14)

These reports were discussed during 1025–1300 and 1430–1515 on Tuesday 11 July 2023 (chaired by JRO).

[JVET-AE0001](https://jvet-experts.org/doc_end_user/current_document.php?id=13073) JVET AHG report: Project Management (AHG1) [J.-R. Ohm (chair), G. J. Sullivan (vice-chair)]

**Goals and activity**

The work of the JVET overall had proceeded well in the interim period with lower 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 site (<http://wftp3.itu.int/av-arch/jvet-site/2023_04_AD_Antalya/>). 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/> are still accessible, but were converted to read-only.

The list of documents produced included the following, particularly:

* The meeting report (JVET-AD1000) [Posted 2023-05-26, also submitted as WG 5 N 196]
* [JVET-AD1003](https://jvet-experts.org/doc_end_user/current_document.php?id=12566) Coding-independent code points for video signal type identification (Draft 2 of 3rd edition) [Posted 2023-07-10, also issued as WG 5 preliminary FDIS N 206]
* [JVET-AD1004](https://jvet-experts.org/doc_end_user/current_document.php?id=12567) Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP [Posted 2023-07-04]
* [JVET-AD1008](https://jvet-experts.org/doc_end_user/current_document.php?id=12568) Additional colour type identifiers for AVC and HEVC (Draft 4) [Posted 2023-07-10, also issued as WG 5 preliminary WD N 200]
* [JVET-AD1012](https://jvet-experts.org/doc_end_user/current_document.php?id=12569) Overview of IT systems used in JVET [Posted 2023-06-12]
* [JVET-AD2002](https://jvet-experts.org/doc_end_user/current_document.php?id=12572) Algorithm description for Versatile Video Coding and Test Model 20 (VTM 20) [Posted 2023-06-29]
* [JVET-AD2005](https://jvet-experts.org/doc_end_user/current_document.php?id=12574) New level and systems-related supplemental enhancement information for VVC (Draft 5) [Posted 2023-06-02]
* [JVET-AD2006](https://jvet-experts.org/doc_end_user/current_document.php?id=12574) Additional SEI messages for VSEI (Draft 4) [Posted 2023-05-29]
* [JVET-AD2007](https://jvet-experts.org/doc_end_user/current_document.php?id=12575) Guidelines for NNVC software development [Posted 2023-05-29]
* [JVET-AD2016](https://jvet-experts.org/doc_end_user/current_document.php?id=12576) Common test conditions and evaluation procedures for neural network-based video coding technology [Posted 2023-05-12]
* [JVET-AD2017](https://jvet-experts.org/doc_end_user/current_document.php?id=12576) Common test conditions and evaluation procedures for enhanced compression tool testing [Posted 2023-05-10, last update 2023-06-15]
* [JVET-AD2019](https://jvet-experts.org/doc_end_user/current_document.php?id=12563) Description of algorithms and software in neural network-based video coding (NNVC) version 3 [Posted 2023-05-29]
* [JVET-AD2021](https://jvet-experts.org/doc_end_user/current_document.php?id=12578) Verification test plan for VVC multilayer coding [Posted 2023-06-12]
* [JVET-AD2022](https://jvet-experts.org/doc_end_user/current_document.php?id=12579) Draft plan for subjective quality testing of FGC SEI message [Posted 2023-06-24]
* [JVET-AD2023](https://jvet-experts.org/doc_end_user/current_document.php?id=12560) Exploration experiment on neural network-based video coding (EE1) [Posted 2023-04-28, last update 2023-05-15]
* [JVET-AD2024](https://jvet-experts.org/doc_end_user/current_document.php?id=12562) Exploration experiment on enhanced compression beyond VVC capability (EE2) [Posted 2023-04-28, last update 2023-05-29]
* [JVET-AD2025](https://jvet-experts.org/doc_end_user/current_document.php?id=12580) Algorithm description of Enhanced Compression Model 9 (ECM 9) [Posted 2023-07-03]
* [JVET-AD2027](https://jvet-experts.org/doc_end_user/current_document.php?id=12582) SEI processing order SEI message in VVC (Draft 4) [Posted 2023-05-19, also issued as WG 5 preliminary WD N 203]
* [JVET-AD2028](https://jvet-experts.org/doc_end_user/current_document.php?id=12583) Additional conformance bitstreams for VVC multilayer configurations [Posted 2023-05-26]
* [JVET-AD2030](https://jvet-experts.org/doc_end_user/current_document.php?id=12564) Optimization of encoders and receiving systems for machine analysis of coded video content (draft 2) [Posted 2023-06-07, also issued as WG 5 preliminary WD N 199]
* [JVET-AD2031](https://jvet-experts.org/doc_end_user/current_document.php?id=12584) Common test conditions for optimization of encoders and receiving systems for machine analysis of coded video content [Posted 2023-05-11]

Furthermore, the following documents were submitted to the ISO/IEC JTC1/SC29 parent body on behalf of its WG 5:

* Recommendations of the 11th WG 5 meeting (WG 5 N 195)
* Draft disposition of comments received on ISO/IEC 23002-7:2022 DAM 1 (WG 5 N 197)
* Preliminary text of ISO/IEC FDIS 23002-7:202x Versatile supplemental enhancement information messages for coded video bitstreams (3rd edition) (WG 5 N 198)
* Draft disposition of comments received on ISO/IEC 23090-3:2022 DAM 1 (WG 5 N 201)
* Preliminary text of ISO/IEC FDIS 23090-3:202x Versatile video coding (3rd edition) (WG 5 N 202)
* Draft disposition of comments received on ISO/IEC DIS 23091-2:202X (WG 5 N 205)
* Liaison statement to SC 29/WG 1 (JPEG) on JPEG AI and neural network-based video coding (WG 5 N 211)
* List of AHGs established at the 11th WG 5 meeting (WG 5 N 212)

It is noted that documents JVET-AC 2026 Conformance testing for VVC operation range extensions (Draft 4), and the corresponding FDIS 23090-15 as WG 5 N 185 (VVC conformance 2nd ed.) have still not been delivered yet as outputs from the 29th JVET (10th WG 5) meeting of January 2023.

The fourteen *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 31st meeting had been announced in the JVET reflector, in the JVET logistics document (<https://www.itu.int/wftp3/av-arch/jvet-site/2023_07_AE_Geneva/JVET-AE_Logistics.docx>), and in the WG 5 calling notice (N 213) and agenda (N 214) for the 12th 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 160 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 31st meeting had been made publicly available on the ITU-hosted ftp site as <http://wftp3.itu.int/av-arch/jvet-site/2023_07_AE_Geneva/JVET-AE_notes_d0.docx>.

**Recommendations**

The AHG recommends its continuation.

[JVET-AE0002](https://jvet-experts.org/doc_end_user/current_document.php?id=13074) 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)]

**Output documents produced**

**JVET-AD1003 - Coding-independent code points for video signal type identification (Draft 2 of 3rd edition)**

This document contains the draft text for the specification of additional colour type identifiers for CICP (Rec. ITU-T H.273 | ISO/IEC 23091-2). Text modifications are provided for specification of code point identifiers for YCgCo-R colour representation with equal luma and chroma bit depths and for Draft SMPTE ST 2128. The new code points for YCgCo-R are referred to as YCgCo-Re and YCgCo-Ro, where the number of bits added to a source RGB bit depth is 2 (i.e., even) and 1 (odd), respectively. Draft SMPTE ST 2128 specifies a colour representation referred to as IPT-PQ-C2. The affected sections are provided relative to the corresponding basis text of ISO/IEC DIS 23091-2:202x (3rd Ed.). Change marks are included to show the changes since the January output (JVET-AC1008), which reflect a renumbering of the new MatrixCoefficients values and some minor refinements and corrections since the previous draft.

**JVET-AD1004 - 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, Video CICP, and the TR on usage of video signal type code points, 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-AD meeting:

* For AVC, incorporated the following items:
  + JVET-AD0078 item 4 and 5 on fixing the equation to derive the filter hint cross correlation matrix.

**JVET-AD1008 - Additional colour type identifiers for AVC and HEVC (Draft 4)**

This document contains the draft text for the specification of additional colour type identifiers for AVC (Rec. ITU-T H.264 | ISO/IEC 14496-10) and HEVC (Rec. ITU-T H.265 | ISO/IEC 23008-2). Text modifications are provided for specification of code point identifiers for YCgCo-R colour representation with equal luma and chroma bit depths and for Draft SMPTE ST 2128. The new code points for YCgCo-R are referred to as YCgCo-Re and YCgCo-Ro, where the number of bits added to a source RGB bit depth is 2 (i.e., even) and 1 (odd), respectively. Draft SMPTE ST 2128 specifies a colour representation referred to as IPT-PQ-C2. The affected sections are provided relative to corresponding basis texts (based on the 2021-08 edition of Rec. ITU-T H.264 (the most recent edition), the 2020 edition of ISO/IEC 23008-2 (not the 5th edition). Equation numbers and their cross-references have been updated without revision marking. Change marks are included to show the changes since the January output (JVET-AC1008), which reflect a renumbering of the new matrix coefficients values and some minor refinements and corrections since the previous draft.

**JVET-AD2002 - Algorithm description for Versatile Video Coding and Test Model 20 (VTM 20)**

The JVET established the VVC Test Model 21 (VTM21) software at its 30th meeting (21-28 April 2023, Antalya, TR). This document serves as a source of general tutorial information on the VVC design and also provides an algorithm description and encoding method description of VTM21 software. It is noted that, as no update of the algorithm description document was released at the 27th JVET meeting, the numbering of the VTM software version is one higher than that of this description document. In the main body of the text, numbering refers to the software version. The VVC has been developed by a joint collaborative team of ITU-T and ISO/IEC experts known as the Joint Video Experts Team (JVET), which is a partnership of ITU-T Study Group 16 Question 6 (known as VCEG) and ISO/IEC JTC 1/SC 29/WG 11 (known as MPEG). This new standard has been designed with two primary goals. The first of these is to specify a video coding technology with a compression capability that is substantially beyond that of the prior generations of such standards, and the second is for this technology to be highly versatile for effective use in a broadened range of applications. In addition to the applications that have commonly been addressed by prior video coding standards, some key application areas for the use of this standard include in particular ultra-high-definition video (e.g., with 3840×2160 or 7620×4320 picture resolution and bit depth of 10 or 12 bits as specified in Rec. ITU-R BT.2100), video with a high dynamic range and wide colour gamut (e.g., with the perceptual quantization or hybrid log-gamma transfer characteristics specified in Rec. ITU-R BT.2100), and video for immersive media applications such as 360° omnidirectional video projected using a common projection format such as the equirectangular or cubemap projection format.

Ed. Notes:

VVC Test Model 20 (VTM20) algorithm description and encoding method v1

* General improvements to reference picture resampling sections
* Incorporated JVET-AD0169: AHG12: RPR filters for scale factors below 1.5x

Incorporated JVET-AD0045: AHG10: Encoder MV selections and DMVR revisited.

**JVET-AD2005 - New level and systems-related supplemental enhancement information for VVC (Draft 5)**

This document contains a draft amendment for changes to the Versatile Video Coding (VVC) standard (ITU‑T H.266 | ISO/IEC 23090-3). This amendment includes the following: 1) the support of an unlimited level for video profiles; 2) some technical corrections and editorial improvements to the 2nd edition text of VVC; and 3) the VVC-specific supports for some supplemental enhancement information (SEI) messages that may be included in VVC bitstreams but are not to be specified in the VVC specification. These SEI messages include two systems-related SEI messages, one for signalling of "green metadata" as to be specified in ISO/IEC 23001-11 and the other for signalling of an alternative video decoding interface for immersive media as to be specified in ISO/IEC 23090-13, and five other SEI messages, namely the shutter interval information SEI message, the neural network post-filter characteristics SEI message SEI message, the neural-network post-processing filter activation SEI message, the phase indication SEI message, and the post filter hint SEI message, that are to be specified in a new edition of the Versatile Supplemental Enhancement Information messages for coded video bitstreams (VSEI) standard (ITU‑T H.274 | ISO/IEC 23002-7).

Draft 5 items from the BoG report JVET-AD0362 yet to be integrated:

None.Other draft 5 items yet to be integrated:

* Spec bug ticket [#1594](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1594) at editors’ discretion

Draft 5 incorporated items:

* Bug ticket [#1593](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1593): Replaced Figures 5 to 7 for correction to illustrated slice partitioning.
* JVET-AD0362, action on item 2)b.ii: Allowed input pictures to an NNPF activated for any current picture to come from an earlier CLVS than the CLVS containing the current picture, and for the derivation of which pictures are used as in the input pictures, instead of using picture order count values, plain language phrases like "the previous picture in output order" and "the n-th previous picture in output order" are used.
* JVET-AD0362, action on item 2)b,i.2: Added a constraint that NNPF shall not be activated so that it creates an interpolated picture before the first picture in the bitstream in output order.
* JVET-AD0362, action on item 2)d: At the end of a bitstream, added repeated inference of NNPF for creating interpolated pictures: When an NNPF is active until the end of the bitstream, the NNPF is applied up to but excluding a set of input pictures that would cause creation of any interpolated picture after the last picture of the bitstream in output order.
* JVET-AD0362, action on item 2)e.i: Limited the special treatment of temporal interleaving frame packing arrangement to picture rate upsampling only rather than for any multi-picture-input NNPF.
* JVET-AD0362, action on item 2)f with the change of "the CLVS" to "the bitstream": Fixed an asserted bug in the derivation of input pictures for the case when there are multiple input pictures for an NNPF activated for currCodedPic and currCodedPic is associated with a frame packing arrangement SEI message with fp\_arrangement\_type equal to 5, taking into account that there can be multiple pictures preceding an input picture in output order.
* JVET-AD0362, action on item 2)g: Specified that any multi-picture-input NNPF can only be used when all input pictures have the same dimensions after applying super resolution post-filter applicable to the input pictures, if any, using text provided in JVET-AD0057 item 6.
* JVET-AD0362, action on item 6)g: In the VVC text, disallowed the NNPFA SEI message from being associated with a non-output picture and disallowed any input pictures being a non-output picture.
* JVET-AD0362, action on item 14)a: Specified an array of filtering strength control variables for multiple input pictures (as StrengthControlVal[ idx ] values for the input pictures with index idx in the range of 0 to numInputPics − 1, inclusive).
* JVET-AD0362, action on item 14)d: Changed the derivation of StrengthControlVal to be equal to ( SliceQpY + QpBdOffset ) ÷ ( 63 + QpBdOffset), instead of SliceQpY ÷ 63, where SliceQpY is that of the first slice of currCodedPic.
* JVET-AD0362, action on item 13)a: Allow inclusion of NNPFC and NNPFA SEI messages in suffix SEI NAL units.
* JVET-AD0362, action on item 13)b: Disallow the post-filter hint SEI message to be present in both a prefix SEI NAL unit and a suffix SEI NAL unit in the same picture unit.
* Editorial change from JVET-AD0077
* VclAssociatedSeiList consists of the payloadType values of the SEI messages that, when non-scalable-nested, infer constraints on the NAL unit header of the SEI NAL unit on the basis of the NAL unit header of the associated VCL NAL unit. It is proposed to add the payloadType value of the green metadata SEI message in VclAssociatedSeiList. (JVET-AD0057, item 3)
* PicUnitRepConSeiList consists of the payloadType values of the SEI messages that are subject to the restriction on 4 repetitions per picture unit (PU). It is proposed to add the payloadType values of post-filter hint, green metadata, scalability dimension information, NNPFC, NNPFA, and phase indication SEI messages in PicUnitRepConSeiList. (JVET-AD0057, item 4)

Alignments of SEI messages between VVC and HEVC, regarding inclusion in prefix and/or suffix SEI NAL units (including for the user data SEI messages), and inclusion into the lists VclAssociatedSeiList and PicUnitRepConSeiList. (In-meeting suggestion).

**JVET-AD2006 - Additional SEI messages for VSEI (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 the shutter interval information SEI message, neural-network post-filter characteristics SEI message, neural-network post-filter activation SEI message, phase indication SEI message, and post-filter hint SEI message.

**Changes that have been integrated:**

Items noted with adoptions in the JVET meeting notes but not in the BoG report JVET-AD0362-v4:

* (JVET-AD0078 items 4 and 5) Fix two bugs in the equation in the semantics of the post-filter hint SEI message.

Items noted with adoptions in the BoG report JVET-AD0362-v4:

* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Add an indication nnpfc\_absent\_input\_pic\_zero\_flag in the NNPFC SEI message that indicates how pictures that would not originate from the bitstream are expected to be replaced in the input tensor. nnpfc\_absent\_input\_pic\_zero\_flag equal to 1 indicates that the NNPF expects an input picture that is not present in the bitstream to be represented sample arrays with sample values equal to 0. nnpfc\_absent\_input\_pic\_flag equal to 1 indicates that the NNPF expects an input picture that is not present in the bitstream to be represented by the closest input picture in output order within the bitstream.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Add nnpfa\_output\_flag[ i ] syntax elements in the NNPFA SEI message for controlling which NNPF-generated pictures that correspond to input pictures are output by the NNPF process, and impose both of the following constraints:
  + For a particular NNPF, disallow it to output a picture multiple times for any particular output order or output time instance.
  + For all NNPFs activated for at least one picture in a bitstream, disallow the NNPFs (when alternative, only one is used) to output a picture multiple times for any particular output order or output time instance.
* Action on items **Fehler! Verweisquelle konnte nicht gefunden werden.**, **Fehler! Verweisquelle konnte nicht gefunden werden.**, and **Fehler! Verweisquelle konnte nicht gefunden werden.**: Specify a general post-processing filtering process using NNPFs, wherein the input to the process is a bitstream BitstreamToFilter, and the output of the process is a list of NNPF output pictures. Then clearly specify the above constraints (as part of the recommended action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**) in the context of the general post-processing filtering process.
* Action on items **Fehler! Verweisquelle konnte nicht gefunden werden.**, **Fehler! Verweisquelle konnte nicht gefunden werden.**, and **Fehler! Verweisquelle konnte nicht gefunden werden.**: Include nnpfc\_input\_pic\_output\_flag[ i ] in the NNPFC SEI message when nnpfc\_num\_input\_pics\_minus1 is greater than 0 regardless of the nnpfc\_purpose value, move the flag to be immediately after nnpfc\_num\_input\_pics\_minus1, and skip signalling of the flag when nnpfc\_num\_input\_pics\_minus1 is equal to 0. Require at least one of the flags to be equal to 1 when there are multiple of them and picture rate upsampling is not indicated.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Delegate to the editor to decide to either specify an inference rule for nnpfc\_input\_pic\_output\_flag[ i ] or directly consider the possible absence in the equations for derivation of the number of output pictures.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Add the following constraints: It is a requirement of bitstream conformance that the value of nnpfc\_pic\_width\_in\_luma\_samples % outSubWidthC shall be equal to 0. It is a requirement of bitstream conformance that the value of nnpfc\_pic\_height\_in\_luma\_samples % outSubHeightC shall be equal to 0.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: For spatial resolution upsampling, signal the scaling ratios instead of the NNPF output picture width and height.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Specify that the width of the NNPF output picture shall be in the range of CroppedWidth to CroppedWidth \* 16, inclusive (instead of CroppedWidth to CroppedWidth \* 16 − 1, inclusive), and the width of the NNPF output picture shall be in the range of CroppedHeight to CroppedHeight \* 16, inclusive (instead of CroppedHeight to CroppedHeight \* 16 − 1, inclusive).
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Add a constraint on NNPF input tensor and output tensor bit depth values to avoid changing in opposite directions from input to output (e.g., luma bit depth getting bigger while chroma bit depth getting smaller).
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Add a metadata extension mechanism in the NNPFC SEI message, to allow new metadata syntax elements to be added in the NNPFC SEI message in future VSEI amendments or editions, and require parsers of v3 of VSEI to skip and ignore such additional metadata syntax elements.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Move nnpfc\_base\_flag next to nnpfc\_id in the NNPFC SEI message syntax.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Make changes to avoid a repetition of an NNPFC SEI message affecting the persistence of an NNPF update.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Fix the incorrect inference of upsampled resolution, colour primaries, transfer characteristics, and matrix coefficients syntax elements when nnpfc\_property\_present\_flag is equal to 0
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Enable the activation of the base NNPF with a particular nnpfc\_id value after an NNPF update for the same nnpfc\_id is available.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Change the persistence scope of the NNPFC SEI message to "The CLVS containing the SEI message".
* Action on items **Fehler! Verweisquelle konnte nicht gefunden werden.** and **Fehler! Verweisquelle konnte nicht gefunden werden.**: Move nnpfc\_inp\_order\_idc and nnpfc\_out\_order\_idc before the bit depth syntax elements, and gate the bit depth syntax elements so that the luma bit depth is present only when the input or output tensor contains luma, and the chroma bit depth is present only when the input or output tensor contains chroma. And update the related constraints on the values of nnpfc\_out\_tensor\_luma\_bitdepth\_minus8 and nnpfc\_out\_tensor\_chroma\_bitdepth\_minus8 accordingly.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Add syntax conditions such that nnpfc\_luma\_padding\_val is present only when the input tensor contains a luma sample array, and nnpfc\_cb\_padding\_val and nnpfc\_cr\_padding\_val are present only when the input tensor contains chroma sample arrays.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Signal nnpfc\_auxiliary\_inp\_idc before nnpfc\_inp\_order\_idc, as the semantics of nnpfc\_inp\_order\_idc syntax element use nnpfc\_auxiliary\_inp\_idc.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Add the following constraint: When nnpfc\_purpose & 0x02 is not equal to 0, nnpfc\_inp\_order\_idc shall not be equal to 0.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Remove "(from the 4:0:0 chroma format to the 4:2:0, 4:2:2, or 4:4:4 chroma format)" in the table describing the NNPF purposes, such that colourization can be applied even when the cropped decoded pictures are of the 4:0:0 chroma format.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Add a constraint that disallows nnpfc\_inp\_order\_idc indicating that the input tensor has chroma component when ChromaFormatIdc is equal to 0.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Make the following change (additions indicated with underline): When nnpfc\_purpose & 0x02 is not equal to 0, nnpfc\_out\_order\_idc shall not be equal to 0 or 3.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Add a constraint such that when the purpose indicates colorization, there shall be chroma output by the NNPF.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Move nnpfc\_separate\_colour\_description\_present\_flag, nnpfc\_colour\_primaries, nnpfc\_transfer\_characteristics, and nnpfc\_matrix\_coeffs to be after nnpfc\_out\_order\_idc. Include the nnpfc\_matrix\_coeffs syntax element in the syntax under the constraint that nnpfc\_out\_format\_idc is equal to 1 (indicating integer sample values in the output tensor). For the semantics of nnpfc\_matrix\_coeffs, replicate the semantics of vui\_matrix\_coeffs with reference to the chroma format, chroma bit depth, and luma bit depth of the output tensor.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Include a syntax element nnpfc\_full\_range\_flag in the NNPFC message when nnpfc\_separate\_colour\_description\_present\_flag is equal to 1 and when nnpfc\_out\_format\_idc is equal to 1.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Add a syntax element indicting the location of chroma samples, and the presence of the syntax element is controlled by a new presence flag. Add a constraint to require the value of the new presence flag to be equal to 0 when nnpfc\_output\_color\_format\_idc is not equal to 1.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Change the definition of the function InpSampleVal( ) to take the component index (0 for luma, 1 for Cb, 2 for Cr) as an additional input argument. For nnpfc\_padding\_type equal to 4, change the function InpSampleVal to return the fixed luma, Cb, or Cr padding value corresponding to the component index given as input argument. For each invocation of the function InpSampleVal, add the component index among the input arguments.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Remove the following constraints on nnpfa\_target\_id values, which were asserted to be enforced by other constraints and are therefore asserted to be redundant: Values of nnpfa\_target\_id from 256 to 511, inclusive, and from 231 to 232 − 2, inclusive, are reserved for future use by ITU-T | ISO/IEC. Decoders conforming to this edition of this document encountering an NNPFA SEI message with nnpfa\_target\_id in the range of 256 to 511, inclusive, or in the range of 231 to 232 − 2, inclusive, shall ignore the SEI message.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Specify an array of filtering strength control variables for multiple input pictures (as StrengthControlVal[ idx ] values for the input pictures with index idx in the range of 0 to numInputPics − 1, inclusive), and derive an array scaled strengthControlScaledVal[ i ] values for multiple input pictures and use them to derive the input tensors.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Make changes to Equation (83) such that when nnpfc\_inp\_order\_idc is equal to 1 (i.e., only chroma is present in the input tensor), the variable strengthControlScaledVal is derived from the variable inpTensorBitDepthC (instead of from inpTensorBitDepthY).
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Extend the current resolution upsampling to be resolution resampling such that both resolution upsampling and resolution downsampling capabilities are supported.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Clarify that the output from the PostProcessingFilter (i.e. outputTensor) does not include overlap regions.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Update the specification of the value range of nnpfc\_padding\_type to take into account that the values in the range 5 to 15, inclusive, are reserved.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.**: Add missing value ranges of ue(v)-coded syntax elements nnpfc\_luma\_padding\_val, nnpfc\_cb\_padding\_val, and nnpfc\_cr\_padding\_val, with the lower limit being 0 and the upper limit being two power of the corresponding bit depth values of the cropped output pictures minus 1.
* Action on items **Fehler! Verweisquelle konnte nicht gefunden werden.** and **Fehler! Verweisquelle konnte nicht gefunden werden.** (editorial): Fix typos by changing all instances of "nnpfc\_parameter\_parameter\_type\_idc" to "nnpfc\_parameter\_type\_idc" and changing "4:2:4 format" to "4:4:4 format".
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.** (editorial): Restructure the semantics of nnpfc\_purpose including table 20 with the definition of the nnpfc\_purpose.
* Action on item **Fehler! Verweisquelle konnte nicht gefunden werden.** (editorial): Avoid repeatedly using "&" operations in many places, e.g., by deriving and using variables, e.g., spatialResolutionResamplingIncludedFlag.

**JVET-AD2027 - SEI processing order SEI message in VVC (draft 4)**

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

**Changes yet to be integrated:**

Items noted with adoptions in the BoG report JVET-AD0362-v4: none.

**Changes that have been integrated:**

Items noted with adoptions in the BoG report JVET-AD0362-v4: 1d (using text in JVET-AD0386)

**Related input contributions**

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

* [JVET-AE0054](file:///C:\Users\rosewarc\Desktop\current_document.php%3fid=13002) AHG2/AHG9: Editorial improvements of annotated regions SEI message
* [JVET-AE0101](file:///C:\Users\rosewarc\Desktop\current_document.php%3fid=13049) [AHG2][AHG9] Neural network post filter and phase indication SEI messages for AVC and HEVC
* [JVET-AE0155](file:///C:\Users\rosewarc\Desktop\current_document.php%3fid=13118) AHG2/AHG9: Editor commentary on the post-filter hint SEI message semantics

**Remaining bug tickets**

* [#1606](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1606) [VVC v2 wrongly uses zero out block size](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1606)
* [#1602](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1602) [Need clarification for 8.4.3 Derivation process for chroma intra prediction mode](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1602)

Fixes are available, will be included in next edition.

**Recommendations**

The AHG recommends to:

* Approve JVET-AD1003, JVET-AD1004, JVET-AD1008, JVET-AD2002, JVET-AD2005, JVET-AD2006, and JVET-AD2027 documents as JVET outputs,
* Compare the VVC documents with the VVC software and resolve any discrepancies that may exist, in collaboration with the software AHG,
* Encourage the use of the issue tracker to report issues with the text of both the VVC specification text and the algorithm and encoder description,
* Continue to improve the editorial consistency of VVC text specification and Test Model documents,
* Ensure that, when considering changes to VVC, properly drafted text for addition to the VVC Test Model and/or the VVC specification text is made available in a timely manner,
* Review bug tickets, and other AHG2 related inputs and act on them if found to be necessary.

[JVET-AE0003](https://jvet-experts.org/doc_end_user/current_document.php?id=13075) 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 21.0](https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/-/releases/VTM-21.0) (Jul. 2023)
* [HM-18.0](https://vcgit.hhi.fraunhofer.de/jvet/HM/-/releases/HM-18.0) (Apr. 2023)
* [HM-16.21+SCM-8.8](https://vcgit.hhi.fraunhofer.de/jvet/HM/-/tags/HM-16.21+SCM-8.8) (Mar. 2020)
* [SHM 12.4](https://vcgit.hhi.fraunhofer.de/jvet/SHM/-/tags/SHM-12.4) (Jan. 2018)
* [HTM 16.3](https://vcgit.hhi.fraunhofer.de/jvet/HTM/-/tags/HTM-16.3) (Jul. 2018)
* [JM 19.1](https://vcgit.hhi.fraunhofer.de/jvet/JM/-/releases/JM-19.1) (Apr. 2023)
* [JSVM 9.19.15](https://vcgit.hhi.fraunhofer.de/jvet/jsvm/-/tags/JSVM_9_19_15)
* [JMVC 8.5](https://vcgit.hhi.fraunhofer.de/jvet/jmvc/-/tags/JMVC_8_5)
* [3DV ATM 15.0](https://vcgit.hhi.fraunhofer.de/jvet/3dv-atm/-/tags/3DV-ATM_v15.0) (no version history)
* [HDRTools 0.24](https://gitlab.com/standards/HDRTools/-/tags/v0.24) (March 2023)

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

**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.

**VTM related activities**

The VTM software can be found at

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

The software development continued on the GitLab server. VTM versions 20.1 and 20.2 were tagged on May 8, and VTM version 21.0 was tagged on Jul. 9. VTM 21.1 is expected during the 31st JVET meeting.

VTM 20.1 was tagged on May 8, 2023. Changes include:

* Undefine macro from HDRTools
* Fix build when GDR\_ENABLED is 0
* Some fixes after [!2541](file:///\\jvet\VVCSoftware_VTM\-\merge_requests\2541) (Use strong-typed enum for chroma format)
* Fix parsing of SDI SEI message
* Fix parsing of MAI SEI message
* Fix parsing of multiple SEI payloads in one NAL unit
* Add command line option to print per-frame encoding time in float format
* Fix [#1595](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1595): Set ptl\_multi\_layer\_enabled\_flag when more than one layer is encoded
* Rollback SEI bits computation as it is erroneous
* Fix compilation when EXTENSION\_360\_VIDEO is enabled
* Shorten Neural Network Post Filter Characteristics SEI parameters
* Fix type of literal in assert expression
* Fixes for multilayer coding
* Port changes from JVET-Z0072 to performance low delay configuration file
* Revert [!2573](file:///\\jvet\VVCSoftware_VTM\-\merge_requests\2573) and add MultiLayerEnabledFlag=1 to two\_layers.cfg

VTM 20.2 was tagged on May 8, 2023. Changes include:

* Remove macros from previous cycle

VTM 21.0 was tagged Jul. 9, 2023. Changes include:

* Fix #1599: include frame rate info in bitstream when provided in y4m file
* JVET-AD0169: Downsampling filters in range 1.1 to 1.35 based on Kaiser(7) windowed sinc
* JVET-AD0386: Processing Order SEI message update
* Fix #1598: signal chroma location type when input is y4m
* Fix #1597 nnpfc\_constant\_patch\_size\_flag and nnpfc\_overlap and nnpfc\_overlap parsing.
* JVET-AD0045: Encoder MV selections and DMVR revisited
* Avoid seg fault Slice::constructRefPicList
* Code cleanup: fix variable names, use proper data types, etc.
* Replace EncAppCfg::OptionalValue with std::optional
* Remove JVET\_O0549\_ENCODER\_ONLY\_FILTER\_POL macro
* Fix output frame rate when using --MaxTemporalLayer in decoder
* Rename df::program\_options\_lite as ProgramOptionsLite
* Rename and initialize m\_RCCpbSaturationEnabled
* JVET-AD0056 item 5: don't signal nnpfc\_input\_pic\_output\_flag[i] flags when nnpfc\_num\_input\_pics\_minus1==0
* Fix variable names in VideoIOYuv
* Fix parsing of frame rate parameter
* JVET-AD0067 item2: Swap signaling order for nnpfc\_aux\_inp\_idc and nnpfc\_inp\_order\_idc
* JVET-AD0129: Fix padding
* JVET-AD0054: nnpfc\_absent\_input\_pic\_zero\_flag
* JVET-AD0141: NNPFA is associated with picture that is an output picture
* JVET-AD0383: Signalling scaling ratio for output picture size instead of absolute width and height
* JVET-AD0067 Item 7: Include syntax element nnpfc\_full\_range\_flag
* Fix #1596 on encoding crash when DeltaQpRD is used
* Avoid CU copy in EncModeCtrlMTnoRQT::useModeResult
* JVET-AD0056 Item 4: picture width and height constraints
* Fix default values of SEI NNFPC parameters to avoid crash and fix variable names
* Check config parameters after they have been set
* JVET-AD0056 Item 6: move nnpfc\_inp\_order\_idc and nnpfc\_out\_order\_idc
* JVET-AD0056 Item 1: move nnpfc\_base\_flag next to nnpfc\_id
* JVET-AD0056 Item 11: nnpfc SEI metadata extension
* JVET-AD0056 Item 9: add condition for nnpfc padding signaling
* Fix compile issues with GCC 13 on Windows
* Remove singleComp function and clean up singleChan function
* JVET-AD0056 Item 7: nnpfc\_inp\_order\_idc and nnpfc\_out\_order\_idc constraints
* JVET-AD0056 Item 8: Move nnpfc\_separate\_colour\_description\_present\_flag
* Add CHECK statement to avoid possible buffer overflow
* fix MSAN issues
* JVET-AD0091 items 4-6: add constraints for nnfpc padding
* Fix RPL construction
* JVET-AD0056 Item 12: remove redundant constraints on nnpfa\_target\_id
* JVET-AD0233 Items 1,2,6: Add NNFPC constraints and chroma sample location
* JVET-AD0056 Item 13: implementation of nnpfa\_target\_base\_flag
* Fix #1603 Overflow in importMergeInfo()
* JVET-AD0057 item 6: multi-picture-input constraint
* Fix #1604: NNPFC constraints check
* JVET-AD0057 items 3 and 4: Update VclAssociatedSeiList and PicUnitRepConSeiList
* JVET-AD0057 item 2: disallow the post-filter hint SEI message to be present in both a prefix SEI NALU and a suffix SEI NALU in the same picture unit
* JVET-AD0388: add nnfpa\_num\_output\_entries and nnfpa\_output\_flag syntax elements

VTM 21.1 is expected to be tagged during the 31st JVET meeting. Changes are expected to include the remaining implementations of the 30th meeting.

* 1. ***CTC Performance***

The following tables shows **VTM 20.0** performance over **HM 17.0** (not updated).

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

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

There is no change in coding performance or run time between **VTM 21.0** and **VTM 20.1** using SDR CTC.

HDR and high bit depth CTC were stated. Partial results show no performance change as well.

The following tables show **VTM 20.0** performance over **HM 17.0** using HDR CTC (not updated):

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access** | | | | | | | | | |
|  | **Over HM17.0** | | | | | | | | | |
|  |  |  | **wPSNR** |  |  | **PSNR** |  |  |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | -39.54% | -37.56% | -36.67% | -54.98% | -48.45% | -33.68% | -49.48% | -40.84% | 286% | 97% |
| Class H2 |  |  |  |  |  | -31.97% | -57.69% | -63.36% | 271% | 97% |
| **Overall** | -39.54% | -37.56% | -36.67% | -54.98% | -48.45% | -33.06% | -52.47% | -49.03% | 280% | 97% |
|  |  |  |  |  |  |  |  |  |  |  |
|  | **All Intra** | | | | | | | | | |
|  | **Over HM17.0** | | | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | -41.45% | -27.31% | -26.79% | -57.74% | -52.79% | -24.00% | -52.93% | -45.13% | 1538% | 113% |
| Class H2 |  |  |  |  |  | -21.76% | -47.21% | -50.57% | 1320% | 109% |
| **Overall** | -41.45% | -27.31% | -26.79% | -57.74% | -52.79% | -23.18% | -50.85% | -47.11% | 1455% | 112% |

***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.

**HM related activities**

HM 18.0 was tagged on Apr. 20, 2023. Changes include:

* Update to Ubuntu 22.04 build server
* Update the HDR configuration files to align with the HDR CTC document (JVET-AC2011)
* Update to FilmGrainCharacteristics SEI support, including applying film grain noise onto the decoded output.

The performance of HM 18.0 is identical to HM 17.0 under SDR and HDR CTCs. Initial results also indicate that the performance is identical to HM 17.0 under RExt non-4:2:0 and RExt high bit depth CTCs.

The following MRs are pending [with status indicated]:

* 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.

Two newer tickets were filed for HM in relation to 360Lib and the NoOutPrior\_A\_Qualcomm\_1 conformance stream.

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

* 43 tickets for “HM”, most of which are more than 5 years,
* 1 ticket for “HM RExt”,
* 9 tickets for “HM SCC”, most of which are at least 3 years old,

Help to address these tickets would be appreciated.

One merge request (MR#40) is available related to HM SCC for ticket [#1511](https://hevc.hhi.fraunhofer.de/trac/hevc/ticket/1511) on SCC reference picture marking. We would appreciate help to confirm that the proposed change matches the SCC text.

**360Lib related activities**

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

There had not been any further developments to 360Lib during this meeting cycle.

**SCM related activities**

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

**SHM related activities**

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

**HTM related activities**

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

The next release will include the following changes:

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

**HDRTools related activities**

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

**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. An possible encoder/decoder mismatch was reported privately, but needs more investigation.

**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.

**Software repositories**

Git repositories that were previously assigned to the JCT-VC group on the GitLab server were re-assigned to the JVET group. The old URLs are still working and will forward the user to the new location, with the display of a warning suggesting to update bookmarks to the new location. The SVN repository for 360Lib was converted to git and development was moved to the GitLab server. Historical branches can still be accessed in the SVN repository.

**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

Older CTC documents are:

JCT3V-G1100 3DV test conditions

JVET-AC1013 was a planned output document for updating JCT3V-G1100, which was not produced. JVET-AD0201 was submitted with a description of the problems that were found during editing.

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.

**Guidelines for reference software development**

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

**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.
* Consider documents (including late documents) related to AHG3 activities.

It was noted that JVET-AE0179 renews licensing conditions for 3DV content, JVET-AE1013 could be issued.

[JVET-AE0004](https://jvet-experts.org/doc_end_user/current_document.php?id=13076) JVET AHG report: Test material and visual assessment (AHG4) [V. Baroncini, T. Suzuki, M. Wien (co-chairs), S. Iwamura, S. Liu, S. Puri, A. Segall, P. Topiwala, S. Wenger, J. Xu, Y. Ye (vice-chairs)]

**Activities**

**Verification *test preparation for VVC multilayer coding***

A teleconference was organized by AG 5 and JVET AhG4 on 2023-05-26 07:00h UTC. The topic was the Verification Test preparation for VVC Multilayer Coding. The report of the AhG meeting is available in JVET-AE0041.

***Test sequences***

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

In the report period, test material related to the CTC for SHVC and HEVC SCC has been made available in that directory.

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).

**Related contributions**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| [JVET number](https://jvet-experts.org/doc_end_user/current_meeting.php?id_meeting=195&type_order=&sql_type=document_number) | MPEG number | [Created](https://jvet-experts.org/doc_end_user/current_meeting.php?id_meeting=195&type_order=&sql_type=document_date_time) | First upload | [Last upload](https://jvet-experts.org/doc_end_user/current_meeting.php?id_meeting=195&type_order=&sql_type=upload_document_date_time) | [Title](https://jvet-experts.org/doc_end_user/current_meeting.php?id_meeting=195&type_order=&sql_type=title) | [Source](https://jvet-experts.org/doc_end_user/current_meeting.php?id_meeting=195&type_order=&sql_type=authors) |
| [JVET-AE0004](https://jvet-experts.org/doc_end_user/current_document.php?id=13076) | m63843 | 2023-07-04 17:26:52 |  |  | JVET AHG report: Test material and visual assessment (AHG4) | V. Baroncini, T. Suzuki, M. Wien (co-chairs), S. Iwamura, S. Liu, S. Puri, A. Segall, P. Topiwala, S. Wenger, J. Xu, Y. Ye (vice-chairs) |
| [JVET-AE0041](https://jvet-experts.org/doc_end_user/current_document.php?id=12988) | m63654 | 2023-05-28 22:37:10 | 2023-05-28 22:43:21 | 2023-05-28 22:56:04 | AHG4: Report on AhG meeting on verification testing for VVC multilayer coding | M. Wien (AhG Co-Chair) |
| [JVET-AE0092](https://jvet-experts.org/doc_end_user/current_document.php?id=13040) | m63797 | 2023-07-04 09:57:25 | 2023-07-04 13:34:47 | 2023-07-04 13:34:47 | [AHG4] Occupancy-only PSNR calculations for V3C V-PCC coding evaluation | S. Schwarz, M. M. Hannuksela |
| [JVET-AE0179](https://jvet-experts.org/doc_end_user/current_document.php?id=13142) | m63912 | 2023-07-05 06:14:29 | 2023-07-05 15:13:30 | 2023-07-05 15:13:30 | AHG4: Renewed license statement for Ghost Town Fly and Undo Dancer 3D video test sequences | M. M. Hannuksela, N. Salonen (Nokia) |
| [JVET-AE0219](https://jvet-experts.org/doc_end_user/current_document.php?id=13182) | m64064 | 2023-07-09 17:23:42 |  |  | AHG4: Results of visual checking of SVVC VT streams with and without DMVR fix | M. Wien (RWTH Aachen University) |
| [JVET-AE0227](https://jvet-experts.org/doc_end_user/current_document.php?id=13190) | m64155 | 2023-07-10 11:28:54 |  |  | AHG4: experiments in preparation of scalable quality ladder visual testing | P. de Lagrange, F. Urban (InterDigital) |
| [JVET-AE0250](https://jvet-experts.org/doc_end_user/current_document.php?id=13213) | m64415 | 2023-07-11 10:00:47 |  |  | Film Grain Synthesis: Material Selection for Ground Truth Content Creation | D. Podborski, D. Ugur, B. Williams, L. Levinson, R. Molholm, A.M. Tourapis |
| [JVET-AE0251](https://jvet-experts.org/doc_end_user/current_document.php?id=13214) | m64416 | 2023-07-11 10:09:09 |  |  | Evaluating Denoising methods in the context of Film Grain Analysis/Synthesis | D. Podborski, D. Ugut, A. M. Tourapis |

**Recommendations**

The AHG recommends:

* To review and consider JVET-AE0219 in the finalization of the SVVC verification test plan.
* To review JVET-AD0179 and consider further steps towards resolving licensing issues with 3DV test content.
* To review and consider JVET-AE0250 and JVET-AE0251 in the development of the draft plan for subjective quality testing of FGC SEI message
* To collect volunteers to conduct further verification tests and subjective quality tests, including volunteers to encode.
* To continue to discuss and to update the non-finalized categories of the verification test plan and subjective quality test plan for FGS, including those which have not been addressed yet.
* To 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.

Test in multi-layer expected after the meeting.

[JVET-AE0005](https://jvet-experts.org/doc_end_user/current_document.php?id=13077) JVET AHG report: Conformance testing (AHG5) [D. Rusanovskyy, I. Moccagatta (co-chairs), F. Bossen, K. Kawamura, T. Ikai, S. Iwamura, H.-J. Jhu, K. Sühring, Y. Yu (vice-chairs)]

**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 30th and 31st meetings.

**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, ready to proceed to FDAM at current meeting 2023-01
  + H.266.1 V2 Consent planned for 2023-07

**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
  + As reported in JVET-AB0005, HRD\_A\_Fujitsu (HRD\_A\_Fujitsu\_4) has been regenerated to avoid CPB underflow
  + The 10 bitstreams reported in Ticket #1581 have been re-generated
  + The problems reported in Ticket #1601 for 2 FDIS\_r1 packages have been fixed
* 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 30th and 31st meeting.

**Activities and Discussion**

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

VVC activities:

The remaining 3 of the 10 bitstreams reported in Ticket #1581

* OLS\_A\_Tencent\_5
* OLS\_B\_Tencent\_5
* OLS\_C\_Tencent\_5

have been re-generated after the VTM problem reported in Ticket #1595 has been fixed.

The re-generated bitstreams, HRD\_A\_Fujitsu\_4, and the remaining conformance streams are decoded correctly using VTM-20.2 w/o range extension support. The streams have been uploaded to <https://www.itu.int/wftp3/av-arch/jvet-site/bitstream_exchange/VVC/under_test/VTM-20.2/>

Ticket #1601 has been filed to report that the .yuv.md5 and .opl files of 2 FDIS\_r1 packages (STILL\_B\_1 and STILL444\_B\_1) are different from the same files in the corresponding FDIS packages (STILL\_B\_ ERICSSON\_1 and STILL444\_B\_ ERICSSON\_1). Test bitstreams and collateral files in the FDIS and FDIS\_r1 directories are supposed to be the same except for removal of company names and email addresses. The differences reported in Ticket #1601 are unintentional and have been fixed. The fixed packages have been uploaded to <https://www.itu.int/wftp3/av-arch/jvet-site/bitstream_exchange/VVC/under_test/VTM-20.2_r1/>

VVC operation range extensions activities:

All bitstreams are decoded correctly using VTM-20.2 and are available at <https://www.itu.int/wftp3/av-arch/jvet-site/bitstream_exchange/VVCv2/under_test/VTM-20.2/>

VVC Multilayer activities:

The 6 additional conformance bitstreams for VVC multilayer configurations in JVET-AD2028 are decoded correctly using VTM-20.2 and are available in <https://www.itu.int/wftp3/av-arch/jvet-site/bitstream_exchange/VVCMultilayer/under_test/VTM-20.2/>

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.

**Contributions**

JVET-AE0111 - Additional conformance tests of spatial scalability for multilayer coding [C. Salmon-Legagneur, P. de Lagrange, F. Urban].

JVET-AE0181 - Scaling window support for VTM [Shunsuke Iwamura, Shimpei Nemoto, Atsuro Ichigaya].

**Ftp site information**

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

* VVC:

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

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

* VVC operation range extensions:

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

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

The ftp site for uploading bitstream file is as follows.

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

(user id: avguest, passwd: Avguest201007)

If using FileZilla, the following configuration is suggested:

Graphical user interface, text, application, email

Description automatically generated

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

**Recommendations**

The AHG recommends the following:

* Proceed with the generation, cross-checking, and documentation of the additional conformance bitstreams for VVC, in particular conformance bitstreams for VVC multilayer configurations (JVET-AD2028).
* Maintain and update the conformance bitstream database and contribute to report problems in JVET document 1004.
* Start the generation, cross-checking, and documentation of the conformance streams for the HEVC multiview profiles supporting extended bit depth (JVET-AA1011).

All issues with V2 bitstreams have been resolved. Text integration was started. Intent to submit to ITU consent, ITU version of the integrated text to be provided until Wednesday 19. Gary Sullivan will provide the basis text, and provide help in differences relative to ISO version.

[JVET-AE0006](https://jvet-experts.org/doc_end_user/current_document.php?id=12989) JVET AHG report: ECM software development (AHG6) [V. Seregin (chair), J. Chen, F. Le Léannec, K. Zhang (vice-chairs)]

**Software development**

ECM software repository is located at [https://vcgit.hhi.fraunhofer.de/ecm/ECM.](https://vcgit.hhi.fraunhofer.de/ecm/ECM.E)

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 adopted aspects were integrated into ECM-9.0:

JVET-AD0202 (EE2-1.2): CCCM with multiple downsampling filters (MR 399)

JVET-AD0208 (EE2 Test-1.8c): IBC adaptation for coding of natural content (MR 405)

JVET-AD0195: High-Precision MV Refinement for BDOF (MR 406)

JVET-AD0222 (EE2-4.4b): ALF: additional fixed filter, residual-based classifier, increase the number of taps that are applied to the fixed filter results (MR 402)

JVET-AD0206: CABAC init at GDR picture (MR 407)

JVET-AD0123 (EE2 Test 5.1): reference frame padding for GDR (MR 409)

JVET-AD0193 (EE2 Test 2.11e): Adaptive OBMC control (MR 408)

JVET-AD0213 (EE2-Test2.7b): bi-predictive LIC and the combination of LIC and OBMC (MR 404)

JVET-AD0082: add configuration option for TMRL (MR 401)

JVET-AD0169: Downsampling filters in range 1.1 to 1.35 based on Kaiser(7) windowed sinc (MR 412)

JVET-AD0184: Removal of Division Operations (MR 413)

JVET-AD0085: Combinations of MPM related tests (MR 400)

JVET-AD0086 (EE2-1.20j combination test): enhanced intra TMP (MR 415)

JVET-AD0188 (EE2 Test-1.6c): Non-local cross-component prediction and cross-component merge mode (MR 418)

JVET-AD0105 (Aspect1): Cap the maximum NumSignPred to 4 for QP<=22 (AI and RA) (MR 403)

JVET-AD0182(EE2-2.1l): Affine Parameter refinement and mode extension (MR 419)

JVET-AD0239/AD0124: Fixes for IntraTMP template availability (MR 421)

JVET-AD0120 (EE2 Test 1.3e): Local-Boosting Cross-Component Prediction (MR 411)

JVET-AD0140: MVD (sign and suffix) prediction, test 2.12a (MR 424)

Bug fixes:

Revert MR383 and fix (MR 396)

Fix #45 on IBC AMVR off (MR 397)

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

JVET-AD0169: Downsampling filters in range 1.1 to 1.35 based on Kaiser(7) windowed sinc (MR 422)

It does not change VTM-11ecm anchor performance under CTC.

ECM-9.0 and VTM-11ecm9.0 were tagged on May 24, 2023.

The following bug fixes were integrated into ECM-9.0:

Fix mismatch with TIMD is 0 (MR 434)

Fix luma offsets in AD0188 (MR 436)

RPR-related fixes for JVET\_AD0140\_MVD\_PREDICTION (MR 435)

CCCM code refactoring (MR 431)

Fix mismatch when --EnableTMTools=0 (MR 438)

Fix crash when --EnableTMTools=0 (MR 440)

Avoid temporary cu copy in EncModeCtrlMTnoRQT::useModeResult (MR 441)

Fix AD0188 for GDR (MR 439)

RPR fixes for JVET\_AD0213\_LIC\_IMP and JVET\_AD0140\_MVD\_PREDICTION (MR 437)

ECM-9.0 was tagged on July 5, 2023.

***CTC* Performance**

In this section, ECM test results following ECM CTC configuration descried in JVET-AD2017 are summarized.

Next tables show ECM-9.0 performance over ECM-8.1 anchor.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | |
|  | **Over ECM-8.1** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -0.70% | -3.10% | -4.16% | 122.0% | 114.0% |
| Class A2 | -2.48% | -6.26% | -6.55% | 121.9% | 115.1% |
| Class B | -1.91% | -5.16% | -4.85% | 121.8% | 115.6% |
| Class C | -1.16% | -2.24% | -2.41% | 126.2% | 122.0% |
| Class E | -3.50% | -5.92% | -5.33% | 128.6% | 126.1% |
| **Overall** | -1.90% | -4.48% | -4.56% | 123.9% | 118.3% |
| Class D | -1.20% | -2.49% | -2.55% | 127.9% | 126.2% |
| Class F | -2.80% | -4.30% | -4.39% | 120.8% | 126.2% |
| Class TGM | -2.62% | -4.09% | -4.04% | 120.8% | 137.6% |
|  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | |
|  | **Over ECM-8.1** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -1.22% | -1.88% | -3.64% | 116.3% | 111.1% |
| Class A2 | -1.96% | -3.81% | -3.63% | 117.9% | 109.4% |
| Class B | -1.51% | -3.75% | -3.55% | 125.2% | 110.3% |
| Class C | -1.19% | -1.24% | -1.37% | 126.7% | 117.3% |
| Class E |  |  |  |  |  |
| **Overall** | -1.46% | -2.72% | -3.00% | 122.3% | 112.1% |
| Class D | -1.40% | -1.18% | -1.79% | 131.8% | 116.7% |
| Class F | -2.53% | -3.40% | -3.91% | 134.4% | 119.7% |
| Class TGM | -2.90% | -4.48% | -4.16% | 119.3% | 113.4% |
|  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | |
|  | **Over ECM-8.1** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -0.78% | -1.87% | -1.75% | 125.4% | 113.1% |
| Class C | -0.99% | -0.35% | -0.65% | 120.0% | 115.8% |
| Class E | -0.83% | -1.10% | -3.12% | 123.6% | 113.0% |
| **Overall** | -0.86% | -1.17% | -1.73% | 123.1% | 114.0% |
| Class D | -0.75% | -2.08% | -0.28% | 129.7% | 119.9% |
| Class F | -1.72% | -3.46% | -2.95% | 131.6% | 123.5% |
| Class TGM | -2.67% | -3.96% | -3.70% | 119.3% | 113.2% |
|  |  |  |  |  |  |
|  | **Low delay P Main 10** | | | | |
|  | **Over ECM-8.1** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -0.41% | -1.24% | -1.57% | 123.1% | 106.7% |
| Class C | -0.52% | -0.42% | -0.80% | 126.5% | 116.1% |
| Class E | -0.85% | -1.96% | -3.58% | 133.4% | 116.8% |
| **Overall** | -0.56% | -1.15% | -1.81% | 126.7% | 112.2% |
| Class D | -0.52% | -1.77% | -0.13% | 138.7% | 120.6% |
| Class F | -1.19% | -2.47% | -2.56% | 130.1% | 120.6% |
| Class TGM | -2.65% | -4.36% | -3.91% | 114.0% | 108.0% |

The below tables show ECM-9.0 performance comparing to VTM-11.0ecm9.0 anchor.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | |
|  | **Over VTM-11.0ecm9.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -9.67% | -21.16% | -28.19% | 827.7% | 406.4% |
| Class A2 | -15.77% | -27.44% | -30.55% | 813.4% | 405.9% |
| Class B | -10.06% | -25.61% | -23.82% | 773.3% | 423.4% |
| Class C | -10.27% | -15.86% | -16.33% | 783.1% | 442.4% |
| Class E | -13.65% | -23.55% | -22.06% | 759.1% | 481.4% |
| **Overall** | -11.59% | -22.66% | -23.71% | 788.5% | 430.8% |
| Class D | -8.61% | -13.65% | -13.74% | 803.2% | 462.1% |
| Class F | -24.34% | -33.43% | -33.28% | 529.7% | 445.8% |
| Class TGM | -37.29% | -45.21% | -44.55% | 443.8% | 443.1% |
|  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | |
|  | **Over VTM-11.0ecm9.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -20.86% | -24.25% | -32.21% | 757.7% | 741.6% |
| Class A2 | -23.90% | -32.20% | -35.28% | 697.8% | 882.0% |
| Class B | -19.50% | -31.71% | -29.62% | 648.2% | 739.3% |
| Class C | -20.90% | -24.61% | -24.58% | 672.8% | 779.1% |
| Class E |  |  |  |  |  |
| **Overall** | -21.03% | -28.42% | -29.93% | 685.5% | 777.1% |
| Class D | -21.87% | -25.90% | -26.55% | 699.1% | 849.6% |
| Class F | -26.77% | -34.78% | -35.14% | 578.7% | 528.1% |
| Class TGM | -35.17% | -41.37% | -41.35% | 547.2% | 419.4% |
|  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | |
|  | **Over VTM-11.0ecm9.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -17.01% | -32.11% | -30.48% | 601.4% | 658.9% |
| Class C | -18.50% | -26.42% | -26.63% | 590.5% | 661.3% |
| Class E | -15.40% | -25.47% | -26.14% | 595.7% | 440.1% |
| **Overall** | -17.10% | -28.56% | -28.11% | 596.3% | 596.4% |
| Class D | -20.39% | -28.41% | -27.33% | 629.6% | 770.9% |
| Class F | -23.36% | -33.28% | -31.66% | 562.9% | 526.4% |
| Class TGM | -33.93% | -40.53% | -40.11% | 511.1% | 404.8% |
|  |  |  |  |  |  |
|  | **Low delay P Main 10** | | | | |
|  | **Over VTM-11.0ecm9.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -15.87% | -31.42% | -29.54% | 537.3% | 625.1% |
| Class C | -17.58% | -25.77% | -25.54% | 521.8% | 661.5% |
| Class E | -14.93% | -24.37% | -25.89% | 579.6% | 462.0% |
| **Overall** | -16.20% | -27.77% | -27.29% | 542.3% | 590.6% |
| Class D | -20.36% | -28.35% | -27.51% | 563.4% | 739.4% |
| Class F | -22.24% | -33.15% | -31.57% | 541.3% | 486.8% |
| Class TGM | -30.83% | -38.46% | -38.16% | 526.4% | 378.7% |

Next tables show ECM-9.1 performance over ECM-9.0 anchor.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | |
|  | **Over ECM-9.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -0.03% | -0.22% | -0.36% | 101.8% | 105.0% |
| Class A2 | -0.03% | -0.11% | -0.08% | 101.9% | 106.8% |
| Class B | -0.02% | -0.22% | -0.02% | 101.4% | 103.8% |
| Class C | 0.00% | -0.13% | -0.06% | 101.7% | 100.3% |
| Class E | -0.02% | -0.05% | -0.06% | 99.1% | 98.1% |
| **Overall** | -0.02% | -0.15% | -0.10% | 101.2% | 102.7% |
| Class D | 0.00% | -0.05% | 0.03% | 97.9% | 98.7% |
| Class F | -0.02% | -0.16% | -0.03% | 98.3% | 99.7% |
| Class TGM | -0.03% | -0.06% | -0.04% | 98.3% | 102.9% |
|  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | |
|  | **Over ECM-9.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -0.01% | -0.21% | -0.07% | 103.8% | 104.0% |
| Class A2 | -0.02% | -0.21% | -0.09% | 103.7% | 103.8% |
| Class B | 0.00% | -0.16% | -0.07% | 100.5% | 102.2% |
| Class C | -0.01% | -0.12% | -0.02% | 99.1% | 99.7% |
| Class E |  |  |  |  |  |
| **Overall** | -0.01% | -0.17% | -0.06% | 101.4% | 102.2% |
| Class D | 0.00% | -0.33% | 0.16% | 99.2% | 99.1% |
| Class F | -0.02% | -0.14% | -0.06% | 95.6% | 93.3% |
| Class TGM | 0.01% | 0.04% | -0.03% | 101.4% | 105.5% |
|  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | |
|  | **Over ECM-9.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | 0.05% | -0.09% | -0.04% | 103.0% | 104.3% |
| Class C | 0.02% | -0.42% | -0.59% | 99.3% | 102.2% |
| Class E | 0.13% | -0.16% | 1.16% | 99.1% | 101.7% |
| **Overall** | 0.06% | -0.22% | 0.08% | 100.8% | 102.9% |
| Class D | -0.12% | 1.38% | -0.03% | 99.0% | 99.3% |
| Class F | -0.09% | -0.02% | 0.29% | 96.9% | 97.6% |
| Class TGM | -0.01% | 0.08% | -0.06% | 103.4% | 108.5% |
|  |  |  |  |  |  |
|  | **Low delay P Main 10** | | | | |
|  | **Over ECM-9.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | 0.00% | 0.37% | -0.01% | 103.9% | 108.4% |
| Class C | -0.03% | -0.16% | 0.08% | 100.1% | 102.4% |
| Class E | 0.02% | -0.61% | 0.87% | 98.2% | 99.3% |
| **Overall** | 0.00% | -0.05% | 0.24% | 101.2% | 104.0% |
| Class D | 0.07% | 1.22% | 1.05% | 99.9% | 100.9% |
| Class F | 0.15% | 0.38% | -0.40% | 98.9% | 97.4% |
| Class TGM | -0.07% | -0.06% | -0.16% | 107.8% | 110.8% |

The Excel files with the complete test results are attached to this report.

**ECM memory consumption**

ECM memory consumption is provided in ECM encoder log files and is summarized in the table below.

|  |  |  |  |
| --- | --- | --- | --- |
|  | RA | LD | AI |
| class A | 25 GB |  | 12 GB |
| class B | 14 GB | 12 GB | 10 GB |
| class C | 6 GB | 6 GB | 6 GB |
| class E |  | 8 GB | 7 GB |
| class D | 5 GB | 5 GB | 5 GB |
| class F | 14 GB | 12 GB | 11 GB |
| class TGM | 16 GB | 14 GB | 11 GB |

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 allocated memory wherever possible, rather than systematically allocating new memory.

**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.

[JVET-AE0007](https://jvet-experts.org/doc_end_user/current_document.php?id=13078) JVET AHG report: ECM tool assessment (AHG7) [X. Li (chair), L.-F. Chen, Z. Deng, J. Gan, E. François, H.-J. Jhu, X. Li, H. Wang (vice-chairs)]

**Group off tests**

***Test settings and crosschecking***

Based on the discussion in the previous meeting, the same four groups were used in this meeting cycle.

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

Accordingly, five group-off tests were performed and crosschecked on top of ECM-9.1.The two anchors are ECM 9.0 and VTM-11ECM9.0. The cfg files used are also attached with this report.

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

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

***Group 1 off***

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



***Group 2 off***

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



***Group 3 off***

Group 3 includes intra and IBC template matching (with search) related tools. The attached offgroup3.cfg was used in addition to ECM CTC settings. Note that some parts of template matching for intra and IBC cannot be correctly switched off in ECM-9. The impacted tools are LBCCP and ARMC.



***Group 4 off***

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

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



***Group 1-4 off***

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



***Summary***

The average RA results compared to ECM-9 are summarized below. Note that SCC results are not included. More figures, such as using ECM-9 as the anchor, may be found in the attached spreadsheet summary.

**BD-rate Y (over VTM11-ECM9) vs Encoding Time**

**Weighted BD-rate YUV (over VTM11-ECM9) vs Encoding Time**

Note that weighting factor 6:1:1 is used in the figure below. In the attached summary, different weighting factor may be specified.

**Bd-rate Y (over VTM11-ECM9) vs Decoding Time**

**Resolved issues and new issues**

***Resolved issues***

* Software issues #27 and #45 with tool off tests were resolved
* It has been verified in the tests that the memory usage ratio is relatively stable

***New issues***

* In ECM-9 some parts of template matching for intra and IBC cannot be correctly switched off, which impact several coding tools, such as LBCCP, ARMC
* Memory usage ratio keeps 100% in the group off tests, though 7-11% memory increase by ECM-9 over ECM-8.1 was observed
  + Two possible reasons 1) The memory usage increase in ECM-9 over ECM-8 was not introduced by the tested tools; 2) Memory was always allocated even a tool is disabled
* The variation of runtime over VTM11ECM9 is relatively large, for which the version of compiler matters

**Input contributions**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| [JVET-AE0174](https://jvet-experts.org/doc_end_user/current_document.php?id=13137) | m63905 | 2023-07-05 02:00:58 | 2023-07-05 04:12:11 | 2023-07-05 04:12:11 | AHG7: On TM control for non-inter tools | Z. Deng, K. Zhang, L. Zhang (Bytedance) |
| [JVET-AE0180](https://jvet-experts.org/doc_end_user/current_document.php?id=13143) | m63913 | 2023-07-05 09:23:52 | 2023-07-05 09:33:31 | 2023-07-05 09:37:18 | On ECM SW memory consumption | R. Chernyak, S. Liu (Tencent), Y. Yasugi, T. Ikai (Sharp) |
| [JVET-AE0195](https://jvet-experts.org/doc_end_user/current_document.php?id=13158) | m63962 | 2023-07-06 15:05:00 | 2023-07-06 18:05:21 | 2023-07-06 18:05:21 | AhG7: ECM stages complexity assessment | F. Galpin, F. Leleannec, C. Salmon-Legagneur, E. François (InterDigital), |

**Recommendations**

* Continue and improve tool assessment
* Report memory usage ratio in ECM CTC spread sheet, as suggested in the BoG report in JVET-AD0401
* Resolve identified software issues related to the tool assessment
* Review all the input contributions

It was commented that recording the run times with two digits after decimal point may not be useful in assessing complexity.

[JVET-AE0008](https://jvet-experts.org/doc_end_user/current_document.php?id=13080) JVET AHG report: Optimization of encoders and receiving systems for machine analysis of coded video content (AHG8) [C. Hollmann, S. Liu, S. Wang, M. Zhou (AHG chairs)]

**Activities**

The AHG used the main JVET reflector, jvet@lists.rwth-aachen.de, for email discussion. There were a few emails exchanged on the reflector uner the context of AHG 8 during this period. There were more email exchanges and discussions among co-chairs, editors and interested experts related to technical report preparation, common test conditions and anchor generation, etc. There was one joint meeting with WG 4 held on 2023-06-21. There are 12 input contriubtions related to AHG 8 mandates submitted to this meeting. They are listed in Section 3.

***Common Test Conditions***

Common test conditions (CTC) for optimization of encoders and receiving systems for machine analysis of coded video content, are summarized in output document JVET-AD2031. This document includes detailed descriptions of test datasets, anchor software and configurations, anchor generation processes, machine task networks used, test and training conditions, evaluation methodologies and metrics. A reporting template in Excel format is enclosed in the document package. This output document package was uploaded on 2023-05-11, with updated anchor results, and is also available at <https://vcgit.hhi.fraunhofer.de/jvet-ahg-ofm>/ofm-ctc.

**Test datasets**

Two video datasets, referred as SFU-HW and TVD, are included in the CTC and results on both are expected to be reported. In addition, three image datasets, referred as TVD (image), OpenImageV6 and FLIR, are included in the CTC and can be tested using intra-only configuration. The results on the image datasets are optional.

It was decided in the last JVET meeting to add “Kimono” sequence (back to) the SFU-HW dataset, as a separate optional (O) class, and thus the SFU-HW dataset now consist of 14 test video sequences, as shown in Table 1. These sequences can be found on <ftp://hevc@mpeg.tnt.uni-hannover.de>. The annotations are available at <https://data.mendeley.com/datasets/hwm673bv4m/1>.

Table 1. Test sequences in SFU-HW dataset

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Class** | **Sequence name** | **Frame count** | **Frame rate** | **Bit depth** | **Frames skipped** | **Frames coded** | **Intra** | **RA** | **LD** |
| A | Traffic | 150 | 30 | 8 | 117 | 33 | M | M | M |
| B | ParkScene | 240 | 24 | 8 | 207 | 33 | M | M | M |
| B | Cactus | 500 | 50 | 8 | 403 | 97 | M | M | M |
| B | BasketballDrive | 500 | 50 | 8 | 403 | 97 | M | M | M |
| B | BQTerrace | 600 | 60 | 8 | 471 | 129 | M | M | M |
| C | RaceHorses | 300 | 30 | 8 | 235 | 65 | M | M | M |
| C | BQMall | 600 | 60 | 8 | 471 | 129 | M | M | M |
| C | PartyScene | 500 | 50 | 8 | 403 | 97 | M | M | M |
| C | BasketballDrill | 500 | 50 | 8 | 403 | 97 | M | M | M |
| D | RaceHorses | 300 | 30 | 8 | 235 | 65 | M | M | M |
| D | BQSquare | 600 | 60 | 8 | 471 | 129 | M | M | M |
| D | BlowingBubbles | 500 | 50 | 8 | 403 | 97 | M | M | M |
| D | BasketballPass | 500 | 50 | 8 | 403 | 97 | M | M | M |
| O | Kimono | 240 | 24 | 8 | 207 | 33 | O | O | O |

*Note: M – mandatory; O – optional.*

It was decided in the last JVET meeting to replace some of the compressed source video sequences in Tencent Video Dataset (TVD) by their uncompressed versions. The updated source video sequences are renamed, as shown in Table 2, and can be downloaded from <https://multimedia.tencent.com/resources/tvd> with corresponding annotations.

Table 2. Test sequences in TVD

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Sequence name** | **Frame count** | **Frame rate** | **Bit depth** | **Frames skipped** | **Frames coded** | **Intra** | **Random access** | **Low delay** |
| TVD-01-1 | 3000 | 50 | 8 | 1500 | 500 | M | M | M |
| TVD-01-2 | 3000 | 50 | 8 | 2000 | 500 | M | M | M |
| TVD-01-3 | 3000 | 50 | 8 | 2500 | 500 | M | M | M |
| TVD-02-1 | 636 | 50 | 10 | 0 | 636 | M | M | M |
| TVD-03-1 | 2334 | 50 | 10 | 0 | 500 | M | M | M |
| TVD-03-2 | 2334 | 50 | 10 | 500 | 500 | M | M | M |
| TVD-03-3 | 2334 | 50 | 10 | 1000 | 500 | M | M | M |

As the original sequences are available in .mp4 format, they shall be converted to YUV420 using FFmpeg. The format conversion process was cleaned up as follows:

ffmpeg -i {input.mp4} {output.yuv}

The TVD (image) dataset is an image dataset of 166 images of 1920x1080 resolution that have annotations for object detection and instance segmentation. The dataset with corresponding annotations is available at <https://multimedia.tencent.com/resources/tvd>.

The OpenImages dataset consists of around 9 million images. A subset of the validation set of its version 6 containing 5000 images are selected for testing object detection in this activity. The dataset with corresponding annotations is available at <https://storage.googleapis.com/openimages/web/index.html>.

The FLIR dataset used in the WG 4 VCM activities is a dataset consisting of 300 infrared images. The images, annotations and the fine-tuned model for thermal images can be found on the MPEG content repository (<https://content.mpeg.expert/data/>).

More detailed information about these test datasets can be found in JVET-AD2031.

**Anchor software and configuration**

Version 12.0 of the VTM software is used for generating anchor results. The VTM software is available at <https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/>.

In addition, version 4.2.2 of the FFmpeg is used for the format conversion. The FFmpeg software is available at <https://ffmpeg.org/releases>.

Three test conditions are used for experimenting encoder and receiving system optimizations on video coding for machine consumptions, i.e., random-access, low-delay and all-intra. The default configuration files provided with the VTM software are used for anchor generation.

* “Random access” (RA): encoder\_randomaccess\_vtm.cfg
* “Low delay” (LD): encoder\_lowdelay\_vtm.cfg
* “All Intra” (AI): encoder\_intra\_vtm.cfg

It was decided in the last JVET meeting that temporal subsampling is disabled for “All Intra” configuration for this activity, thus the following line

TemporalSubsampleRatio : 8

in “encoder\_intra\_vtm.cfg” needs to be modified to

TemporalSubsampleRatio : 1

A subset of these configurations might be used for a particular experiment. However, as these test conditions are for video coding experiments, results for at least one of the random access or low delay configurations are expected to be provided.

**Anchor generation**

The anchor generation process is illustrated in Figure 1. The task network Faster R-CNN X101-FPN which is part of Detectron2 (<https://github.com/facebookresearch/detectron2>) is used to evaluate object detection (SFU-HW dataset and image datasets), the model is available [here](https://dl.fbaipublicfiles.com/detectron2/COCO-Detection/faster_rcnn_X_101_32x8d_FPN_3x/139173657/model_final_68b088.pkl). The task network JDE-1088x608 which is part of the Towards Realtime MOT framework (<https://github.com/Zhongdao/Towards-Realtime-MOT>) is used to evaluate object tracking (TVD), the model is available [here](https://drive.google.com/open?id=1nlnuYfGNuHWZztQHXwVZSL_FvfE551pA). More details about anchor generation process and these machine task networks are referred to JVET-AD2031.

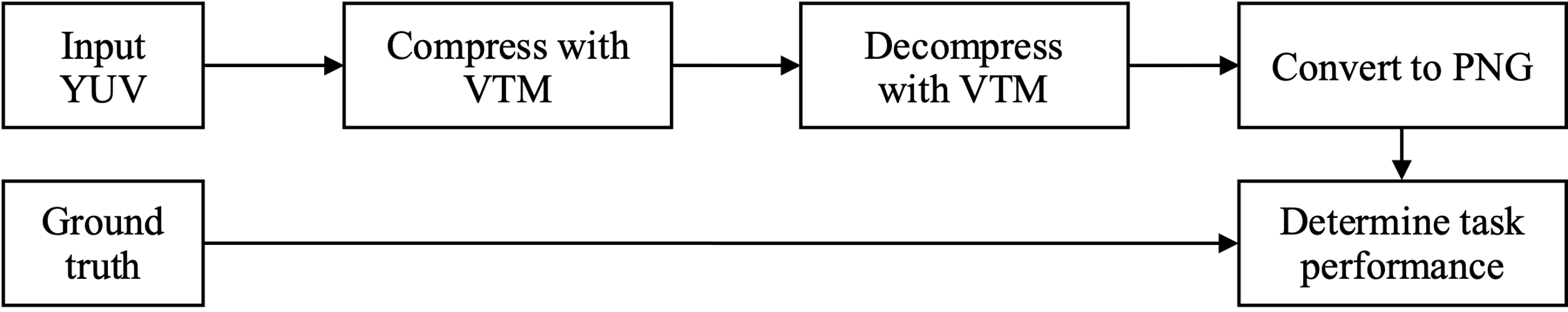


Figure 1. Anchor generation pipeline

Anchor results were generated following the CTC described in JVET-AD2031 and crosschecked by Ericsson and Tencent.

**Evaluation and reporting**

Proposed technologies are evaluated based on their compression performance, measured by bitrate, PSNR, mAP and MOTA, as well as encoding and decoding runtime to reflect the complexity of the proposed technology, to some extent. Definitions and detailed descriptions of these metrics can be found in JVET-AD2031.

It is noted that the mean Average Precision (mAP) is used to measure object detection performance, and Multiple Object Tracking Accuracy (MOTA) is used to measure object tracking performance. For the purpose of reporting encoding and decoding running times, the anchor and proposal should be simulated on the same platform, e.g. similar CPU and GPU configuration, to have reliable time comparison. Parallel encoding and decoding as described in JVET-B0036 may be applied for RA configurations.

In addition, relevant inference and training information should be reported if the proposed technology consists of learning-based components, such as network structure, number of network parameters, precision of network parameters, number of multiply–accumulate operations (MAC) per pixel, patch size, batch size, epoch, training time, training datasets, lost function, number of iterations, optimizer, and any pre- and post-processings used. More details can be found in JVET-AD2031.

A reporting template in Excel file format has been prepared to illustrate results following CTC and evaluation methodology described in JVET-AD2031. It is enclosed in JVET-AD2031 package.







Figure 2. Reporting template summary

***Technical Report***

The draft 2 of the technical report (TR) has been prepared and uploaded as JVET-AD2030 based on discussions in the last JVET meeting on 2023-06-07. More descriptions were added to use cases and applications, pre-processing technologies and encoding technologies. Two software implementation examples were included in the TR draft 2 as Annex.

***Git Management***

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

* JVET-AB0275: a region of interest-based method that uses adaptive QP to reduce the quality in background areas
* JVET-AC0086: a method that uses a pre-analysis to perform content adaptive machine vision oriented preprocessing

***AHG Coordination***

Collaborative discussions were held between this AHG and the VCM AHG in WG4 to synchronize common test conditions including test datasets, anchor QP points, encoder configurations, etc. A joint meeting was held on 2023-06-21 0500UTC to discuss metrics, existing issues and potential solutions for machine task performance evaluation. Two contributions were presented and discussed during this meeting:

**m63676 - [VCM] Proposed performance evaluation method for VCM CE1 (ETRI, Konkuk Univ., Myongji Univ.)**

This contribution proposes a performance evaluation method for VCM CE1.

Regarding the performance evaluation for VCM CE1, there are some difficulties as follows:

1. The current BD-rate metric provided by the reporting template is not able to produce a value if either the anchor or a proposal is non-monotonic.
2. To compensate for this limitation, the reporting template requires submission of BD-rate for multiple subcases (combination of 4 or 5 points out of 6 total points), but it is still difficult to make a direct comparison between proposals since the metric is not singular.

The following rules were proposed as the performance evaluation method for VCM CE1:

1. For each test case, a proponent should submit 4 points that correspond to the high 4 rate points of the VTM anchor.
2. Only if the result at the high 4 rate points is non-monotonic, then it is allowed to adjust QP in a range of [-2, +2] to produce a monotonic result.
3. If a monotonic result cannot be obtained even with the adjustment of QP in the range, then it is not considered as a candidate technology in this round of CE1 evaluation.
4. Evaluate the proposed technologies using the summed scores based on the class-wise (SFU) or sequence-wise (TVD) BD-rate rankings of the proposals.

**m63692 - [VCM] Improvements of the BD-rate model using monotonic curve-fitting method (Tencent)**

This contribution provides improvements of the BD-rate model used in current Excel templates. Specifically, we propose a monotonic fitted curve to replace the PCHIP (Piecewise Cubic Hermit Interpolation) curve for BD-rate and BD-metric calculation, such as BD-PSNR, BD-mAP, BD-MOTA, etc. The fitted curve guarantees monotonicity even if the input points are nonmonotonic, thus makes it possible to calculate BD-rate and BD-metric values in various testing conditions.

This contribution was also submitted to JVET (AHG 8) as document JVET-AE0107.

**Input contributions**

There are 12 input contributions related to AHG 8 mandates. They are listed below.

|  |  |  |
| --- | --- | --- |
| **Report** | | |
| JVET-AE0008 | JVET AHG report: Optimization of encoders and receiving systems for machine analysis of coded video content (AHG8) | C. Hollmann, S. Liu, S. Wang, M. Zhou (AHG chairs) |
| **Proposal** | | |
| JVET-AE0053 | AHG8/AHG9: Neural-network post-filter regions SEI message | T. Chujoh, Y. Yasugi, T. Ikai (Sharp) |
| JVET-AE0064 | AHG8/AHG9: Signalling encoder preprocessing and human / machine viewing indications | C. Kim, D. Gwak, Hendry, J. Lim, S. Kim (LGE), M. M. Hannuksela, F. Cricri, H. Zhang (Nokia) |
| JVET-AE0079 | AHG8/AHG9: Source picture timing information SEI message | S. McCarthy, G.J. Sullivan, P. Yin (Dolby) |
| JVET-AE0081 | [AHG8] De-noising filter as pre-processing for machine task | C. Kim, D. Gwak, J.Lim (LGE) |
| JVET-AE0090 | AHG8/AHG9: On machine vision indication | J. Gao, H.-B. Teo, C.-S. Lim, K. Abe (Panasonic) |
| JVET-AE0095\* | AHG8/AHG9: proposed changes to the candidate new object mask information SEI message | P. de Lagrange, E. François, D. Doyen (InterDigital), J. Chen, S. Wang, Y. Ye (Alibaba) |
| JVET-AE0096 | [AHG8] Study on using different VTM versions | C. Hollmann, J. Ström (Ericsson) |
| JVET-AE0099 | [AHG8] NNPF and post-filter hint SEI messages for the technical report | C. Hollmann, M. Pettersson, R. Sjöberg, M. Damghanian (Ericsson) |
| JVET-AE0107 | AHG8: Improvements of the BD-rate model using monotonic curve-fitting method | H. Wang, X. Pan, Z. Liu, X. Xu, S. Liu (Tencent) |
| JVET-AE0143 | AHG8: A spatial resampling algorithm and an exemplar software implementation | S. Wang, B. Li, J. Chen, Y. Ye (Alibaba), S. Wang (CityU) |
| **Crosscheck** | | |
| JVET-AE0234 | AHG8: Crosscheck of JVET-AE0107 | Honglei Zhang (Nokia) |

\* Late.

**Recommendations**

The AHG recommends to:

* Review all input contributions.
* Discuss and refine test conditions, evalution and reporting procedures.
* Discuss the non-monotonic issue and potential solutions.
* Continue investigating non-normative technologies and their suitability for machine analysis applications.
* Continue developing draft technical report on optimization of encoders and receiving systems for machine analysis of coded video content.
* Continue collecting new test materials.

It was pointed out that a potential timeline for the technical report should be discussed. This however may also relate to the level of completeness in terms of using video compression standards for a variety of machine usage applications, which somewhat relies on the availability of test material of sufficient variety.

[JVET-AE0009](https://jvet-experts.org/doc_end_user/current_document.php?id=13082) JVET AHG report: SEI message studies (AHG9) [S. McCarthy, Y.-K. Wang (co-chairs), T. Chujoh, S. Deshpande, C. Fogg, Hendry, P. de Lagrange, G. J. Sullivan, A. Tourapis, S. Wenger (vice-chairs)]

**Related contributions**

A total of 42 contributions and 1 bug report are identified relating to the mandates of AHG9. Some contributions also relate to the work of AHG2 and AHG8.

The number of contributions relating to each AHG9 mandate is as follows

* 31 contributions relate to the mandate to study the SEI messages in VSEI, VVC, HEVC, and AVC. 2 of the 31 contributions relate to more than one SEI message.
  + 28 contributions relate to the neural-network post-filter SEI messages, including 1 summary of proposals on NNPF SEI messages.
  + 1 contribution relates to the annotated region SEI message
  + 2 contribution relates to the post filter hint SEI message
  + 1 contribution relates to the phase indication SEI message
  + 1 contribution relates to the SEI processing order SEI message
* 1 contribution relates to the mandate to collect software and showcase information for SEI messages;
* 10 contributions relate to the mandate to identify potential needs for additional SEI messages.
  + 2 contributions relate to human/machine viewing
  + 1 contribution relates to multiplane image information
  + 2 contributions relate to alternative picture timing
  + 3 contributions relate to generative face video
  + 1 contribution relates to object masking information
  + 1 contribution relates to object-wave compression and computer-generated hologram use
* 1 bug report relates to the mandate to coordinate with AHG3 for software support of SEI messages

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

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

**Neural-network post filter characteristics and activation SEI messages**

**Summary of proposals on NNPF SEI messages**

[JVET-AE018](https://www.jvet-experts.org/doc_end_user/current_document.php?id=13150)7 AHG9: A summary of SEI proposals on NNPF [Y.-K. Wang (Bytedance)]

**Proposals on NNPF SEI messages**

[JVET-AE0048](https://jvet-experts.org/doc_end_user/current_document.php?id=12996) AHG9: Miscellaneous VSEI changes on neural-network post-filter SEI messages [M. M. Hannuksela, F. Cricri, M. Santamaria (Nokia)]

[JVET-AE0049](https://jvet-experts.org/doc_end_user/current_document.php?id=12997) AHG9: Miscellaneous VVC changes on neural-network post-filter SEI messages [M. M. Hannuksela, F. Cricri, M. Santamaria (Nokia)]

[JVET-AE0050](https://jvet-experts.org/doc_end_user/current_document.php?id=12998) AHG9: On NNPF input picture selection [M. M. Hannuksela, F. Cricri (Nokia)]

[JVET-AE0051](https://jvet-experts.org/doc_end_user/current_document.php?id=12999) AHG9: On persistent NNPF activation [M. M. Hannuksela, F. Cricri (Nokia)]

[JVET-AE0052](https://jvet-experts.org/doc_end_user/current_document.php?id=13000) AHG9: NNPF cascades and alternatives [M. M. Hannuksela, F. Cricri (Nokia)]

[JVET-AE0053](https://jvet-experts.org/doc_end_user/current_document.php?id=13001) AHG8/AHG9: Neural-network post-filter regions SEI message [T. Chujoh, Y. Yasugi, T. Ikai (Sharp)]

[JVET-AE0060](https://jvet-experts.org/doc_end_user/current_document.php?id=13008) [AHG9] Comments on NNPFC [S. Deshpande (Sharp)]

[JVET-AE0061](https://jvet-experts.org/doc_end_user/current_document.php?id=13009) [AHG9] On NNPFC Application Purpose [S. Deshpande (Sharp)]

[JVET-AE0062](https://jvet-experts.org/doc_end_user/current_document.php?id=13010) [AHG9] On NNPF for Deinterlacing [A. Sidiya, S. Deshpande (Sharp)]

[JVET-AE0063](https://jvet-experts.org/doc_end_user/current_document.php?id=13011) [AHG9] On operations for multiple NNPFs [L. Chen, O. Chubach, Y.-W. Huang, S. Lei (MediaTek)]

[JVET-AE0068](https://jvet-experts.org/doc_end_user/current_document.php?id=13016) AHG9: On extensibility of purpose syntax element in NNPFC SEI message [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

[JVET-AE0069](https://jvet-experts.org/doc_end_user/current_document.php?id=13017) AHG9: On the signalling of output pictures in NNPFA SEI message [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

[JVET-AE0070](https://jvet-experts.org/doc_end_user/current_document.php?id=13018) AHG9: On input pictures that are not present in the bitstream for NNPF SEI messages [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

[JVET-AE0099](https://jvet-experts.org/doc_end_user/current_document.php?id=13047) [AHG8] NNPF and post-filter hint SEI messages for the technical report [C. Hollmann, M. Pettersson, R. Sjöberg, M. Damghanian (Ericsson)]

(JVET-AD0099 also relates to the post-filter hint SEI message)

[JVET-AE0101](https://jvet-experts.org/doc_end_user/current_document.php?id=13049) [AHG2][AHG9] Neural network post filter and phase indication SEI messages for AVC and HEVC [T. Ikai, T. Chujoh (Sharp), Y.-K. Wang, J. Xu, W. Jia (Bytedance)]

(JVET-AD0101 also relates to the phase indication SEI message)

[JVET-AE0106](https://jvet-experts.org/doc_end_user/current_document.php?id=13054) AHG9: On missing value ranges for some syntax elements in the NNPFC SEI message [C. Lin, Y.-K. Wang, J. Xu, W. Jia, J. Li, Y. Li, K. Zhang, L. Zhang (Bytedance)]

[JVET-AE0113](https://jvet-experts.org/doc_end_user/current_document.php?id=13061) AHG9: Extendibility and code word length of nnpfc\_purpose [R. Sjöberg, M. Pettersson (Ericsson)]

[JVET-AE0126](https://jvet-experts.org/doc_end_user/current_document.php?id=13081) AHG9: NNPF cleanup and editorial changes for VSEI [Y.-K. Wang, W. Jia, J. Xu, C. Lin (Bytedance)]

[JVET-AE0127](https://jvet-experts.org/doc_end_user/current_document.php?id=13083) AHG9: NNPF editorial changes for VVC [Y.-K. Wang, W. Jia, J. Xu (Bytedance)]

[JVET-AE0128](https://jvet-experts.org/doc_end_user/current_document.php?id=13084) AHG9: On NNPFC extensibility and base filter signalling [Y.-K. Wang (Bytedance)]

[JVET-AE0134](https://jvet-experts.org/doc_end_user/current_document.php?id=13097) AHG9: Align the design of NNPF with multiple input pictures to NNPF including picture rate upsampling [J. Xu, Y.-K. Wang (Bytedance)]

[JVET-AE0135](https://jvet-experts.org/doc_end_user/current_document.php?id=13098) AHG9: On NNPF picture rate upsampling constraints [J. Xu, Y.-K. Wang (Bytedance)]

[JVET-AE0141](https://jvet-experts.org/doc_end_user/current_document.php?id=13104) AHG9: Fix a bug in NNPFC SEI message for colourization [J. Xu, Y.-K. Wang (Bytedance)]

[JVET-AE0142](https://jvet-experts.org/doc_end_user/current_document.php?id=13105) AHG9: On derivation of NNPF input pictures and the value of nnpfc\_purpose [W. Jia, Y.-K. Wang, J. Xu, L. Zhang (Bytedance)]

[JVET-AE0173](https://jvet-experts.org/doc_end_user/current_document.php?id=13136) [AHG9] Clarification and improvements of signaling of NNPF update [Y. Lim (Samsung)]

[JVET-AE0175](https://jvet-experts.org/doc_end_user/current_document.php?id=13138) [AHG9] Editorial improvements of nnpfc\_mode\_idc [Y. Lim (Samsung)]

[JVET-AE0189](file:///C:\Users\smcca\Box\Sean%20McCarthy\MPEG\MPEG%20143\AHG9\current_document.php%3fid=13152) AHG9: On the design of nnpfa\_target\_base\_flag in NNPFA SEI message [Hendry (LGE)]

**Annotated regions SEI message**

[JVET-AE0054](https://jvet-experts.org/doc_end_user/current_document.php?id=13002) AHG2/AHG9: Editorial improvements of annotated regions SEI message [T. Chujoh, T. Aono, T. Ikai (Sharp)]

**Post filter hint SEI message**

[JVET-AE0099](https://jvet-experts.org/doc_end_user/current_document.php?id=13047) [AHG8] NNPF and post-filter hint SEI messages for the technical report [C. Hollmann, M. Pettersson, R. Sjöberg, M. Damghanian (Ericsson)]

(JVET-AD0099 also relates to NNPF SEI messages)

[JVET-AE0155](https://jvet-experts.org/doc_end_user/current_document.php?id=13118) Editor commentary on the post-filter hint SEI message semantics [G. J. Sullivan, Y.-K. Wang (Editors)

**Phase indication SEI message**

[JVET-AE0101](https://jvet-experts.org/doc_end_user/current_document.php?id=13049) [AHG2][AHG9] Neural network post filter and phase indication SEI messages for AVC and HEVC [T. Ikai, T. Chujoh (Sharp), Y.-K. Wang, J. Xu, W. Jia (Bytedance)]

(JVET-AD0101 also relates to NNPF SEI messages)

**SEI processing order SEI message**

[JVET-AE0156](https://jvet-experts.org/doc_end_user/current_document.php?id=13119) AHG9: Message wrapping and importance indication for the SEI processing order SEI message [G. J. Sullivan, S. McCarthy, P. Yin (Dolby Labs)]

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

[JVET-AE0143](https://jvet-experts.org/doc_end_user/current_document.php?id=13106) AHG8: A spatial resampling algorithm and an exemplar software implementation [S. Wang, B. Li, J. Chen, Y. Ye (Alibaba), S. Wang (CityU)]

***Identify potential needs for additional SEI messages***

**Human/machine viewing**

[JVET-AE0064](https://jvet-experts.org/doc_end_user/current_document.php?id=13012) AHG8/AHG9: Signalling encoder preprocessing and human / machine viewing indications [C. Kim, D. Gwak, Hendry, J. Lim, S. Kim (LGE), M. M. Hannuksela, F. Cricri, H. Zhang (Nokia)]

[JVET-AE0090](https://jvet-experts.org/doc_end_user/current_document.php?id=13038) AHG8/AHG9: On machine vision indication [J. Gao, H.-B. Teo, C.-S. Lim, K. Abe (Panasonic)]

**Multiplane image information**

[JVET-AE0066](https://jvet-experts.org/doc_end_user/current_document.php?id=13014) AHG9: Multiplane Image Information SEI [T. Lu, P. Yin, G. Su, D. Lee, T. Huang, S. McCarthy, W. Husak, G. J. Sullivan (Dolby)]

**Alternative picture timing**

[JVET-AE0079](https://jvet-experts.org/doc_end_user/current_document.php?id=13027) AHG8/AHG9: Source picture timing information SEI message [S. McCarthy, G. J. Sullivan, P. Yin (Dolby)]

[JVET-AE0089](https://jvet-experts.org/doc_end_user/current_document.php?id=13037) AHG9: Alternative Output Timing Hint SEI [H.-B. Teo, J. Gao, C.-S. Lim, K. Abe (Panasonic)]

**Generative face video**

[JVET-AE0080](https://jvet-experts.org/doc_end_user/current_document.php?id=13028) AHG9: Generative Face Video SEI message [S. McCarthy, P. Yin, G.-M. Su, A. K. Choudhury, W. Husak (Dolby)]

[JVET-AE0083](https://jvet-experts.org/doc_end_user/current_document.php?id=13031) AHG9: Common SEI Message of Generative Face Video [B. Chen, J. Chen, Y. Ye (Alibaba), S. Wang (CityU)]

[JVET-AE0088](https://jvet-experts.org/doc_end_user/current_document.php?id=13036) AHG9: A study on Generative Face Video SEI Message [H.-B. Teo, J.-Y Thong, K. Jayashree, C.-S. Lim, K. Abe (Panasonic)]

**Object mask information**

[JVET-AE0095](https://jvet-experts.org/doc_end_user/current_document.php?id=13043) AHG8/AHG9: proposed changes to the candidate new object mask information SEI message [P. de Lagrange, E. François, D. Doyen (InterDigital), J. Chen, S. Wang, Y. Ye (Alibaba)]

**Object-wave compression and computer-generated hologram use**

[JVET-AE0182](https://jvet-experts.org/doc_end_user/current_document.php?id=13145) AHG9: SEI message extension of VVC for object-wave compression and computer-generated hologram use [K. Nonaka, R. Koiso, H. Kojima, K. Kawamura, H. Kato (KDDI)]

***Coordinate with AHG3 for software support of SEI messages***

Bug report <https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1605>

**Activities**

The regular JVET e-mail reflector was used for discussions ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de)) with [AHG9] in message headers. There were no emails sent to the JVET reflector during the AHG period with [AHG9] in the message header.

**Recommendations**

The AHG recommends to:

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

It was agreed that no technical aspects regarding new functionality should be added to the v3 NNPF SEI message; main focus on bug fixes. In terms of purpose, no new items should be added to v3, as long as extensibility is possible.

BoG on VSEI V3 aspects, and VVC V3 SEI message aspects (G. Sullivan), to meet from Tuesday afternoon. The BoG should also sort out which proposals in the NNPF category would rather be candidates for a future version.

[JVET-AE0010](https://jvet-experts.org/doc_end_user/current_document.php?id=13085) JVET AHG report: Encoding algorithm optimization (AHG10) [P. de Lagrange, A. Duenas, R. Sjöberg, A. Tourapis (AHG chairs)]

**Related contributions**

A total of 5 contributions, not including cross-checks, are identified relating to AHG10, and summarized in the following sections.

***Adaptive resolution***

**JVET-AE0104 - AHG10: GOP-based RPR encoder control using parallel resolution encoding**

This contribution proposes multi-pass GOP-based adaptive resolution across full-resolution and 4/5, 2/3, and ½ reduced resolutions, using a rate-distortion cost based on squared error (=PSNR). The reported BD-rate difference in RA class A (using CTC QPs) are about -1.5% when compared to regular full-resolutions, and around -1% when compared to the current single-pass GOP-based RPR technique implemented in the VTM. When adding a lower quality point on top of the CTC QPs (QP=42 for UHD), gain increases compared to regular coding, but decreases compared to current GOP-based RPR. Gains are generally lower for lower resolutions and for LD and AI configurations.

***Low delay***

**JVET-AE0114 - AHG10: Low-Delay configuration improvements**

This contribution proposes changes to the low-delay configurations regarding pictures and QPs for improved coding efficiency. In addition to a -3 QP offset to the first intra picture (instead of -1), it is proposed to introduce a multi-level QP structure with a periodicity of 16 pictures, plus a picture with a -2 offset every 64 pictures. Reference picture list is changed to take advantage of higher quality reference pictures. A BD-rate difference of around -6% is reported.

***Bug fixes, more encoder control***

**JVET-AE0111 - Additional conformance tests of spatial scalability for multilayer coding**

Besides proposing an additional conformance bitstream, this contribution reports about fixing a bug in reference picture list when combining scalability and low-delay configuration: first picture of enhancement layer is intra-coded (not using the base layer as a reference), and for the remaining pictures, intra-layer prediction is removed when inter-layer referencing is activated; the proposed fix is in MR !2633.

**JVET-AE0122 - AHG10: GOP-based RPR and scalable coding**

This contribution proposes a bug fix and more control over the encoder when combining RPR and multi-layer. One use case is to restrict reduced resolution coding to the base layer. A marginal BD-rate gain is reported when testing the proposed changes using 1.5x scalability and GOP-based RPR for the base layer, over the test sequences currently considered for scalability visual testing.

**JVET-AE0181 - Scaling window support for VTM**

This contribution reports about adding control of the scaling window to the VTM, with additional configuration parameters. It also reports about an experiment using dual-layer coding with the enhancement layer containing a scaled-down, windowed version of the base layer. When coding without scaling window, affine coding mode is mostly used, while merge mode is used when using scaling window. A lower rate with a higher PSNR is obtained.

**Recommendation**

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

[JVET-AE0011](https://jvet-experts.org/doc_end_user/current_document.php?id=13086) JVET AHG report: Neural network-based video coding (AHG11) [E. Alshina, S. Liu, A. Segall (co chairs), F. Galpin, J. Li, R.-L. Liao, D. Rusanovskyy, T. Shao, M. Wien, P. Wu (vice-chairs)]

**Activities**

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

***Common Test Conditions***

**Document**

The AHG released revised common test conditions as decided at the 30th meeting. The final version was uploaded as document JVET-AD2016 on May 12, 2023.

**Anchor Encoding**

Anchors for the NN-based video coding activity made available on the Git repository used for the AHG activity: <https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/nnvc-ctc/-/tree/master>.

***EE Coordination***

The AHG finalized, conducted and discussed the EE on NN based video coding. The final version of the EE description was uploaded to the document repository on May 15, 2023.

A summary report for the EE is available at this meeting as:

|  |  |  |
| --- | --- | --- |
| JVET-AE0023 | EE1: Summary report of exploration experiment on neural network-based video coding | E. Alshina, F. Galpin, Y. Li, D. Rusanovskyy, M. Santamaria, J. Ström, L. Wang, Z. Xie (EE coordinators) |

***Teleconferences***

The AHG conducted two teleconferences during the interim period. The teleconferences were held on June 14, 2023, and June 28, 2023 and attended by approximately 20 participants.

A summary report for the teleconferences is available at this meeting as:

|  |  |  |
| --- | --- | --- |
| JVET-AE0042 | AhG14 & AHG11: Report on AhG teleconference on high operation point (HOP) unified filter training | E. Alshina, F. Galpin |

***Workshop on Neural Network-Based Technologies***

The AHG contributed to the “Third AG4 Workshop on JPEG and MPEG Emerging Activities” that was conducted by the JPEG and MPEG Collaboration Advisory Group, or ISO/IEC JTC1/SC29/AG4.

The workshop took place on June 1st, 2023, and its agenda was:

|  |  |
| --- | --- |
| **Speaker** | **Topic** |
| Elena Alshina | AI-based Video Coding |
| Werner Bailer | Neural Network Compression |
| Marc Antonini | Coding for DNA Storage |
| Tim Bruylants | Event-based Vision |
| Arianne Hids | Scene-based Interchange and Support of Immersive Displays |

The AHG contributed the presentation titled “AI-based Video Coding”. Slides for the presentation are available in JVET-AE0237.

***Performance Evaluation***

The performance of the NNVC-5.0 anchor compared to NNVC-4.0 is reported below.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |  |
|  | **BD-rate Over NNVC-4.0** | | | | | | | |  |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU | PSNR Overlap |
| Class A1 | -7.23% | -4.86% | -6.00% | -8.72% | -4.05% | -5.06% | 134% | 7667% | 96% |
| Class A2 | -6.56% | -5.78% | -5.09% | -6.53% | -4.14% | -2.90% | 130% | 7183% | 97% |
| Class B | -6.29% | -8.19% | -7.53% | -5.97% | -6.06% | -5.18% | 132% | 7834% | 98% |
| Class C | -6.62% | -10.53% | -9.55% | -6.26% | -7.91% | -5.58% | 132% | 7422% | 98% |
| Class E | - | - | - | - | - | - | - | - | - |
| **Overall** | -6.62% | -7.66% | -7.28% | -6.71% | -5.77% | -4.81% | 132% | 7557% | 97% |
| Class D | -7.57% | -9.45% | -9.68% | -5.76% | -6.81% | -5.94% | 129% | 8193% | 98% |
| Class F | -3.07% | -5.88% | -4.94% | -3.05% | -4.76% | -3.54% | 148% | 9739% | 99% |
| Class H | - | - | - | - | - | - | - | - | - |
|  |  |  |  |  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | | | | |  |
|  | **BD-rate Over NNVC-4.0** | | | | | | | |  |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU | PSNR Overlap |
| Class A1 | - | - | - | - | - | - | - | - | - |
| Class A2 | - | - | - | - | - | - | - | - | - |
| Class B | -4.85% | -8.13% | -7.84% | -5.15% | -6.16% | -6.65% | 134% | 6812% | 99% |
| Class C | -5.24% | -11.13% | -9.00% | -5.61% | -10.74% | -4.02% | 126% | 6233% | 99% |
| Class E | -4.88% | -4.09% | -5.17% | -5.64% | -2.18% | -3.36% | 162% | 9877% | 98% |
| **Overall** | -4.98% | -8.12% | -7.56% | -5.42% | -6.69% | -4.95% | 138% | 7257% | 99% |
| Class D | -6.54% | -10.76% | -9.77% | -5.75% | -9.48% | -6.83% | 125% | 6400% | 98% |
| Class F | -2.05% | -4.88% | -4.46% | -2.29% | -3.08% | -4.12% | 148% | 8941% | 99% |
| Class H | - | - | - | - | - | - | - | - | - |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  | **All Intra Main10** | | | | | | | |  |
|  | **BD-rate Over NNVC-4.0** | | | | | | | |  |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU | PSNR Overlap |
| Class A1 | -8.46% | -7.23% | -8.67% | -10.00% | -8.00% | -7.50% | 210% | 5134% | 97% |
| Class A2 | -6.61% | -8.04% | -7.94% | -6.90% | -6.44% | -5.44% | 204% | 4416% | 98% |
| Class B | -7.03% | -8.58% | -8.71% | -7.04% | -7.61% | -6.99% | 203% | 4493% | 98% |
| Class C | -7.34% | -9.76% | -9.92% | -7.61% | -8.38% | -7.64% | 190% | 3770% | 98% |
| Class E | -10.29% | -9.04% | -8.85% | -10.14% | -7.52% | -6.38% | 199% | 5149% | 96% |
| **Overall** | -7.81% | -8.60% | -8.87% | -8.16% | -7.64% | -6.86% | 201% | 4507% | 97% |
| Class D | -7.57% | -8.11% | -9.33% | -7.16% | -6.81% | -7.56% | 180% | 3760% | 98% |
| Class F | -4.55% | -5.97% | -5.65% | -4.24% | -4.75% | -4.38% | 155% | 3961% | 98% |
| Class H | - | - | - | - | - | - | - | - | - |

**Input contributions**

There are 47 input contributions related to the AHG mandates. Ten of the contributions are part of the EE activity, while the remaining 37 contributions are related to AHG11 but not part of the EE. The list of input contributions is provided below.

***EE and Related Input Contributions***

|  |  |  |
| --- | --- | --- |
| **Reporting** | | |
| [JVET-AE0023](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13091) | EE1: Summary report of exploration experiment on neural network-based video coding | E. Alshina, F. Galpin, Y. Li, D. Rusanovskyy, M. Santamaria, J. Ström, L. Wang, Z. Xie |
| **EE Technology** | | |
| [JVET-AE0067](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13015) | EE1-4.1: Neural-network loop filters with further complexity reduction | J. N. Shingala, A. Shyam, A. Suneja, S. Badya (Ittiam), T. Shao, A. Arora, P. Yin, F. Pu, T. Lu, Sean McCarthy (Dolby) |
| [JVET-AE0112](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13060) | EE1-5.1: Deep Reference Frame Generation for Inter Prediction Enhancement | W. Bao, W. Meng, J. Jia, Y. Zhang, H. Wang, Z. Chen (Wuhan Univ.), [Z. Liu](mailto:zizhengliu@tencent.com), X. Xu, S. Liu (Tencent) |
| [JVET-AE0144](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13107) | EE1-6.1: neural network-based intra prediction with reduced complexity | T. Dumas, F. Galpin, P. Bordes (Interdigital) |
| [JVET-AE0160](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13123) | EE1-1.5: Optimization for complexity-performance trade-off of HOP network | R. Chang, L. Wang, X. Xu, S. Liu (Tencent) |
| [JVET-AE0164](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13127) | EE1-1.2 Complexity-performance tradeoff of decomposition | D. Rusanovskyy, Y. Li, M. Karczewicz (Qualcomm) |
| [JVET-AE0165](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13128) | EE1-4.4: Low complexity NN filter with design elements of Unified Filter Architecture and EE1-1.2 and EE1-1.3 | Y. Li, D. Rusanovskyy, M. Karczewicz (Qualcomm) |
| [JVET-AE0191](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13154) | AhG11: EE1-0 High Operation Point model | F. Galpin (InterDigital), S. Eadie, D. Rusanovskyy (Qualcomm), Y. Li, J. Li (ByteDance), L. Wang, R. Chang (Tencent), Z. Xie (Oppo), E. Alshina (Huawei) |
| **Cross Checks** | | |
| [JVET-AE0183](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13146) | Crosscheck of JVET-AE0067 (EE1-4.1: Neural-network loop filters with further complexity reduction) | J. Ström (Ericsson) |
| JVET-AE0229 | Crosscheck of JVET-AE0112(EE1-5.1: Deep Reference Frame Generation for Inter Prediction Enhancement) | X.Jie (OPPO) |

***Non-EE Input Contributions***

|  |  |  |
| --- | --- | --- |
| **Reporting** | | |
| JVET-AE0011 | JVET AHG report: Neural network-based video coding (AHG11) | E. Alshina, S. Liu, A. Segall (co chairs), F. Galpin, J. Li, R.-L. Liao, D. Rusanovskyy, T. Shao, M. Wien, P. Wu (vice chairs) |
| [JVET-AE0042](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=12990) | AhG14 & AHG11: Report on AhG teleconference on high operation point (HOP) unified filter training | E. Alshina, F. Galpin |
| JVET-AE0237 | Presentation of AI-based video Coding in "Third AG4 Workshop on JPEG and MPEG Emerging Activities" | E. Alshina |
| **Loop Filtering** | | |
| [JVET-AE0072](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13020) | [AHG11] On residual adjustments for NNLF | Z. Dai, Y. Yu, H. Yu, D. Wang (OPPO) |
| [JVET-AE0093](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13041) | AHG11: Content-adaptive neural network loop-filter | R. Yang, M. Santamaria, F. Cricri, H. Zhang, J. Lainema, R. G. Youvalari, M. M. Hannuksela (Nokia) |
| [JVET-AE0161](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13124) | AHG11: Input and output rotation of model for NNVC in-loop filter | R. Chang, L. Wang, X. Xu, S. Liu (Tencent) |
| [JVET-AE0171](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13134) | AHG11: Neural network-based in-loop filter with layer normalization | Y. Li, K. Zhang, L. Zhang (Bytedance) |
| **Post Filtering** | | |
| [JVET-AE0048](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=12996) | AHG9: Miscellaneous VSEI changes on neural-network post-filter SEI messages | M. M. Hannuksela, F. Cricri, M. Santamaria (Nokia) |
| [JVET-AE0049](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=12997) | AHG9: Miscellaneous VVC changes on neural-network post-filter SEI messages | M. M. Hannuksela, F. Cricri, M. Santamaria (Nokia) |
| [JVET-AE0050](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=12998) | AHG9: On NNPF input picture selection | M. M. Hannuksela, F. Cricri (Nokia) |
| [JVET-AE0051](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=12999) | AHG9: On persistent NNPF activation | M. M. Hannuksela, F. Cricri (Nokia) |
| [JVET-AE0052](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13000) | AHG9: NNPF cascades and alternatives | M. M. Hannuksela, F. Cricri (Nokia) |
| [JVET-AE0053](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13001) | AHG8/AHG9: Neural-network post-filter regions SEI message | T. Chujoh, Y. Yasugi, T. Ikai (Sharp) |
| [JVET-AE0060](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13008) | [AHG9] Comments on NNPFC | S. Deshpande (Sharp) |
| [JVET-AE0061](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13009) | [AHG9] On NNPFC Application Purpose | S. Deshpande (Sharp) |
| [JVET-AE0062](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13010) | [AHG9] On NNPF for Deinterlacing | A. Sidiya, S. Deshpande (Sharp) |
| [JVET-AE0063](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13011) | [AHG9] On grouping and operations for multiple NNPFs | L. Chen, O. Chubach, Y.-W. Huang, S. Lei (MediaTek) |
| [JVET-AE0068](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13016) | AHG9: On extensibility of purpose syntax element in NNPFC SEI message | Hendry, J. Nam, S. Kim, J. Lim (LGE) |
| [JVET-AE0069](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13017) | AHG9: On the signalling of output pictures in NNPFA SEI message | Hendry, J. Nam, S. Kim, J. Lim (LGE) |
| [JVET-AE0070](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13018) | AHG9: On input pictures that are not present in the bitstream for NNPF SEI messages | Hendry, J. Nam, S. Kim, J. Lim (LGE) |
| [JVET-AE0101](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13049) | [AHG2][AHG9] Neural network post filter and phase indication SEI messages for AVC and HEVC | T. Ikai, T. Chujoh (Sharp), Y.-K. Wang, J. Xu, W. Jia (Bytedance) |
| [JVET-AE0106](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13054) | AHG9: On missing value ranges for some syntax elements in the NNPFC SEI message | C. Lin, Y.-K. Wang, J. Xu, W. Jia, J. Li, Y. Li, K. Zhang, L. Zhang (Bytedance) |
| [JVET-AE0113](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13061) | AHG9: Extendibility and code word length of nnpfc\_purpose | R. Sjöberg, M. Pettersson (Ericsson) |
| [JVET-AE0126](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13081) | AHG9: NNPF cleanup and editorial changes for VSEI | Y.-K. Wang, W. Jia, J. Xu, C. Lin (Bytedance) |
| [JVET-AE0127](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13083) | AHG9: NNPF editorial changes for VVC | Y.-K. Wang, W. Jia, J. Xu (Bytedance) |
| [JVET-AE0128](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13084) | AHG9: On NNPFC extensibility and base filter signalling | Y.-K. Wang (Bytedance) |
| [JVET-AE0134](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13097) | AHG9: Align the design of NNPF with multiple input pictures to NNPF including picture rate upsampling | J. Xu, Y.-K. Wang (Bytedance) |
| [JVET-AE0135](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13098) | AHG9: On NNPF picture rate upsampling constraints | J. Xu, Y.-K. Wang (Bytedance) |
| [JVET-AE0141](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13104) | AHG9: Fix a bug in NNPFC SEI message for colourization | J. Xu, Y.-K. Wang (Bytedance) |
| [JVET-AE0142](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13105) | AHG9: On derivation of NNPF input pictures and the value of nnpfc\_purpose | W. Jia, Y.-K. Wang, J. Xu, L. Zhang (Bytedance) |
| [JVET-AE0155](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13118) | AHG2/AHG9: Editor commentary on the post-filter hint SEI message semantics | G. J. Sullivan, Y.-K. Wang (Editors) |
| [JVET-AE0173](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13136) | [AHG9] Clarification and improvements of signalling of NNPF update | Y. Lim (Samsung) |
| [JVET-AE0175](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13138) | [AHG9] Editorial improvements of nnpfc\_mode\_idc | Y. Lim (Samsung) |
| [JVET-AE0187](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13150) | AHG9: A summary of SEI proposals on NNPF | Y.-K. Wang (Bytedance) |
| [JVET-AE0189](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13152) | AHG9: On the design of nnpfa\_target\_base\_flag in NNPFA SEI message | Hendry (LGE) |
| **Test Conditions** | | |
| [JVET-AE0162](file:////Users/shanl/Documents/contribution/jvet31ae/current_document.php%3fid=13125) | AHG11/AHG14: Fix MS-SSIM calculation for SR | R. Chang, L. Wang, X. Xu, S. Liu (Tencent) |
| **Cross Checks** | | |
| JVET-AE0230 | Crosscheck of JVET-AE0162(AHG11/AHG14 : Fix MS-SSIM calculation for SR) | Z. Xie (OPPO) |
| JVET-AE0232 | Crosscheck of JVET-AE0072 ([AHG11] On residual adjustments for NNLF) | T. Shao (OPPO) |

**Recommendations**

The AHG recommends:

* Review all input contributions.
* Continue investigating neural network-based video coding tools, including coding performance and complexity.
* AHG11 encourages EE1 to continue using a workflow similar to the one in JVET-AE0191 with unified training, clearly documented NNVC technology and transparent training process performed by multiple parties in parallel with regular information sharing.

[JVET-AE0012](https://jvet-experts.org/doc_end_user/current_document.php?id=13087) 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, H. Yang (vice-chairs)]

**Activities**

The Common Test Conditions were updated (JVET-AD2017). The run time precision in the reporting template has been increased to 1 digit after decimal point. The primary activity of the AHG was the “Exploration experiment on enhanced compression beyond VVC capability” (JVET-ACD024). The combined improvements of the ECM-9.0 over VTM-11.0ecm8.0 anchorfor AI, RA and LB configurations are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | All Intra Main10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -9.67% | -21.16% | -28.19% | 827.7% | 406.4% |
| Class A2 | -15.77% | -27.44% | -30.55% | 813.4% | 405.9% |
| Class B | -10.06% | -25.61% | -23.82% | 773.3% | 423.4% |
| Class C | -10.27% | -15.86% | -16.33% | 783.1% | 442.4% |
| Class E | -13.65% | -23.55% | -22.06% | 759.1% | 481.4% |
| Overall | -11.59% | -22.66% | -23.71% | 788.5% | 430.8% |
| Class D | -8.61% | -13.65% | -13.74% | 803.2% | 462.1% |
| Class F | -24.34% | -33.43% | -33.28% | 529.7% | 445.8% |
| Class TGM | -37.29% | -45.21% | -44.55% | 443.8% | 443.1% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Random Access Main 10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -20.86% | -24.25% | -32.21% | 757.7% | 741.6% |
| Class A2 | -23.90% | -32.20% | -35.28% | 697.8% | 882.0% |
| Class B | -19.50% | -31.71% | -29.62% | 648.2% | 739.3% |
| Class C | -20.90% | -24.61% | -24.58% | 672.8% | 779.1% |
| Class E |  |  |  |  |  |
| Overall | -21.03% | -28.42% | -29.93% | 685.5% | 777.1% |
| Class D | -21.87% | -25.90% | -26.55% | 699.1% | 849.6% |
| Class F | -26.77% | -34.78% | -35.14% | 578.7% | 528.1% |
| Class TGM | -35.17% | -41.37% | -41.35% | 547.2% | 419.4% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Low Delay B Main 10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -17.01% | -32.11% | -30.48% | 601.4% | 658.9% |
| Class C | -18.50% | -26.42% | -26.63% | 590.5% | 661.3% |
| Class E | -15.40% | -25.47% | -26.14% | 595.7% | 440.1% |
| Overall | -17.10% | -28.56% | -28.11% | 596.3% | 596.4% |
| Class D | -20.39% | -28.41% | -27.33% | 629.6% | 770.9% |
| Class F | -23.36% | -33.28% | -31.66% | 562.9% | 526.4% |
| Class TGM | -33.93% | -40.53% | -40.11% | 511.1% | 404.8% |

The rate reduction for natural sequences over VTM in RA configuration for {Y, U, V} increased from ECM-8.0’s {-19.86%, -26.50%, -26.67%} to ECM-9.0’s {-21.03%, -28.42%, -29.93%} with the encoder complexity increasing from 601% to 686%. For SCC sequences (class TGM) the rate reduction for RA configuration increased from ECM-8.0’s {-33.27%, -38.66%, -38.66%} to ECM-9.0’s {-35.17%, -41.66%, -41.35%}.

**Contributions**

In addition to 24 EE2 contributions, 52 (comparing to 55 last meeting) EE2-related and ECM-related contributions were received and can be subdivided as follows:

***Intra (19)***

JVET-AE0071, "AHG12: DIMD Merge", S. Blasi, I. Zupancic, J. Lainema (Nokia)

JVET-AE0073, "Non-EE2: IBC-LIC Model Merge mode", L. Zhang, Y. Yu, H. Yu, D. Wang (OPPO)

JVET-AE0075, "Non-EE2: On IntraTMP block vector", Y. Yu, L. Zhang, L. Xu, H. Yu, J. Gan, F. Wang, Z. Xie, D. Wang (OPPO)

JVET-AE0085, "Non-EE2: Direct block vector (DBV) mode extension", M. Hong, J. Choi, N. Park, J. Lim, S. Kim (LGE)

JVET-AE0098, "AHG12: On the chroma DBV mode", H. Huang, H. Wang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AE0110, "Non-EE2: An improved method for IntraTMP fusion", S. Peng, D. Jiang, J. Lin, C. Zheng, K. Fu, P. Zhang (Dahua)

JVET-AE0120, "Non-EE2: Intra Angular Prediction Extension", J. Fu, J. Zhang, C. Jia, S. Ma (PKU)

JVET-AE0124, "Non-EE2: Fixes related to Intra TMP", J.-L Lin, P.-H Lin, Y.-J Chang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AE0130, "Non-EE2: DIMD with filtered template", C. Zhou, Z. Lv (vivo)

JVET-AE0137, "Non-EE2: On TIMD fusion", P. Andrivon", M. Blestel (Ofinno)

JVET-AE0138, "Non-EE2: Reference sample interpolation for intra prediction", G. Rath, F. Le Léannec, F. Urban, F. Racape, T. Dumas (InterDigital)

JVET-AE0158, "AHG12: LUT-based angle calculation for DIMD", A. Aminlou, J. Lainema (Nokia)

JVET-AE0166, "AHG12: IntraTMP and IBC search area harmonization", D. Ruiz Coll, B. Chen, P. Andrivon, M. Blestel (Ofinno)

JVET-AE0167, "AHG12: DBV improvement", L. Xu, Y. Yu, H. Yu, D. Wang (OPPO)

JVET-AE0184, "AHG12: TIMD with IntraTMP and IBC candidates", K. Naser, F. Le Léannec, P. Bordes, F. Galpin, A. Robert (InterDigital)

JVET-AE0185, "AHG12: IntraTMP with Neighboring Candidates", K. Naser, F. Le Léannec, A. Robert, F. Galpin (InterDigital)

JVET-AE0186, "Non-EE2-2.1: Block Vector Guided Chroma Direct Mode", K. Naser, F. Le Léannec, P. Bordes, Y. Chen, G. Rath (InterDigital)

JVET-AE0198, "AHG12: Extended Search Region for IntraTMP", K. Naser, P. Bordes, F. Le Léannec, F. Galpin (InterDigital)

JVET-AE0214, "AHG12: on Intra CIIP", K. Naser, P. Bordes, F. Galpin, K. Reuze (InterDigital)

***Inter (11)***

JVET-AE0047, "AHG12: On GPM-MMVD", R. Yu (Ericsson)

JVET-AE0105, "Non-EE2: Local illumination compensation with multiple templates", Y. Wang, K. Zhang, Y. He, H. Liu, L. Zhang (Bytedance)

JVET-AE0108, "AHG12: DMVR with robust MV derivation", K. Andersson, R. Yu (Ericsson)

JVET-AE0109, "Non-EE2: LIC flag derivation of merge candidates with template costs", N. Zhang, K. Zhang, H. Liu, Y. Wang, L. Zhang (Bytedance)

JVET-AE0117, "Non-EE2: Enhanced subblock-based motion compensation", L. Zhao, K. Zhang, L. Zhang (Bytedance)

JVET-AE0129, "AHG12 Template-based CIIP weight derivation", M.-H. Jia, Y.-L. Hu, S.-W. Xie, Y. Gao, C. Huang (ZTE)

JVET-AE0136, "Non-EE2: Fix on TM-based reordering for affine MMVD mode", D. Kim, W. Lim, J. Kim, S.-C. Lim, J. S. Choi (ETRI)

JVET-AE0140, "Non-EE2: LIC extensions", A. Filippov, V. Rufitskiy, K. Suverov (Ofinno)

JVET-AE0148, "Non-EE2: Affine subblock BDOF refinement", Z. Zhang, H. Huang, J.-L Lin, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AE0168, "Non-EE2: Sample-based BDOF for Chroma", C.-C. Chen, H. Huang, Z. Zhang, C. S. Coban, V. Seregin, M. Karczewicz

JVET-AE0177, "Non-EE2: Local illumination compensation with slope adjustment", Y. Wang, K. Zhang, Y. He, H. Liu, L. Zhang (Bytedance)

***Cross Component Prediction (6)***

JVET-AE0074, "Non-EE2: improvement on cross-component prediction merge mode", H. Huang, Y. Yu, H. Yu, D. Wang (OPPO)

JVET-AE0097, "AHG12: On the cross-component merge mode", H. Huang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AE0115, "Non-EE2: Unified intra CC-models parameters precision", P. Bordes, K. Naser, F. Galpin, F. Leleannec (InterDigital)

JVET-AE0170, "Non-EE2: Enhancements on CCP merge", K. Zhang, Z. Deng, L. Zhang (Bytedance)

JVET-AE0176, "EE2-related: Enhancements on CCRM", Z. Deng, K. Zhang, L. Zhang (Bytedance)

JVET-AE0178, "Non-EE2: Cross-component prediction merge mode for chroma inter coding", M.-S. Chiang, H.-Y. Tseng, C.-M. Tsai, C.-Y. Chuang, C.-W. Hsu, C.-Y. Chen, T.-D. Chuang, O. Chubach, Y.-W. Chen, Y.-W. Huang, S.-M. Lei (MediaTek)

***In-Loop Filters (3)***

JVET-AE0044, "AHG12: Dynamic Scaling of Bilateral Filter (BIF)", V. Shchukin, P. Wennersten, J. Ström (Ericsson)

JVET-AE0121, "Non-EE2: Luma Residual Taps in Chroma-ALF and CCALF", W. Yin, K. Zhang, Z. Deng, L. Zhang (Bytedance)

JVET-AE0131, "Non-EE2: Variance based Classification for In-loop Filtering", W. Yin, K. Zhang, L. Zhang (Bytedance)

***Entropy Coding (1)***

JVET-AE0058, "AHG12: Spatial CABAC tuning", J. Lainema, A. Aminlou, P. Astola", R. G. Youvalari, D. Bugdayci Sansli (Nokia)

***Partitioning (1)***

JVET-AE0082, "AHG12: Non-square quadtree partitioning", Y. Ahn, J. Nam, N. Park, J. Lim, S. Kim (LGE)

***GDR (2)***

JVET-AE0145, "AHG 12: Flexible GDR", L. Wang, S. Hong, K. Panusopone (Nokia)

JVET-AE0146, "AHG 12: Fixes H-CCP Table for CUs in Refreshed Areas of GDR/Recovering Pictures", S. Hong, L. Wang, K. Panusopone (Nokia)

***RPR (2)***

JVET-AE0103, "Weighted Edge Enhancement Filtering for Picture Upscaling and RPR", T. ClaÃŸen, M. Wien (RWTH Aachen University)

JVET-AE0153, "Non-EE2: Enabling template-based inter tools for scaled reference pictures in the RPR", X. Xiu, H.-J. Jhu, C.-W. Kuo, C. Ma, N. Yan, W. Chen, X. Wang (Kwai)

***Encoder/CTC (7)***

JVET-AE0055, "AHG12: Sign prediction parameter configuration for low-delay conditions", Y. Yasugi, T. Ikai (Sharp)

JVET-AE0056, "Non-EE2: Changes on TIMD and SGPM for reducing ECM decoder complexity", Z. Fan, Y. Yasugi, T. Ikai (Sharp)

JVET-AE0057, "MTT split modes early termination", W. Ahmad, P. Wennersten, K. Andersson (Ericsson)

JVET-AE0133, "AhG12: On Common Test Conditions change based on EE2 Test 1.1", G. Laroche, P. Onno (Canon)

JVET-AE0147, "AHG12: Updating context model parameters for ECM", R.-L. Liao, Y. Ye, J. Chen, X. Li (Alibaba)

JVET-AE0163, "AHG12: modified CTC proposal for low-delay configuration", S. Puri, C. Bonnineau, F. Le Leannec, T. Poirier, E. Francois (InterDigital)

JVET-AE0180, "On ECM SW memory consumption", R. Chernyak, S. Liu (Tencent), Y. Yasugi, T. Ikai (Sharp)

Recommendations

The AHG recommends to:

* To review all the related contributions.

It was pointed out that the gain gap between luma and chroma became again larger with ECM 9. It was however asserted to be careful shifting gain from chroma to luma without careful study e.g. in terms of visual impact and rate dependency.

[JVET-AE0013](https://jvet-experts.org/doc_end_user/current_document.php?id=13089) JVET AHG report: Film grain technologies (AHG13) [W. Husak, M. Radosavljević (co-chairs), A. Duenas, D. Grois, Y. He, P. de Lagrange, X. Meng, A. Segall, A. Tourapis, W. Zhang (vice-chairs)]

**Discussion**

Much of this AHG period, ISO/IEC 23002-9 CDTR was out for comment. The document was issued on May 15. The expected response date was July 9 – just before the 31st Meeting (Geneva). The editors met on a roughly biweekly cadence throughout the period.

The discussions were related to content for film grain, external activities on film grain technologies and additional software being developed for film grain. There are several companies that are independently exploring content. The content is expected to be used for demonstrations and tests. The following characteristics were enumerated for further study:

* video representative of what could be found on streaming platforms
* 10 or 15s length
* consistent grain / no scene cut if possible
  + perhaps cross-fades?
* balanced scene lighting (not everything dark or bright)
* decent resolution and video quality, so that video quality can be rated
* reasonable motion (not too fast, not too blurry, not fully still)
* reasonable spatial content (not empty sky, not packed with excessive texture)
* "normal" grain (not just blurry blobs or blocks, or just scratches)
* color / BW?

The group agreed the content should include a “ground truth”. However, it was recognized that the “de-noising” process could be controversial. Several content repositories were identified that could be a source of open-source content. In addition, some companies are exploring providing content they own and can assign the rights.

Two applications SDOs have opened or are exploring study items/missions on film grain. 3GPP has started a study item that is scheduled to be completed at the end of 2023. DVB is also exploring opening a study mission on film grain. There is also an expectation that other SDOs will follow suit and explore film grain. The group discussed potential support for those groups.

There were two industry demonstrations with film grain during the AHG period. The first was at Mile High Video 2003 (MHV2003) and the second was at DVB World. The second supported the launch and discussion of the DVB study mission.

Finally, several additional software packages were discussed. Several software packages were updated and some new features were added, such as an autoregressive parameter translator.

**Related contributions**

Three contribution related to AHG13 was identified as of 07/10/2023.

* The one contribution was the AHG report:
  + JVET-AE0013 JVET AHG report: Film grain technologies (AHG13)
* Two other contributions were added but at present without an abstract
  + JVET-AE0250 Film Grain Synthesis: Material Selection for Ground Truth Content Creation
  + JVET-AE0251 Evaluating Denoising methods in the context of Film Grain Analysis/Synthesis

***Contributions***

There were no contributions uploaded as of report time other than the AHG report.

**Recommendations**

The AHG recommends:

* the related input contributions are reviewed;
* testing of FGC be discussed; and
* continue the study of film grain technologies in JVET.

It was pointed out that the software packages mentioned could be useful in the visual assessment activity.

BoG (W. Husak) to develop the DoCR (ballot response in [m63723](https://dms.mpeg.expert/doc_end_user/current_document.php?id=88067&id_meeting=195)), and define which changes will be made in draft 5 relative to draft 4 (draft 5 will also be submitted as WG 5 DTR).

Potential viewing regarding the aspects of JVET-AE0250 and JVET-AE0251.

[JVET-AE0014](https://jvet-experts.org/doc_end_user/current_document.php?id=13090) JVET AHG report: NNVC software development (AHG14) [S. Eadie, F. Galpin, Y. Li, J. Shingala, L. Wang, Z. Xie (AHG chairs)]

**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, and enabling deblocking in the RDO.

NNVC-5.1 anchor at <https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/VVCSoftware_VTM> is used for NNVC performance evaluation (Note: NNVC-5.0 should produce the same results for anchor generation).

***Software changes***

The following changes were integrated:

* Integration of JVET-AD0107 – simplified parameter selection filterset 1
* Integration of JVET-AD0156 – Low complexity filterset
* Integration of JVET-AD0380 – High Operation Point Filterset
* Integration of JVET-AD0163 – content adaptive NN post-filter
* Integration of JVET-AD0164 – sadl v5 and v5.1
* Integration of JVET-AC0089 Use qp slice for inter slcie
* Various fixes (21 issues resolved)

Other changes:

* New configurations files for anchors
* Improvement of the data loader for training
* Improvement of training dataset preparation scripts
* Improvement of training dataset creation scripts

***Software version***

NNVC-5.1 was tagged xxx, 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).

**CTC performance**

See configurations section for naming convention.

***NNVC-4.0 vs NNVC-5.0 VTM mode***

NNVC-5.0 in VTM mode performance are the same as the ones in NNVC-4.0 VTM.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over NNVC-4.0 VTM anchor** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 92% |
| Class A2 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 93% |
| Class B | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 94% |
| Class C | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 96% |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 94% |
| Class D | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 95% |
| Class F | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 101% | 93% |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over NNVC-4.0 VTM anchor** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 101% | 93% |
| Class A2 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 93% |
| Class B | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 95% |
| Class C | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 95% |
| Class E | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 101% | 96% |
| **Overall** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 101% | 94% |
| Class D | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 97% |
| Class F | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 98% |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Low delay B Main10** | | | | | | | |
|  | **BD-rate Over NNVC-4.0 VTM anchor** | | | | | | | |
|  | 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 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 92% |
| Class C | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 94% |
| Class E | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 86% |
| **Overall** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 91% |
| Class D | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 95% |
| Class F | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 102% | 91% |

***NNVC-4.0 vs NNVC-5.0 anchor***

The NNVC-5.0 anchor included LC filter and Intra Prediction tools activated.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over nnvc4.0\_VTM-anchor** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -7.23% | -4.86% | -6.00% | -8.72% | -4.05% | -5.06% | 134% | 7667% |
| Class A2 | -6.56% | -5.78% | -5.09% | -6.53% | -4.14% | -2.90% | 130% | 7183% |
| Class B | -6.29% | -8.19% | -7.53% | -5.97% | -6.06% | -5.18% | 132% | 7834% |
| Class C | -6.62% | -10.53% | -9.55% | -6.26% | -7.91% | -5.58% | 132% | 7422% |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -6.62% | -7.66% | -7.28% | -6.71% | -5.77% | -4.81% | 132% | 7557% |
| Class D | -7.57% | -9.45% | -9.68% | -5.76% | -6.81% | -5.94% | 129% | 8193% |
| Class F | -3.07% | -5.88% | -4.94% | -3.05% | -4.76% | -3.54% | 148% | 9739% |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Low delay B Main10** | | | | | | | |
|  | **BD-rate Over nnvc4.0\_VTM-anchor** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 |  |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |  |
| Class B | -4.85% | -8.13% | -7.84% | -5.15% | -6.16% | -6.65% | 134% | 6812% |
| Class C | -5.24% | -11.13% | -9.00% | -5.61% | -10.74% | -4.02% | 126% | 6233% |
| Class E | -4.88% | -4.09% | -5.17% | -5.64% | -2.18% | -3.36% | 162% | 9877% |
| **Overall** | -4.98% | -8.12% | -7.56% | -5.42% | -6.69% | -4.95% | 138% | 7257% |
| Class D | -6.54% | -10.76% | -9.77% | -5.75% | -9.48% | -6.83% | 125% | 6400% |
| Class F | -2.05% | -4.88% | -4.46% | -2.29% | -3.08% | -4.12% | 148% | 8941% |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over nnvc4.0\_VTM-anchor** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -8.46% | -7.23% | -8.67% | -10.00% | -8.00% | -7.50% | 210% | 5134% |
| Class A2 | -6.61% | -8.04% | -7.94% | -6.90% | -6.44% | -5.44% | 204% | 4416% |
| Class B | -7.03% | -8.58% | -8.71% | -7.04% | -7.61% | -6.99% | 203% | 4493% |
| Class C | -7.34% | -9.76% | -9.92% | -7.61% | -8.38% | -7.64% | 190% | 3770% |
| Class E | -10.29% | -9.04% | -8.85% | -10.14% | -7.52% | -6.38% | 199% | 5149% |
| **Overall** | -7.81% | -8.60% | -8.87% | -8.16% | -7.64% | -6.86% | 201% | 4507% |
| Class D | -7.57% | -8.11% | -9.33% | -7.16% | -6.81% | -7.56% | 180% | 3760% |
| Class F | -4.55% | -5.97% | -5.65% | -4.24% | -4.75% | -4.38% | 155% | 3961% |

***NNVC-5.0 VTM mode vs NNVC-5.1 VTM mode + HOP (HOP only)***

These early results are obtained using model of epoch 18 of the HOP training process in float.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over NNVC-5.0 VTM anchor** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -9.65% | -18.14% | -22.49% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class A2 | -10.08% | -22.08% | -17.55% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class B | -8.92% | -23.34% | -22.97% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class C | -9.65% | -24.98% | -24.39% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -9.49% | -22.48% | -22.17% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class D | -10.89% | -24.02% | -24.54% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class F | -5.30% | -16.51% | -14.82% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-1.0** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -7.04% | -17.89% | -20.06% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class A2 | -6.97% | -19.25% | -16.77% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class B | -7.00% | -18.24% | -19.89% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class C | -7.94% | -18.64% | -21.26% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class E | -10.39% | -20.44% | -21.57% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| **Overall** | -7.77% | -18.81% | -19.98% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class D | -7.83% | -17.73% | -21.23% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class F | -5.53% | -16.45% | -16.17% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |

***NNVC-5.0 anchor vs NNVC-5.1 HOP***

These early results are obtained using model of epoch 18 of the HOP training process in float.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over NNVC-5.0 anchor** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | #VALUE! | #VALUE! | #VALUE! | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class A2 | #VALUE! | #VALUE! | #VALUE! | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class B | -4.32% | -17.27% | -17.62% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class C | -4.65% | -16.79% | -17.31% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | #VALUE! | #VALUE! | #VALUE! | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class D | -4.82% | -16.61% | -17.22% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class F | -3.33% | -12.04% | -10.99% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over NNVC-5.0 anchor** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | #VALUE! | #VALUE! | #VALUE! | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class A2 | #VALUE! | #VALUE! | #VALUE! | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class B | -2.88% | -13.90% | -15.16% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class C | -3.50% | -12.91% | -15.35% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class E | -4.40% | -16.48% | -17.90% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| **Overall** | #VALUE! | #VALUE! | #VALUE! | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class D | -2.90% | -13.35% | -15.66% | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |
| Class F | #VALUE! | #VALUE! | #VALUE! | 100.00% | 100.00% | 100.00% | #NUM! | #NUM! |

***NNVC-5.1 HOP final***

TBA

***NNVC-5.1 VTM mode with Post-filter***

Results are extracted from JVET-AD0163.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over NNVC-4.0 NNPF OFF** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -6.40% | -13.70% | -20.30% | -8.39% | -16.51% | -21.17% | 171% | 22100% |
| Class A2 | -4.22% | -18.38% | -13.60% | -3.45% | -18.07% | -11.95% | 165% | 20486% |
| Class B | -5.13% | -20.59% | -16.95% | -3.25% | -20.82% | -16.74% | 170% | 23102% |
| Class C | -3.59% | -18.92% | -15.99% | -2.40% | -17.62% | -15.89% | 149% | 22203% |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -4.79% | -18.32% | -16.69% | -4.09% | -18.56% | -16.44% | 164% | 22119% |
| Class D | -3.60% | -18.30% | -17.35% | 0.28% | -16.57% | -15.64% | 154% | 22471% |
| Class F | -1.42% | -13.15% | -10.64% | -1.24% | -14.56% | -12.42% | 215% | 14088% |

**Discussions**

***Previous discussions***

The following points were discussed or are still discussed in the group:

* SOLVED: Official configuration files for NNVC software and official configuration files for EE1 anchors.
* SOLVED: To enforce adopted contributions to provide training scripts to train a model from scratch
* SOLVED A MR (MR68) merged.

***New discussions***

* Final training strategy for HOP model is still work in progress.
* Some aspects of encoder optimization (eg RDO models) are not yet aligned with HOP model.

**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-5.0 software
* P: the results are the ones from previous NNVC software basis
* N: not tested.

The column “xcheck” is read as follow:

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

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Name | Tools | Configuration | Tested | Xcheck |
| NNVC-5.0 VTM mode | none | encoder\_xxx\_vtm.cfg | Y | Y |
| NNVC 5.0 Anchor/EE1 | Intra Pred + LC filterset | encoder\_xxx\_nnvc.cfg | Y | Y |
| NNVC-5.1. HOP | Intra Pred + HOP | encoder\_xxx\_nnvc.cfg + nn-based/HOP.cfg | P |  |
| HOP only | HOP | encoder\_xxx\_vtm.cfg + nn-based/HOP.cfg | Y |  |
| set0 | Loop filter set #0 | encoder\_xxx\_vtm.cfg + nn-based/NnlfOption\_1.cfg | P | Y |
| set1 | Loop filter set #1 | encoder\_xxx\_vtm.cfg + nn-based/NnlfOption\_2.cfg | P | Y |
| Intra | Intra prediction | nn-based/encoder\_xxx\_nnvc.cfg | P | Y |
| set0+intra | Loop filter set #0+ Intra prediction | nn-based/encoder\_xxx\_nnvc.cfg + nn-based/NnlfOption\_1.cfg | P | Y |
| set1+intra | Loop filter set #1+ Intra prediction | nn-based/encoder\_xxx\_nnvc.cfg + nn-based/NnlfOption\_2.cfg | P | Y |
| set0+rdo+intra | Loop filter set #0 + Rdo + Intra prediction | nn-based/encoder\_xxx\_nnvc.cfg + nn-based/NnlfOption\_1.cfg + --EncNnlfOpt=1 | P | Y |
| set1+rdo+intra+temporal filter | Loop filter set #1 + Rdo + Intra prediction + Temporal filter | nn-based/encoder\_xxx\_nnvc.cfg + nn-based/NnlfOption\_2.cfg + --EncNnlfOpt=1 + --NnlfSet1Multiframe=1 | P | Y |
| Sr | Super-resolution | encoder\_xxx\_vtm.cfg + nn-based/nnsr.cfg + nn-based/nnsr\_classAx.cfg | P | N |
| set0+sr | Loop filter set #0+Super-resolution | encoder\_xxx\_vtm.cfg + nn-based/NnlfOption\_1.cfg + nn-based/nnsr.cfg + nn-based/nnsr\_classAx.cfg | N | N |
| set1+sr | Loop filter set #1+Super-resolution | encoder\_xxx\_vtm.cfg + nn-based/NnlfOption\_2.cfg + nn-based/nnsr.cfg + nn-based/nnsr\_classAx.cfg | N | N |
| sr+intra | Super-resolution +Intra prediction | nn-based/encoder\_xxx\_nnvc.cfg + nn-based/nnsr.cfg + nn-based/nnsr\_classAx.cfg | P | Y |
| set0+sr+intra | Loop filter set #0+Super-resolution+Intra prediction | nn-based/encoder\_xxx\_nnvc.cfg + nn-based/NnlfOption\_1.cfg + nn-based/nnsr.cfg + nn-based/nnsr\_classAx.cfg | N | N |
| set1+sr+intra | Loop filter set #01+Super-resolution+Intra prediction | nn-based/encoder\_xxx\_nnvc.cfg + nn-based/NnlfOption\_2.cfg + nn-based/nnsr.cfg + nn-based/nnsr\_classAx.cfg | N | N |
| Pf | Adaptive post-filters | encoder\_xxx\_vtm.cfg + nn-based/nnpf\_int16.cfg | P | P |
| Intra+pf | Intra prediction + adaptive post-filter | nn-based/encoder\_xxx\_nnvc.cfg + nn-based/nnpf\_int16.cfg | P | N |

Note: current HOP results used early model of stage2/epoch18.

**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.

It was asked whether the low-operation filter also used three-stage training before it was adopted in the last meeting? Yes.

# Project development (19)

## AHG1: Deployment and advertisement of standards (1)

Contributions in this area were discussed at XXXX–XXXX on XXday XX July 2023 (chaired by XX).

[JVET-AE0154](https://jvet-experts.org/doc_end_user/current_document.php?id=13117) MC-IF VVC technical guidelines [L. Litwic (Ericsson), S. McCarthy (Dolby), S. Wenger (Tencent), J. Ridge (Nokia), B. Bross (Fraunhofer HHI), D. Rusanovskyy (Qualcomm), Alan Stein (InterDigital)]

## AHG2: Text development and errata reporting (1)

Contributions in this area were discussed at XXXX–XXXX on XXday XX July 2023 (chaired by XX).

[JVET-AE0155](https://jvet-experts.org/doc_end_user/current_document.php?id=13118) Editor commentary on the post-filter hint SEI message semantics [G. J. Sullivan, Y.-K. Wang (Editors)]

Addressed by the BoG on SEI messages for VSEI v3 and VVC v3.

## AHG3: Test conditions (0)

This section is kept as a template for future use.

## AHG3: Software development (1)

Contributions in this area were discussed at XXXX–XXXX on XXday XX July 2023 (chaired by XX).

[JVET-AE0181](https://jvet-experts.org/doc_end_user/current_document.php?id=13144) Scaling window support for VTM [S. Iwamura, S. Nemoto, A. Ichigaya (NHK)]

[JVET-AE0275](https://jvet-experts.org/doc_end_user/current_document.php?id=13238) Cross-check of JVET-AE0181 (Scaling window support for VTM) [F. Urban (InterDigital)] [late]

## AHG4: Subjective quality testing and verification testing (3)

Contributions in this area were discussed at XXXX–XXXX on XXday XX July 2023 (chaired by XX).

[JVET-AE0041](https://jvet-experts.org/doc_end_user/current_document.php?id=12988) AHG4: Report on AhG meeting on verification testing for VVC multilayer coding [M. Wien (AhG4 co-chair)]

[JVET-AE0219](https://jvet-experts.org/doc_end_user/current_document.php?id=13182) AHG4: Results of visual checking of SVVC VT streams with and without DMVR fix [M. Wien (AhG4 co-chair)]

[JVET-AE0227](https://jvet-experts.org/doc_end_user/current_document.php?id=13190) AHG4: experiments in preparation of scalable quality ladder visual testing [P. de Lagrange, F. Urban (InterDigital)] [late]

## AHG4: Test material (1)

Contributions in this area were discussed at XXXX–XXXX on XXday XX July 2023 (chaired by XX).

[JVET-AE0179](https://jvet-experts.org/doc_end_user/current_document.php?id=13142) AHG4: Renewed license statement for Ghost Town Fly and Undo Dancer 3D video test sequences [M. M. Hannuksela, N. Salonen (Nokia)] [late]

## AHG4: Quality assessment methodology (1)

Contributions in this area were discussed at XXXX–XXXX on XXday XX July 2023 (chaired by XX).

[JVET-AE0092](https://jvet-experts.org/doc_end_user/current_document.php?id=13040) [AHG4] Occupancy-only PSNR calculations for V3C V-PCC coding evaluation [S. Schwarz, M. M. Hannuksela (Nokia)]

## AHG5: Conformance test development (1)

Contributions in this area were discussed at XXXX–XXXX on XXday XX July 2023 (chaired by XX).

[JVET-AE0111](https://jvet-experts.org/doc_end_user/current_document.php?id=13059) Additional conformance tests of spatial scalability for multilayer coding [C. Salmon-Legagneur, P. de Lagrange, F. Urban (InterDigital)]

## AHG7: ECM tool assessment (2)

Contributions in this area were discussed at XXXX–XXXX on XXday XX July 2023 (chaired by XX).

[JVET-AE0174](https://jvet-experts.org/doc_end_user/current_document.php?id=13137) AHG7: On TM control for non-inter tools [Z. Deng, K. Zhang, L. Zhang (Bytedance)]

[JVET-AE0212](https://jvet-experts.org/doc_end_user/current_document.php?id=13175) Cross-check of JVET-AE0174 On TM control for non-inter tools [X. Li (Google)] [late]

[JVET-AE0256](https://jvet-experts.org/doc_end_user/current_document.php?id=13219) Crosscheck of JVET-AE0174 (On TM control for non-inter tools) [H. Wang (Qualcomm)] [late]

[JVET-AE0195](https://jvet-experts.org/doc_end_user/current_document.php?id=13158) AhG7: ECM stages complexity assessment [F. Galpin, F. Le Léannec, C. Salmon-Legagneur, E. François (InterDigital)] [late]

## AHG8: Optimization of encoders and receiving systems for machine analysis of coded video content (5)

Contributions in this area were discussed at XXXX–XXXX on XXday XX July 2023 (chaired by XX).

[JVET-AE0081](https://jvet-experts.org/doc_end_user/current_document.php?id=13029) [AHG8] De-noising filter as pre-processing for machine task C. Kim, D. Gwak, J. Lim (LGE)

[JVET-AE0221](https://jvet-experts.org/doc_end_user/current_document.php?id=13184) Crosscheck of JVET-AE0081 ([AHG8] De-noising filter as pre-processing for machine task) [C. Hollmann (Ericsson)] [late]

[JVET-AE0096](https://jvet-experts.org/doc_end_user/current_document.php?id=13044) [AHG8] Study on using different VTM versions [C. Hollmann, J. Ström (Ericsson)]

[JVET-AE0099](https://jvet-experts.org/doc_end_user/current_document.php?id=13047) [AHG8] NNPF and post-filter hint SEI messages for the technical report [C. Hollmann, M. Pettersson, R. Sjöberg, M. Damghanian (Ericsson)]

[JVET-AE0107](https://jvet-experts.org/doc_end_user/current_document.php?id=13055) AHG8: Improvements of the BD-rate model using monotonic curve-fitting method [H. Wang, X. Pan, Z. Liu, X. Xu, S. Liu (Tencent)]

[JVET-AE0234](https://jvet-experts.org/doc_end_user/current_document.php?id=13197) AHG8: Crosscheck of AE-0107 [H. Zhang (Nokia)] [late]

[JVET-AE0143](https://jvet-experts.org/doc_end_user/current_document.php?id=13106) AHG8: A spatial resampling algorithm and an exemplar software implementation [S. Wang, B. Li, J. Chen, Y. Ye (Alibaba), S. Wang (CityU)]

Contributions related to SEI messages:

[JVET-AE0053](https://jvet-experts.org/doc_end_user/current_document.php?id=13001) AHG8/AHG9: Neural-network post-filter regions SEI message [T. Chujoh, Y. Yasugi, T. Ikai (Sharp)]

[JVET-AE0064](https://jvet-experts.org/doc_end_user/current_document.php?id=13012) AHG8/AHG9: Signalling encoder preprocessing and human / machine viewing indications [C. Kim, D. Gwak, Hendry, J. Lim, S. Kim (LGE), M. M. Hannuksela, F. Cricri, H. Zhang (Nokia)]

[JVET-AE0079](https://jvet-experts.org/doc_end_user/current_document.php?id=13027) AHG8/AHG9: Source picture timing information SEI message [S. McCarthy, G. J. Sullivan, P. Yin (Dolby)]

[JVET-AE0090](https://jvet-experts.org/doc_end_user/current_document.php?id=13038) AHG8/AHG9: On machine vision indication [J. Gao, H.-B. Teo, C.-S. Lim, K. Abe (Panasonic)]

[JVET-AE0095](https://jvet-experts.org/doc_end_user/current_document.php?id=13043) AHG8/AHG9: proposed changes to the candidate new object mask information SEI message [P. de Lagrange, E. François, D. Doyen (InterDigital), J. Chen, S. Wang, Y. Ye (Alibaba)] [late]

## AHG10: Encoding algorithm optimization (5)

Contributions in this area were discussed at XXXX–XXXX on XXday XX July 2023 (chaired by XX).

[JVET-AE0057](https://jvet-experts.org/doc_end_user/current_document.php?id=13005) MTT split modes early termination [W. Ahmad, P. Wennersten, K. Andersson (Ericsson)]

[JVET-AE0215](https://jvet-experts.org/doc_end_user/current_document.php?id=13178) crosscheck of JVET-AE0057 MTT split modes early termination [K. Naser (InterDigital)] [late] [miss]

[JVET-AE0104](https://jvet-experts.org/doc_end_user/current_document.php?id=13052) AHG10: GOP-based RPR encoder control using parallel resolution encoding [D. Arai, S. Nemoto, S. Iwamura, A. Ichigaya (NHK)]

[JVET-AE0277](https://jvet-experts.org/doc_end_user/current_document.php?id=13240) Cross-check of JVET-AE0104 (AHG10: GOP-based RPR encoder control using parallel resolution encoding) [K. Andersson (Ericsson)] [late]

[JVET-AE0114](https://jvet-experts.org/doc_end_user/current_document.php?id=13062) AHG10: Low-Delay configuration improvements [J. Enhorn, P. Wennersten (Ericsson)]

[JVET-AE0283](https://jvet-experts.org/doc_end_user/current_document.php?id=13246) Cross check of JVET-AE0114 (AHG10: Low-Delay configuration improvements) [S. Hong (Nokia)] [late]

[JVET-AE0122](https://jvet-experts.org/doc_end_user/current_document.php?id=13070) AHG10: GOP-based RPR and scalable coding [K. Andersson (Ericsson)]

[JVET-AE0123](https://jvet-experts.org/doc_end_user/current_document.php?id=13071) AHG10: Lagrange multiplier optimization for chroma ALF and CCALF [S.-W. Xie, Y. Gao, M.-H. Jia, Y.-L. Hu, C. Huang, P. Wu (ZTE)]

[JVET-AE0262](https://jvet-experts.org/doc_end_user/current_document.php?id=13225) Crosscheck of JVET-AE0123 (AHG10: Lagrange multiplier optimization for chroma ALF and CCALF) [H. Zhang (Tencent)] [late] [miss]

## AHG13: Film grain synthesis (2)

This section is kept as a template for future use.

[JVET-AE0250](https://jvet-experts.org/doc_end_user/current_document.php?id=13213) Film Grain Synthesis: Material Selection for Ground Truth Content Creation [D. Podborski, D. Ugur, B. Williams, L. Levinson, R. Molholm, A. M. Tourapis (Apple)] [late]

[JVET-AE0251](https://jvet-experts.org/doc_end_user/current_document.php?id=13214) Evaluating Denoising methods in the context of Film Grain Analysis/Synthesis [D. Podborski, D. Ugur, A. M. Tourapis (Apple)] [late]

## Implementation studies (1)

Contributions in this area were discussed at XXXX–XXXX on XXday XX July 2023 (chaired by XX).

[JVET-AE0172](https://jvet-experts.org/doc_end_user/current_document.php?id=13135) Update on multilayer coding support for VVenC [S. Iwamura, S. Nemoto, A. Ichigaya (NHK)]

## Profile/tier/level specification (0)

This section is kept as a template for future use.

## Proposed modification of system interface (0)

This section is kept as a template for future use.

## General aspects of standards development and applications of standards (0)

This section is kept as a template for future use.

# Low-level tool technology proposals

## AHG11/AHG14: Neural network-based video coding (17)

### Summary, BoG reports, and information documents (3)

Contributions in this area were discussed at 1430–1540 on Wednesday 12 July 2023 (chaired by JRO).

[JVET-AE0023](https://jvet-experts.org/doc_end_user/current_document.php?id=13091) EE1: Summary report of exploration experiment on neural network-based video coding [E. Alshina, F. Galpin, Y. Li, D. Rusanovskyy, M. Santamaria, J. Ström, L. Wang, Z. Xie (EE coordinators)]

This report summarizes the activities of the Exploration Experiment 1 (EE1) performed between the JVET-AD and JVET-AE 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.

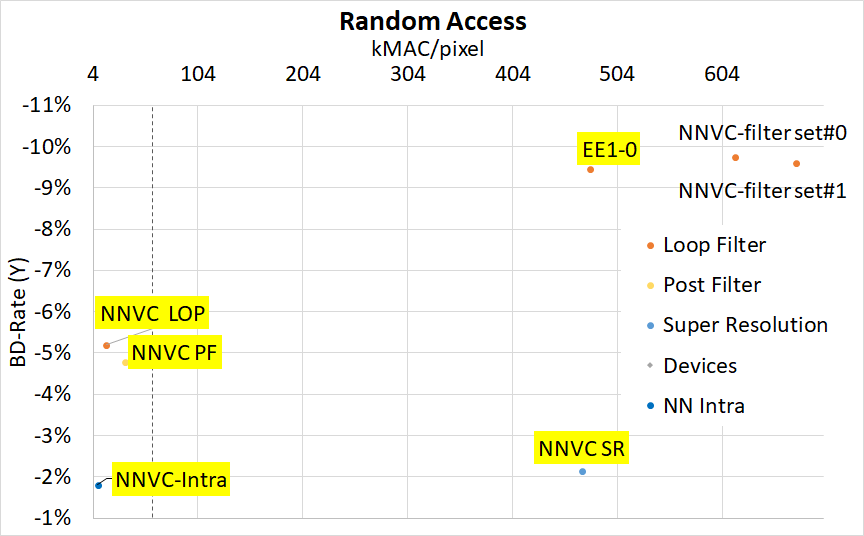
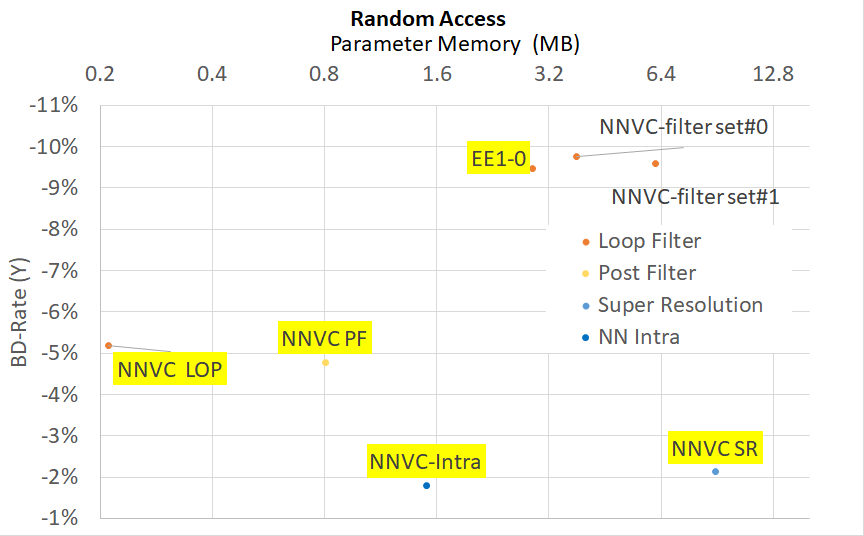
**Introduction**

In total seventeen test been planned for this EE1 round, but only seven were completed. Major reason for the delay is unified filter architecture for high operation point (HOP) and unified training procedure (both greed at JVET-AD meeting and documented in the JVET-AD0380 and JVET-AD2023) implemented from scratch and trained in parallel by multiple companies with careful cross-check of all intermediate steps and performance tracking. In addition, many unfinished EE tests are supposed to be conducted on top of the HOP model, which is not fully ready yet.

EE1 conducted two teleconferences jointly with AhG11 and AhG14 to discuss training of HOP unified filter (EE1-0). The teleconferences were held on June 14 and 28, 2023 and attended by approximately 20 participants.

In JVET-AD meeting low complexity operation point NN-based filter was agreed to be default configuration of NNVC-5.0. Also it was agreed to enable NN-Intra by default in NNVC-5.0 anchor. EE1 tests participant targeting for adoption to Common SW base NNVC are requested to report performance relatively to NNVC-5.0 default configuration.

In order to track NNVC work progress results relatively to VVC (NNVC-5.0 with all tools disabled) are shown in Figure 1.

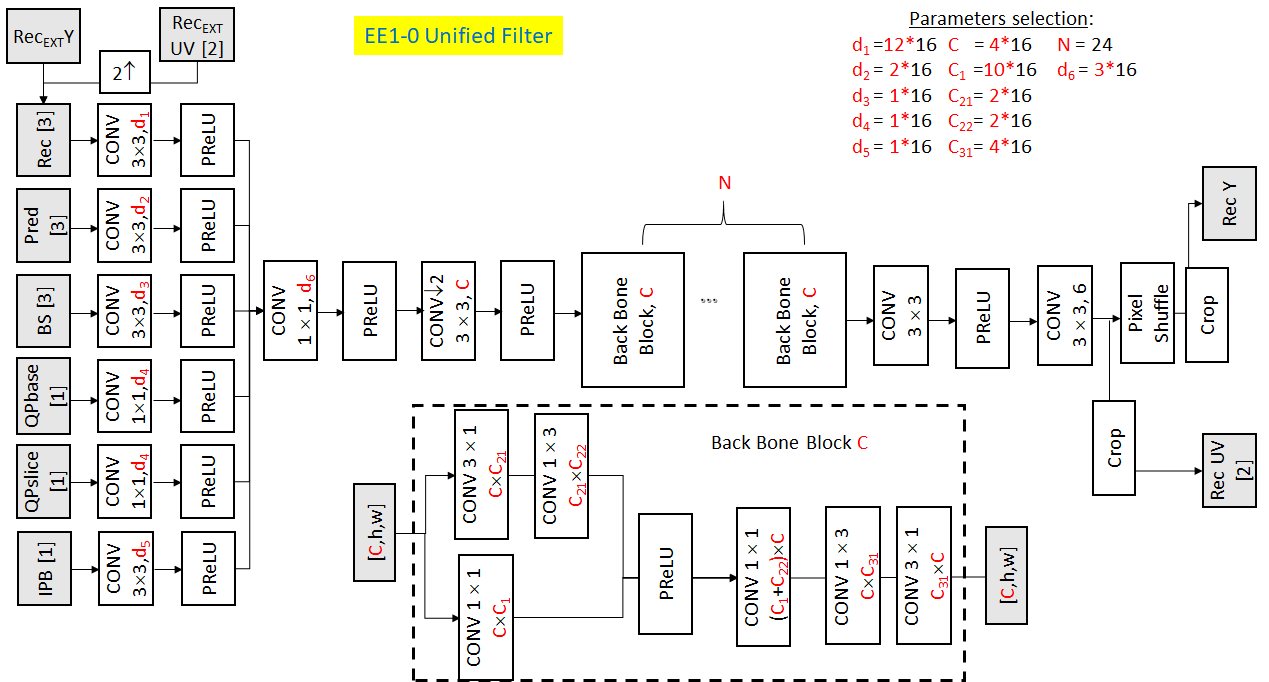
 

*Figure 1 Complexity performance trade-off of each NNVC tool in common SW base (old filters shown for the last time).*

**Unified Filter Training**

***EE1-0 Agreed unified filter design for High Operation Point***

[**JVET-AE0191**](https://jvet-experts.org/doc_end_user/current_document.php?id=13154) **AhG11: EE1-0 High Operation Point model** [F. Galpin (InterDigital)](mailto:franck.galpin@interdigital.com), S. Eadie, D. Rusanovskyy (Qualcomm), Y. Li, J. Li (ByteDance), L. Wang, R. Chang (Tencent), Z. Xie (Oppo), E. Alshina (Huawei)



*Figure 2 Unified filter architecture (HOP).*

Unified filter architecture (EE1.0) was documented in JVET-AD2023 and is shown in Figure 2. Computational complexity: 477 kMAC/pxl; Total Number of parameters: 1.45 M;

Unified training procedure of EE1.0 was also documented in JVET-AD2023 and consisted of following stages

1. Training Model stage I (Intra model)
   1. All Intra coding of DIV2K data coding with vanilla VTM (NN Tools off)
   2. Training data dumping and DIV2K unified database creation size of 2.6TB
   3. Model stage I training from scratch for Intra data (40 epochs, 10 days)
2. Training Model stage II (Intra & Inter model)
   1. Random Access coding of BVI and TVD data with NNVC5.0 (NN Tools off, HOP Stage I model enabled for Intra Slices)
   2. Training data dumping, BVI and TVD data unified databases creation size of 1.5TB
   3. Model stage II training with DIV2K, BVI and TVD databases (20 epochs, 10 days)
   4. Testing (5 days)
3. Training stage III (Intra & Inter model)
   1. Random Access coding of BVI and TVD data with NNVC5.0 (NN Tools off, HOP Stage II model enabled for Intra and B slices)
   2. Training data dumping, BVI and TVD data unified databases creation size of 1.5TB
   3. Model stage III training with DIV2K, BVI and TVD databases (20 epochs, 10 days)

It must be noted that collaborative development of HOP filter (EE1-0) allows achieving significant improvement for complexity-performance trade-off of NN-based filter at high operation point.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| HOP filters: past and future | Parameters, M | kMAC /pixel | Random Access vs VVC | | | All Intra vs VVC | | |
| Y | Cb | Cr | Y | Cb | Cr |
| EE1-0 (Stage I) | 1.5 | 477 | **-4.2%** | -17.4% | -16.2% | **-7.8%** | -19.2% | -20.2% |
| EE1-0 (Stage II) | 1.5 | 477 | **-9.5%** | -22.4% | -22.1% | **-7.8%** | -18.8% | -20.0% |
| filter set#0 (fully trained) | 1.9 | 615 | **-9.8%** | -21.1% | -20.5% | **-7.5%** | -16.8% | -17.4% |
| filter set#1 (fully trained) | 3.1 | 673 | **-9.6%** | -20.9% | -21.5% | **-7.4%** | -18.3% | -19.9% |

Only after two (out of three) stages training unified HOP filter (EE1-0) achieves almost the same performance with ste#0/set#1 filters in NNVC 4.0 with more than 20% lower computational complexity. Roughly 1% gain is expected after completion Stage II training (expected to finish during JVET-AE meeting).

Filter architecture EE1.0 and unified dataset and training process are recommended for adoption to the next version of NNVC common SW base.

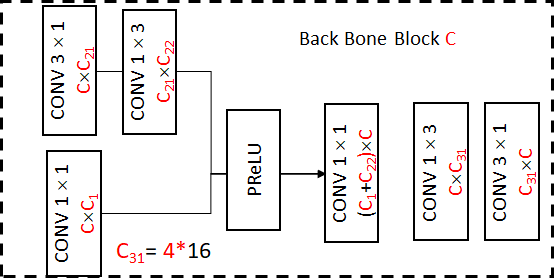
New finding during joint training of HOP are recommended to be presented with details (JVET-AE0191).

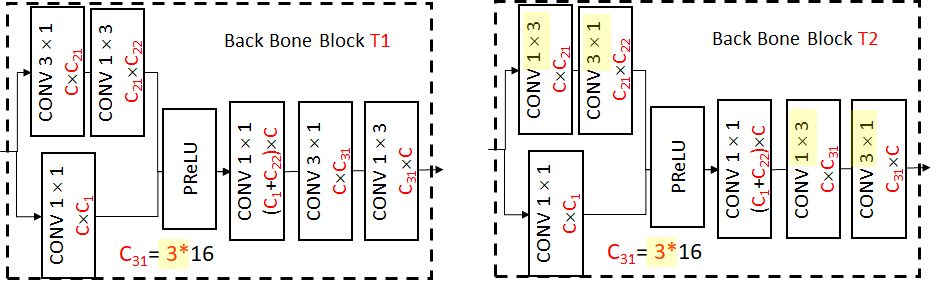
NOTE: “Skip connection” missed in HOP filter diagram, shall be added to diagrams as depicted on slide #6 of attached presentation.

**Unified Filter Architectural changes**

***EE1-1.2 Complexity-performance tradeoff of decomposition***

[**JVET-AE0164**](https://jvet-experts.org/doc_end_user/current_document.php?id=13127) **EE1-1.2 Complexity-performance tradeoff of decomposition** [D. Rusanovskyy](mailto:dmytror@qti.qualcomm.com), [Y. Li](mailto:yli30@qti.qualcomm.com), [M. Karczewicz (Qualcomm)](mailto:martak@qti.qualcomm.com)





*Figure 3 Back Bone Block: (above) EE1-0 design, all 24 BBB sare the same; (below) proposed design: every odd BBBs are T1, even BBBs are T2.*

Proposed modification for unified filter architecture is shown in Figure 3. Type 1 and Type 2 Back Bone Blocks feature reduced complexity, comparing to the anchor BBB, and are alternating (T1 and T2) in order. The model was trained following HOP training procedure and test set. Results for Stage1 have been provided, training for latter stages of HOP is ongoing, results to be provided once available.

Computational complexity: 426 kMAC/pxl (11% reduction compared to EE1-0); Total Number of parameters: 1.3M;

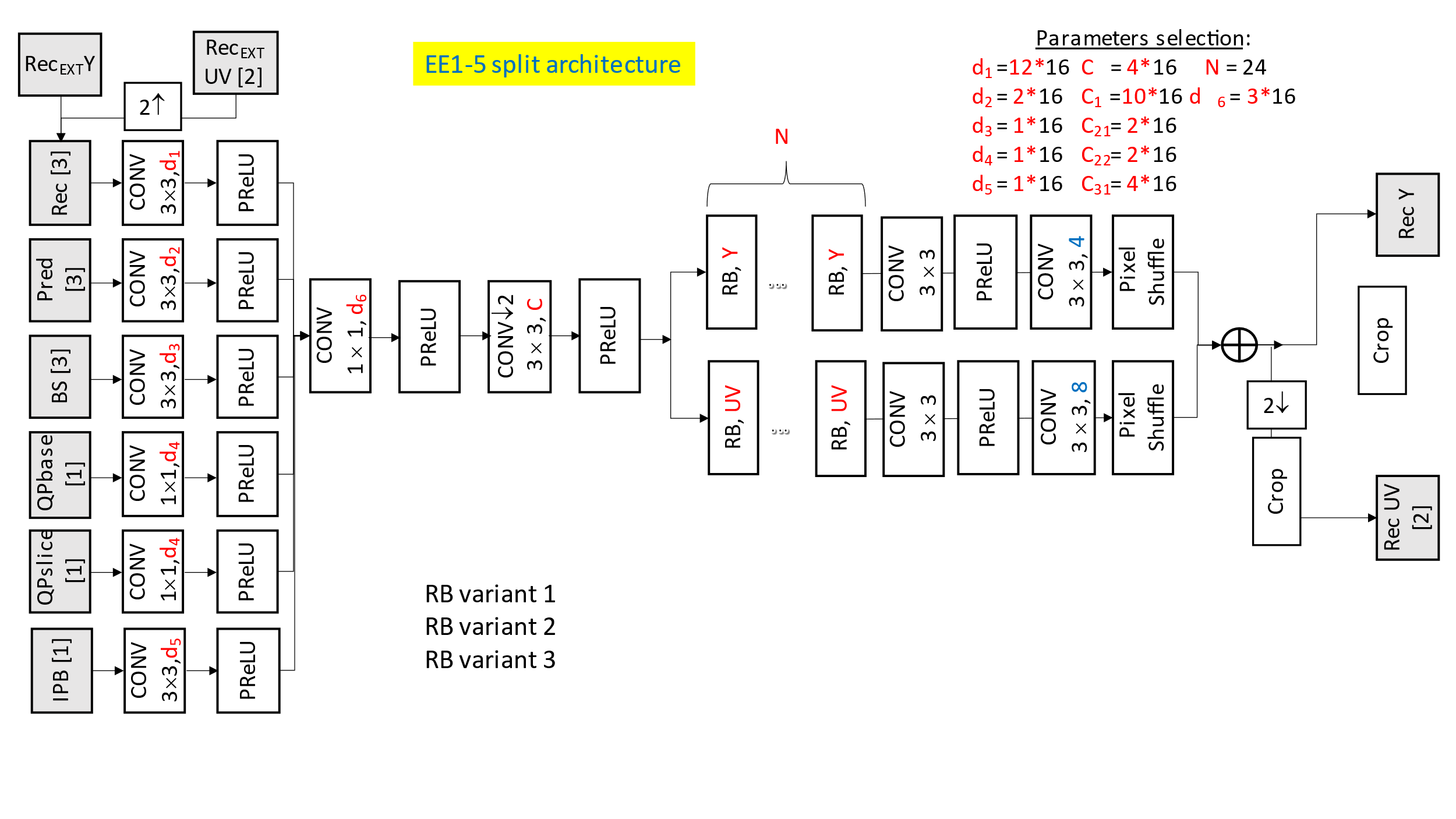
In all Intra cfg performance difference compare to EE1-0 (after same stage of training) is 0.0% (Y), 0.2%(Cb), 0.1%(Cr).

Suggestion: review in the BOG to plan study in EE1, e.g. provide results with 3 stages training, perform training cross-check.

***EE1-1.5 Optimization for complexity-performance trade-off of network***

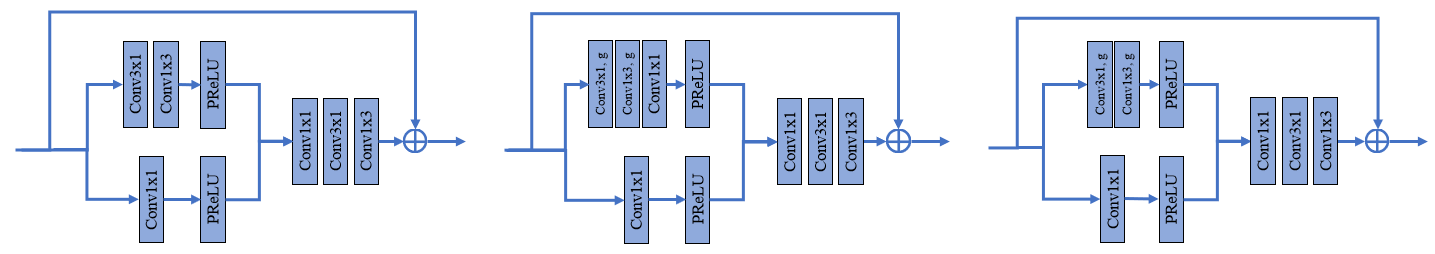
[**JVET-AE0160**](https://jvet-experts.org/doc_end_user/documents/31_Geneva/wg11/JVET-AE0160-v1.zip) **EE1-1.5: Optimization for complexity-performance trade-off of HOP network** [R. Chang](mailto:renjiechang@tencent.com), [L. Wang](mailto:liqiangwang@tencent.com), [X. Xu](mailto:xiaozhongxu@tencent.com), [S. Liu (Tencent)](mailto:shanl@tencent.com)

Proposed modification for unified filter architecture is shown in Figure 4.



*Figure 4 Proposed split Luma – Chroma filter architecture. Key changes: split architecture for luma and chroma compared to HOP architecture.*

Three variants of residual block are proposed. They are shown on Figure 5.



*Figure 5 Proposed variants of Residual Block.*

Based on filter0 training scripts and dataset, retrain the EE1-0 residual block and two improved residual blocks in Test 1, and retrain the EE1-0 architecture and split architecture in Test 2. Compared to the EE1-0 residual block, 11% kMAC/pixel reduction with only 0.1% luma gain loss for RA. Compared to the EE1-0 architecture, 34% kMAC/pixel reduction with only 0.5% luma gain loss for RA.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Test | Parameters, M | kMAC /pixel | Random Access vs NNVC-5.0 anchor | | | All Intra vs NNVC-5.0 anchor | | |
| Y | Cb | Cr | Y | Cb | Cr |
| NNIntra & EE1-1.5.1-subtest 1 | 1.4 | 461 | **-5.1%** | -14.6% | -14.5% | **-3.1%** | -12.4% | -13.5% |
| NNIntra & EE1-1.5.1-subtest 2 | 1.2 | 402 | **-4.8%** | -14.9% | -14.1% | **-2.8%** | -12.0% | -13.0% |
| NNIntra & EE1-1.5.1-subtest 3 | 1.3 | 408 | **-5.0%** | -14.2% | -13.8% | **-2.9%** | -11.8% | -13.0% |
| NNIntra & EE1-1.5.2-subtest 1 | 1.5 | 480 | **-3.6%** | -12.1% | -13.7% | **-2.3%** | -11.4% | -12.2% |
| NNIntra & EE1-1.5.2-subtest 2 | 0.9 | 316 | **-3.1%** | -11.6% | -11.5% | **-2.0%** | -10.6% | -11.5% |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Test** | **Subtest** | **Architecture** | **Back Bone Block** | **Training strategy** | **Training set** | **Interface** | **kMAC/pxl** |
| EE1-0 (HOP) |  | HOP | HOP resblock | HOP | HOP | HOP | 477 |
| EE1-1.5.1 | subtest1 (\*) | Filter set#0 | Same as HOP resblock | Filter set#0 | Filter set#0 | Same as HOP | 461 |
| subtest2 | HOP resblock with depth-wise separable | 408 |
| subtest3 | HOP resblock with group convolution | 402 |
| EE1-1.5.2 | subtest1 (\*\*) | Same as HOP | Same as HOP resblock | 480 |
| subtest2 | HOP with Split | 316 |

Suggestion: review in the BOG to plan study in EE1, e.g. utilizing HOP unified dataset and training process.

1. **Training set modification**

No specific study was conducted in this category. Test EE1-1.5 uses different training data and training strategy.

1. **Unified filter usage aspects**

No tests are possible at this point, since HOP filter is not yet available.

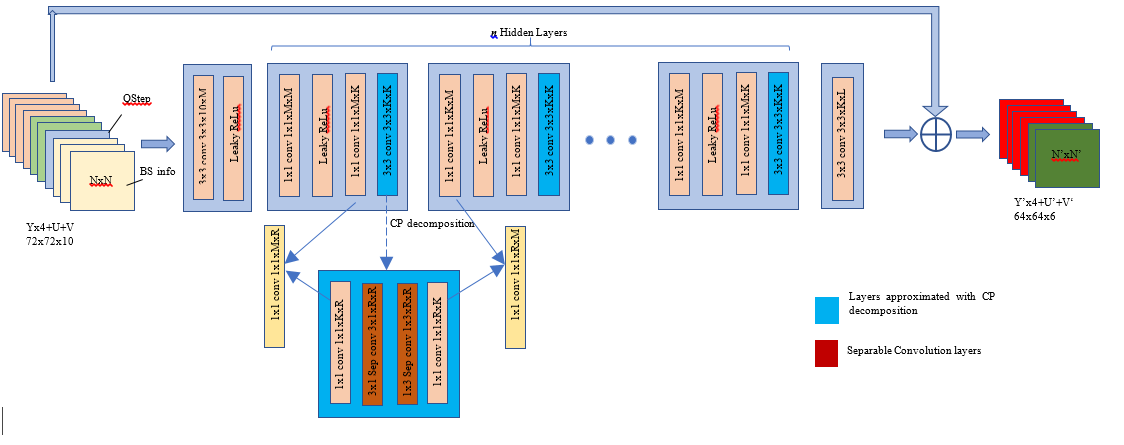
Recommendation to conduct scheduled tests in the next EE round.

1. **Improvement for Low complexity NN-filter**

***EE1-4.1 Neural-network loop filters with further complexity reduction***

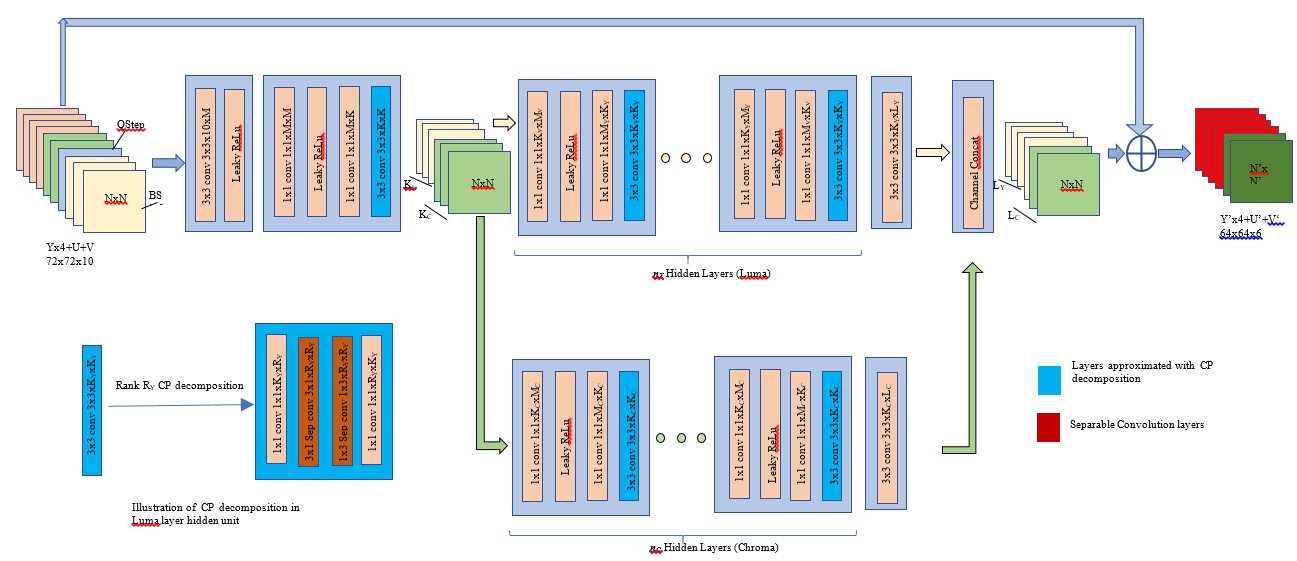
[**JVET-AE0067**](https://jvet-experts.org/doc_end_user/current_document.php?id=13015) **EE1-4.1: Neural-network loop filters with further complexity reduction** [J. N. Shingala](mailto:jay.shingala@ittiam.com), A. Shyam, A. Suneja, S. Badya (Ittiam), [T. Shao](mailto:tshao@dolby.com), A. Arora, [P. Yin](mailto:pyin@dolby.com), F. Pu, T. Lu, Sean McCarthy (Dolby)

Filter architecture of low operation point (LOP) filter in NNVC-5.0 is shown in Figure 6. Computational complexity is 16 kMAX/pxl. Number of parameters 0.21 M.



*Figure 6 Low Operation Point (LOP) filter in NNVC-5.0.*

Proposed modification for LOP are shown in Figure 7. . Parameters of the architecture are following: number of input feature sets M=64, K=24, R=24. Computational complexity is 10 kMAX/pxl. Number of parameters 0.13 M.



*Figure 7 Proposed modification for LOP filter.*

Complexity reduced drastically, performance degradation is 0.3...0.4% (performance gain of LOP filter is ~5%). Only float point model is provided. Performance test results is summarized in Table 1 .

*Table 1 Performance of low-complexity filter EE1-4.1 in comparison with LOP filter in NNVC-5.0*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Test | **Random Access vs VVC** | | | **All Intra vs VVC** | | |
| Y | Cb | Cr | Y | Cb | Cr |
| NNIntra & EE1-4.1 | **-6.2%** | -10.9% | -9.4% | **-7.5%** | -11.0% | -10.8% |
| Anchor(NNIntra & LOP NNVC-5.0) | **-6.6%** | -7.7% | -7.3% | **-7.8%** | -8.6% | -8.9% |
| NNIntra | **-1.8%** | -0.5% | -0.9% | **-3.6%** | -3.2% | -3.3% |
| LOP NNVC5.0 | **-5.2%** | -7.5% | -6.6% | **-4.6%** | -5.7% | -5.7% |

Suggestion: to continue study in EE1, provide quantized model and cross-check training.

***EE1-4.4 Low complexity NN filter with design elements of UF and EE1-1.2 and EE1-1.3***

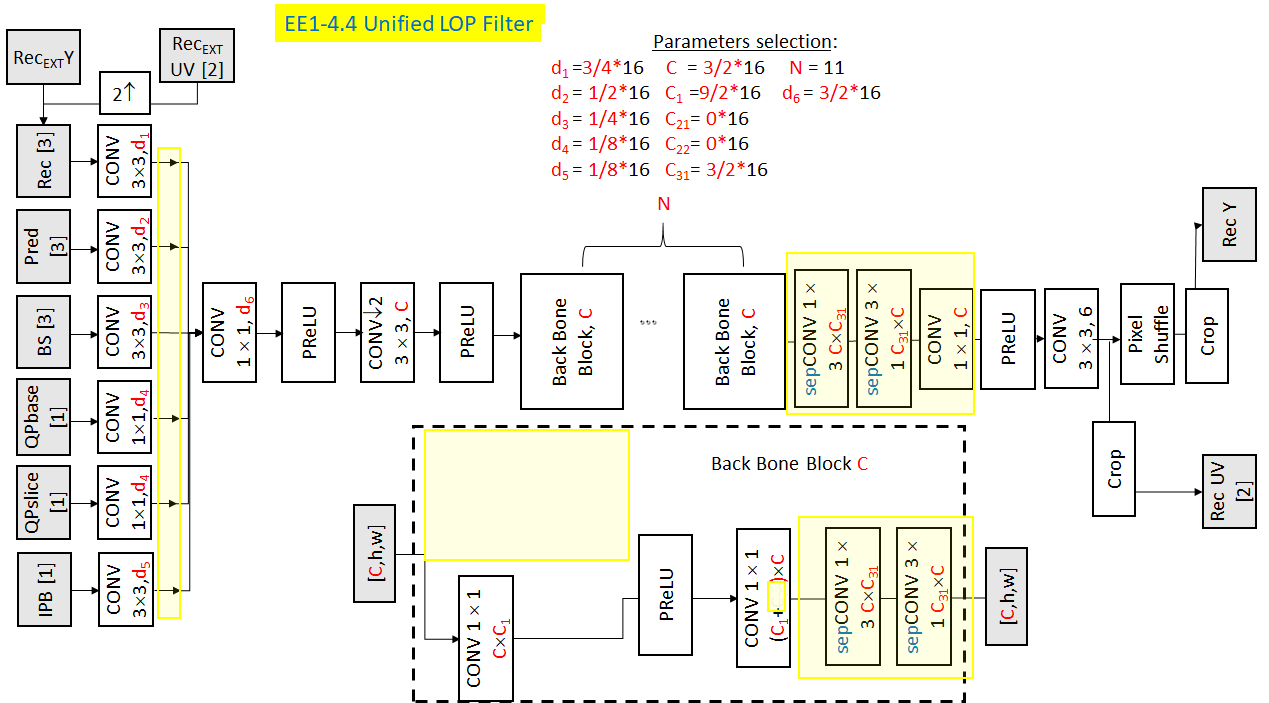
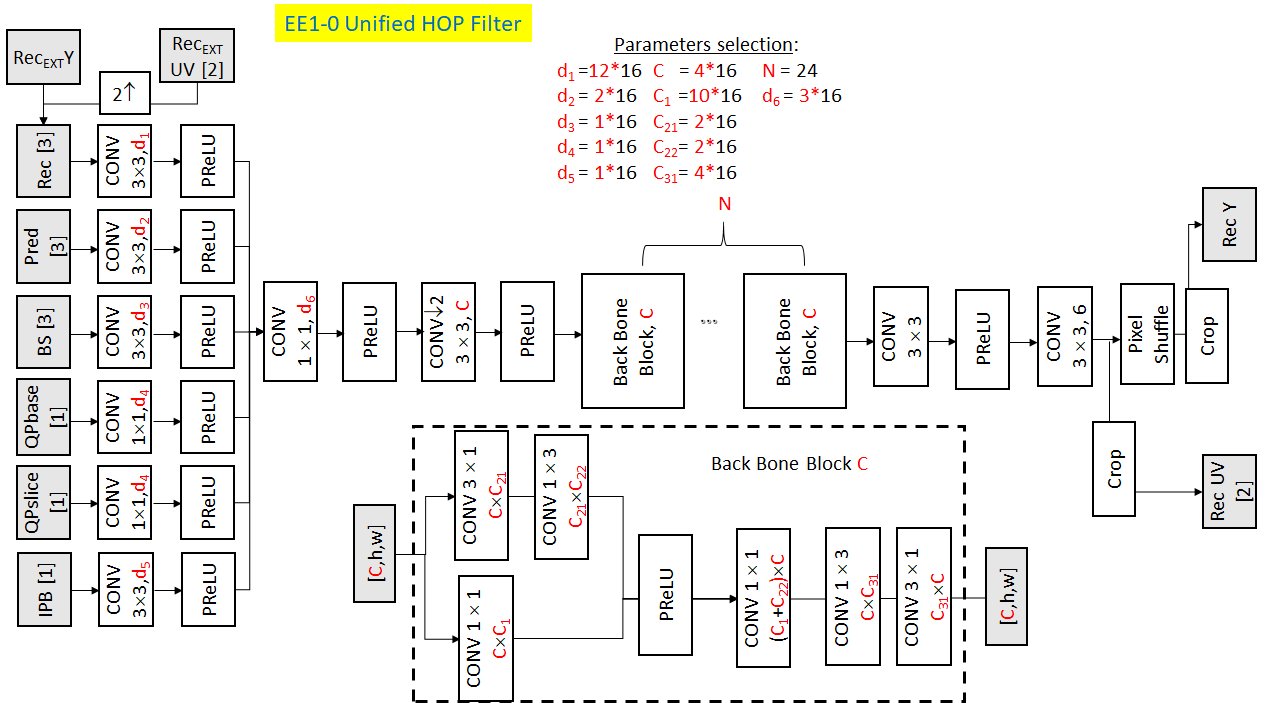
[**JVET-AE0165**](https://jvet-experts.org/doc_end_user/current_document.php?id=13128) **EE1-4.4: Low complexity NN filter with design elements of Unified Filter Architecture and EE1-1.2 and EE1-1.3** [Y. Li](mailto:yli30@qti.qualcomm.com), [D. Rusanovskyy](mailto:dmytror@qti.qualcomm.com), [M. Karczewicz (Qualcomm)](mailto:martak@qti.qualcomm.com) (Dolby)

In this contribution, Unified Filter Architecture of HOP was adjusted to meet complexity constraints of LOP. This was achieved by reducing complexity of the headblock, backbone, and tailblock and utilizing the CP4 decomposition in the backbone blocks. The model was trained End-to-End following HOP training procedure and test set. Results for Stage1 have been provided, training for latter stages of HOP is ongoing, results to be provided once available.

Training for latter stages of HOP is ongoing, results to be provided once available. Proposed filter, comparing to the EE1.0 anchor is shown in Figure 8 (above: EE1-0 filter, below: proposed low complexity filter).

Achieved model complexity in kMAC/pixel: 14.8 (frame-wise); 16.7 (block-wise of block 256x256); 18.7 (block-wise of block 128x128), total amount of parameters is 0.05 M (compared to 1.45 for HOP).

A white sheet with black text and red numbers

Description automatically generated

*Figure 8 Low complexity filter following concept of unified HOP.*

Performance test results are available for all intra configuration only, shown in Table 2. Training procedure follows EE1 recommendation, so results easily comparable with the rest of EE1 tests.

*Table 2 Performance of low-complexity filter EE1-4.4 in comparison with LOP filter in NNVC-5.0*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Test | **Random Access vs VVC** | | | **All Intra vs VVC** | | |
| Y | Cb | Cr | Y | Cb | Cr |
| NNIntra & EE1-4.4 |  |  |  | **-7.7%** | -13.6% | -14.3% |
| Anchor(NNIntra & LOP NNVC-5.0) | **-6.6%** | -7.7% | -7.3% | **-7.8%** | -8.6% | -8.9% |

Channels number in some point of processing pipeline have been selected to match complexity/number of channels of the LOP anchor. Due to this, they are multiple of 16 (even not multiple of 8), which could be sub-optimal for some implementations

Suggestion: review in the BOG to plan study in EE1, e.g for architecture optimization, provide results with HOP 3 stages training, perform training cross-check.

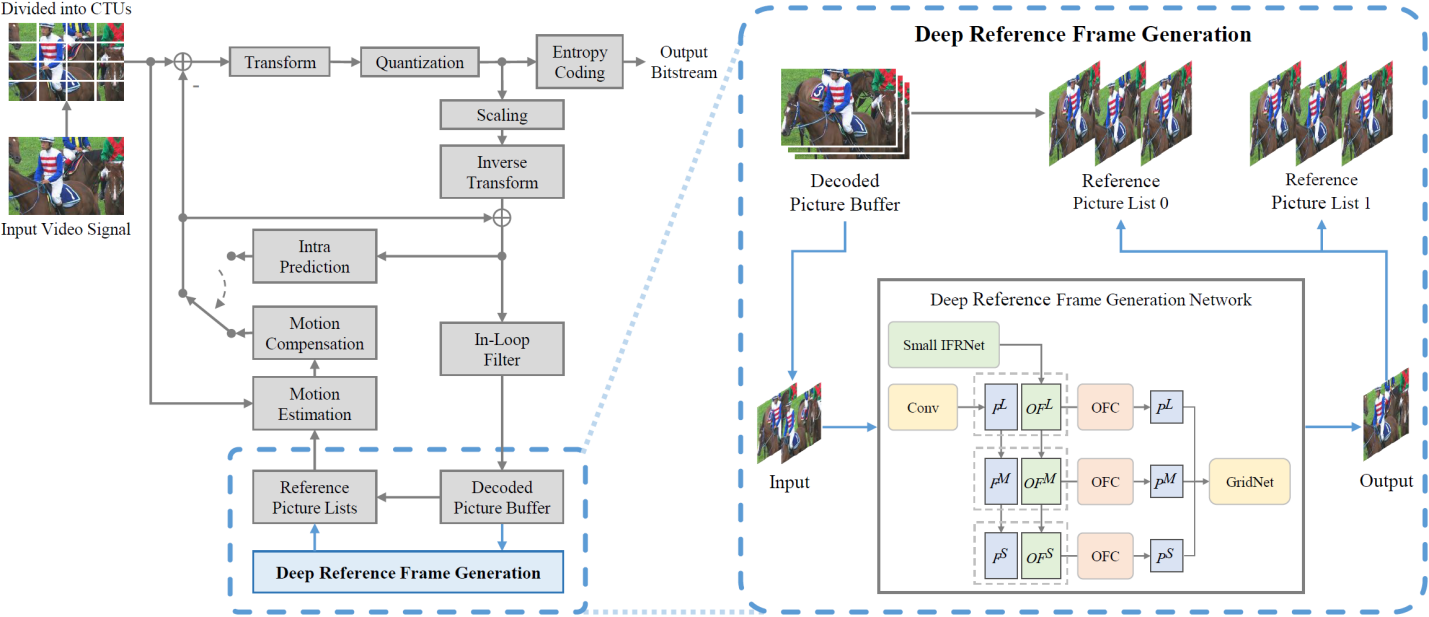
EE1-4.1 and EE1-4.4 still require training cross-check. This should be performed in the next round of EE. It was pointed out that proponents might think about a unification with the target of an improved low operation point filter, as this might speed up the process of training cross-check, and invoke a similar approach of multiple training stages as was done with the high operation point. If the proposals are kept independently, an equivalent training method should be used. Revisit after review

1. **NN-based Inter**

***EE1-5.1 NN-based inter prediction with lower complexity***

[**JVET-AE0112**](https://jvet-experts.org/doc_end_user/current_document.php?id=13060) **EE1-5.1: Deep Reference Frame Generation for Inter Prediction Enhancement** [W. Bao](mailto:baoweijie@whu.edu.cn), [W. Meng](mailto:whmeng@whu.edu.cn), [J. Jia](mailto:jiajh2021@whu.edu.cn), [Y. Zhang](mailto:yuantongzhang@whu.edu.cn), [H. Wang](mailto:wanghr827@whu.edu.cn), [Z. Chen (Wuhan Univ.)](mailto:zzchen@whu.edu.cn), [Z. Liu](mailto:zizhengliu@tencent.com), [X. Xu](mailto:xiaozhongxu@tencent.com), [S. Liu (Tencent)](mailto:shanl@tencent.com)

The framework of the proposed DRF method is depicted in Figure 9. In this approach, two reconstructed frames chosen from the DPB are inputted into the deep reference frame generation network to synthesize a new frame. This synthesized frame is then added to both reference picture list0 and list1, using the same POC as the current encoding frame. It serves as an additional reference frame for subsequent encoding processes.



*Figure 9 position of NN-Inter in video codec (test EE1-5)*

Changes of design compared to previous EE1 tests: removal of PFRB module and ResBlocks, IFRNet🡪 small IFRNet, SADL implementation (32 float).

Computational complexity is 504 kMAC/pxl (was 986 before), total number of parameters is 14.9 M (was 19.3 M).

Performance in random access configuration test (compared to previous design) is shown Table 3.

*Table 3 Performance of reduced complexity NN-Inter (EE1-5) compared to previously studied design.*

Performance after complexity reduction (EE1-5): kMAC/xpl 504, Number of parameters 14.9M.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over VTM-11.0\_NNVC-5.0** | | | | | | | |
| Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 |  |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |  |
| Class B | -3.2% | -7.3% | -6.6% | -3.9% | -9.3% | -8.4% | 700% | 1030222% |
| Class C | -4.5% | -7.0% | -7.2% | -6.0% | -9.7% | -10.0% | 441% | 761956% |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** |  |  |  |  |  |  |  |  |
| Class D | -5.5% | -5.9% | -6.5% | -5.7% | -10.4% | -10.1% | 751% | 1300136% |

Performance demonstrated in JVET-AD0162: kMAC/xpl 989, Number of parameters 19.3M.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | | | | | | | |
| **BD-rate Over VTM-11.0\_nnvc-4.0** | | | | | | | |
| Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -3.7% | -8.0% | -8.1% | -4.1% | -9.4% | -9.0% | 157% | 65726% |
| Class A2 | -5.4% | -11.7% | -9.4% | -6.2% | -11.5% | -9.2% | 147% | 60992% |
| Class B | -4.1% | -10.3% | -9.8% | -4.7% | -11.2% | -10.8% | 159% | 79960% |
| Class C | -5.7% | -13.4% | -13.6% | -7.3% | -13.5% | -14.3% | 123% | 47989% |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -4.7% | -11.0% | -10.4% | -5.6% | -11.5% | -11.0% | 146% | 63562% |
| Class D | -7.2% | -14.2% | -15.6% | -6.8% | -14.8% | -15.2% | 113% | 32008% |

Reduction of computational complexity roughly twice leads to ~1 % compression performance degradation.

Suggestion: Get opinion from cross-checker during the meeting

Revisit after completion of cross-check (training completed, inference ongoing)

Still floating point – integerization should be studied in next EE

Also the interdependency with loop filter should be studied – how much of the gain is retained in combination with low-operation loop filter (which will be one of the anchors).

1. **NN-based intra**

***EE1-6.1 Neural network-based intra prediction with reduced complexity***

[**JVET-AE0144**](https://jvet-experts.org/doc_end_user/current_document.php?id=13107) **EE1-6.1: neural network-based intra prediction with reduced complexity**  [T. Dumas](mailto:thierry.dumas@interdigital.com), F. Galpin, P. Bordes (Interdigital)

NN-Intra is enabled by default in NNVC-5.0 Design of NN-Intra is show in Figure 10.

Note that all the mentioned layers are fully-connected. No convolutional layers exist in the presented neural networks. Changes compared to NN-Intra in common SW base (NNVC0) are summarized in Table 4. The percentage of non-zero weights in the layer returning becomes 10% for the neural network predicting blocks of size , and 50% for the other six neural networks.

Computational complexity is different for different block sizes, the worst case is 4.8 kMAC/pixel (7.8 in NNVC). Total number of parameters is 1.3 M (1.5 in NNVC).

In AI, full test results show ~0.2% BD-rate performance degradation compared to NN-intra in NNVC-5.0. In RA, partial test results (full test results will be available by start of the meeting) show ~0.1% (RA) BD-rate performance degradation compared to NN-intra in NNVC-5.0.

Add full results from an update of the report

Revisit after completion of training cross-check.

Diagram

Description automatically generated

*Figure 10 : prediction of the current block from its context of reference samples via the neural network , parametrized by , belonging to the neural network-based intra prediction mode. Here, and The dimensions of the context of reference samples, “preprocessing”, “postprocessing”, , , are explained in NNVC description document.*

*Table 4 architecture policy for each neural network in common SW base NNVC and proposed solution (EE1-6).*

|  |  |  |
| --- | --- | --- |
| **architecture** | **NN-Intra in common SW base** | **EE1-6** |
| Number of hidden layers | For 16x16, 3.  For the other six neural networks, 2. | Same. |
| Number of neurons per hidden layer | 1216. | For 4x4, 8x4, and 16x16, 1216 for the last hidden layer, and 640 for the other hidden layers.  For the other four neural networks, 1216. |
| Biases | For all layers. | Only for the layer returning and the layer returning and . |

Suggestion: looking at Bytedance’s training cross-check to decide.

1. **Recommendation:**

* Present unified HOP filter (JVET-AE0191) in details
* Adopt HOP filter from JVET-AE0191 (after training is finished) to NNVC-6.0.
* Continue investigating pending EE tests of this round (supposed to be conducted on top of the HOP model) in the next round of EE.
* Conduct BoG to review input contributions and plan EE1 tests, identify potential for joint training in the next EE1 round (could be two designs for LOP and HOP).
* Conduct training cross-check for low-complexity NN-Intra (EE1-6)

Review of EE1 & related documents (5.1.2/5.1.3) on Thursday morning, BoG (E. Alshina, F. Galpin) after that with the mandate of preparing the next EE, software development, etc.

[JVET-AE0042](https://jvet-experts.org/doc_end_user/current_document.php?id=12990) AhG14 & AHG11: Report on AhG teleconference on high operation point (HOP) unified filter training [E. Alshina (AhG11 co-chair)]

No need for presentation – was included in AHG report JVET-AE0011.

[JVET-AE0237](https://jvet-experts.org/doc_end_user/current_document.php?id=13200) Presentation of AI-based video Coding in "Third AG4 Workshop on JPEG and MPEG Emerging Activities" [E. Alshina]

For information – no need for presentation.

### EE1 contributions: Neural network-based video coding (7)

Contributions in this area were discussed at 1130–1340 on Thursday 13 July 2023 (chaired by JRO).

[JVET-AE0191](https://jvet-experts.org/doc_end_user/current_document.php?id=13154) AhG11 - EE1-0 High Operation Point model [F. Galpin (InterDigital), S. Eadie, D. Rusanovskyy (Qualcomm), Y. Li, J. Li (ByteDance), L. Wang, R. Chang (Tencent), Z. Xie (Oppo), E. Alshina (Huawei)] [late]

A Unified Filter Architecture for the High-performance Operation Point (HOP) was defined at JVET-AD0380 and includes filter architecture, training process and training test set. Every design element and training procedure for this architecture were carefully selected though multiple rounds of earlier EE1 tests. Following JVET-AD meeting, a unified dataset and unified training procedure was implemented by multiple participants, and training conducted in parallel independently to allow cross-check.

This contribution reports on the progress of unified filter implementation and achieved simulation results.

Intermediate results demonstrate the performance of the model: -7.78%, -18.81%, -19.98% BD-rate change in Y, U and V components for AI configuration and -9.49%, -22.48%, -22.17% BD-rate change in Y, U and V components for RA configuration.

It is suggested to adopt EE1-0 into NNVC common software and enable it by default for HOP. It is also suggested to make training strategy defined for EE1-0 a part of NNVC SW. It is also suggested to continue practice of joint training activity for most promising NN-based tools studied in EE1. Identified dis-convergence of training at Stage II can be one of aspects to study in further EE1 tests.

In stage 2, problem with switch from L1 to L2 loss function: In the first epoch (18), the loss is still decreasing, and then it increases. It was suggested that a solution might be reduction of the learning rate. Results above are with the model coming from epoch 18 of stage 2.

In each stage, training is performed from scratch. Currently, training of stage 3 is running, which uses the filter model from stage 2 only for generating decoded data, but then a completely new model is trained from scratch.

It was agreed that a unified training strategy shall be used in EE to make proposals in a given category comparable. In the same category (e.g., directly competing methods, such as filter for HOP), identical number of epochs shall be used, training parameters (loss function, batch size, switching point, learning rate decay) shall be identical.

For proposals that are exclusive (e.g., HOP and LOP loop filters), the following shall apply:

* It would be desirable (to be clarified in BoG if mandatory already in next round of EE) to define an equivalent mandatory strategy for LOP (the current LOP was trained somewhat similarly). Data set should be identical, but number of epochs may be different, as well as training parameters.
* not only the data set, but also augmentation, generation of coded data (QP points) should be identical.

Proponents are free to provide additional results with a different training strategy.

In terms of the HOP model for next NNVC SW the following was agreed (Decision):

* Inference results for integerized stage 2 are expected during the meeting, such that a decision could be made. D. Rusanovskyy will send the model to cross-checkers, and provide an input when results are available. This could become the “fallback candidate” for HOP in NNVC6 SW.
* Training for stage 3 is running (plus cross-checks), but will not be available before the end of the meeting. An integerization strategy for stage 3 should be negotiated in the BoG. If the integerized stage 3 model provides better results within 14 days after the meeting, it will replace the “fallback candidate” in NNVC6 (to be confirmed in telco on Aug. 2)

[JVET-AE0067](https://jvet-experts.org/doc_end_user/current_document.php?id=13015) EE1-4.1: Neural-network loop filters with further complexity reduction [J. N. Shingala, A. Shyam, A. Suneja, S. Badya (Ittiam), T. Shao, A. Arora, P. Yin, F. Pu, T. Lu, Sean McCarthy (Dolby)]

Was presented in JVET plenary Thu 13 July.

This contribution reports the experimental results of EE1-4.1. The models for complexity reduction in this EE test are originally proposed in JVET-AD0157. The model uses CP decomposition, fusing adjacent 1x1 convolution and split architecture for luma and chroma. The EE tests focus on two aspects: 1) test different combination of model parameters with further complexity reduction compared to NNVC 5.0 LOP (worse case block level complexity is 16.2 KMAC/Pixel); 2) 16x SIMD friendly. The new proposed model is implemented on top of NNVC common software NNVC-5.0.

The AI and RA results for the new proposed model of {Y, Cb, Cr} BD-Rate gain over NNVC-5.0 tools off are as follows:

EE1-4.1: CP-decomposition with fusing and split luma chroma (16L, 16C), worst case block level complexity of 9.87 KMAC/Pixel

* SADL, fp32: AI { -4.14%, -7.90%, 7.64%}, RA { -4.58%, -10.14%, -8.37%}

The AI and RA BD-Rate results over NNVC-5.0 anchor (NnIntraPred and Nnlf LOP tool on) are as follows:

* SADL, fp32: AI { 0.35%, -2.59%, -2.07%}, RA { 0.48%, -3.29%, -2.14% }

Figure 1 shows NNVC 5.0 NNLF LOP (Low Operation Point) architecture: CP-decomposition of 3x3 convolution layers and fusion of adjacent 1x1 pointwise convolutions. The 3x3 convolutions of each hidden layer are decomposed into 4 layers with rank *R* followed by fusion of adjacent 1x1 convolution:

* 1st layer: 1x1x*K*x*R* pointwise convolution
* 2nd layer: 3x1x*R*x*R* separable convolution
* 3rd layer: 1x3x*R*x*R* separable convolution
* 4th layer: 1x1x*R*x*K* pointwise convolution

Figure 2 shows the EE1-4.1 model: split Luma, Chroma architecture combined with CP decomposition and fusing of adjacent 1x1 conv layer. Various 16x SIMD friendly channel count combinations have been tested at about 10KMAC/Pixel worst case block-level complexity. The proposed new model setting is shown in Table 1.

3x3 conv 10xM

Leaky ReLu

3x3 conv KxL

1x1 conv MxM

Leaky ReLu

3x3 CPD KxK

1x1 conv MxK

1x1 conv KxM

Leaky ReLu

3x3 CPD KxK

1x1 conv MxK

BS info

QStep

Yx4+U+V 72x72x10

NxN

1x1 conv KxM

Leaky ReLu

3x3 CPD KxK

1x1 conv MxK

Y’x4+U’+V‘

64x64x6

N’xN’

***N*** Hidden Layers

1x1 conv KxR

3x1 Sep conv R

1x3 Sep conv R

CP decomposition

Layers approximated with CP decomposition

Separable Convolution layers

1x1 conv MxR

1x1 conv RxM

**Figure 1 NNVC 5.0 LOP: CP Decomposition + fusion of 1x1 conv layers**

Ky

NxN

KC

Channel Concat

LY

NxN

LC

3x3 CPD 10xM

Leaky ReLu

3x3 CPD KYxLY

1x1 conv MxM

Leaky ReLu

3x3 CPD KxK

1x1 conv MxK

Leaky ReLu

3x3 CPD KYxKY

1x1 conv MYxKY

BS info

QStep

Yx4+U+V 72x72x10

NxN

1x1 conv KYxMY

Leaky ReLu

3x3 CPD KYxKY

1x1 conv MYxKY

Y’x4+U’+V‘ 64x64x6

N’xN’

***NY*** Hidden Layers (Luma)

***NC*** Hidden Layers (Chroma)

1x1 conv MCxKC

1x1 conv KCxMC

3x3 CPD KCxKC

1x1 conv MCxKC

3x3 CPD KCxKC

1x1 conv KCxMC

Leaky ReLu

3x3 CPD KCxLC

Leaky ReLu

1x1 conv KYxMY

Layers approximated with CP decomposition

Separable Convolution layers

3x3 CPD KYxKY

Rank RY CP decomposition

1x1 conv KYxRY

3x1 Sep conv RY

1x1conv RYxKY

1x3 Sep conv RY

Illustration of CP decomposition in Luma layer hidden unit

Note: Adjacent 1x1 conv layers after 3x3 CPD of hidden layers are fused into single 1x1 conv layer

**Figure 2 Split luma and chroma model + CP Decomposition + fusion of 1x1 conv layer** **(EE1-4.1)**

**Table 1 EE1-4.1 Split Luma Chroma (16L, 16C) setting and complexity**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model/Exp | K | Ky | Kc | M | My | Mc | Ry | Rc | Ny | Nc | kmac |
| split luma chroma (16L, 16C) | 32 | 16 | 16 | 64 | 48 | 16 | 16 | 16 | 10 | 6 | 9.87 |

2 models for inter, 2 models for intra.

Proponents used the same training strategy as for current LOP filter. It was requested to provide slides for the BoG to understand the commonalities and differences with the HOP strategy.

Loss in luma, gain in chroma. It was suggested to try different weights, or modify number of channels (less in chroma).

[JVET-AE0183](https://jvet-experts.org/doc_end_user/current_document.php?id=13146) Crosscheck of JVET-AE0067 (EE1-4.1: Neural-network loop filters with further complexity reduction) [J. Ström (Ericsson)] [late]

[JVET-AE0112](https://jvet-experts.org/doc_end_user/current_document.php?id=13060) EE1-5.1: Deep Reference Frame Generation for Inter Prediction Enhancement [W. Bao, W. Meng, J. Jia, Y. Zhang, H. Wang, Z. Chen (Wuhan Univ.), Z. Liu, X. Xu, S. Liu (Tencent)]

See EE summary report – next EE to be defined in BoG

[JVET-AE0229](https://jvet-experts.org/doc_end_user/current_document.php?id=13192) Crosscheck of JVET-AE0112(EE1-5.1: Deep Reference Frame Generation for Inter Prediction Enhancement) [Z. Xie (OPPO)] [late]

[JVET-AE0144](https://jvet-experts.org/doc_end_user/current_document.php?id=13107) EE1-6.1: neural network-based intra prediction with reduced complexity [T. Dumas, F. Galpin, P. Bordes (Interdigital)]

Revisit after crosscheck availability

[JVET-AE0160](https://jvet-experts.org/doc_end_user/current_document.php?id=13123) EE1-1.5: Optimization for complexity-performance trade-off of HOP network [R. Chang, L. Wang, X. Xu, S. Liu (Tencent)]

See EE summary report – next EE to be defined in BoG

[JVET-AE0255](https://jvet-experts.org/doc_end_user/current_document.php?id=13218) Crosscheck of JVET-AE0160 (EE1-1.5: Optimization for complexity-performance trade-off of HOP network) [D. Liu (Ericsson)] [late]

[JVET-AE0164](https://jvet-experts.org/doc_end_user/current_document.php?id=13127) EE1-1.2 Complexity-performance tradeoff of decomposition [D. Rusanovskyy, Y. Li, M. Karczewicz (Qualcomm)]

See EE summary report – next EE to be defined in BoG

[JVET-AE0165](https://jvet-experts.org/doc_end_user/current_document.php?id=13128) EE1-4.4: Low complexity NN filter with design elements of Unified Filter Architecture and EE1-1.2 and EE1-1.3 [Y. Li, D. Rusanovskyy, M. Karczewicz (Qualcomm)]

A Unified Filter Architecture for the High-performance Operation Point (HOP) was defined at JVET-AC meeting and includes model architecture, training process and training test set. Unified data set creation, training, and performance assessment of the HOP architecture have been ongoing between JVET-AD and AE meetings, and Stage1 model and its performance was fully evaluated (on AI results) at the time of submission of this contribution. A Low-complexity Operation Point (LOP) in-loop filter, with complexity of 16.2 kMAC/pixel (frame-wise), was also adopted in JVET-AC meeting to the NNVC software to serve as an anchor for further HOP development.

In this contribution, Unified Filter Architecture of HOP was adjusted to meet complexity constraints of LOP. This was achieved by reducing complexity of the headblock, backbone and utilizing the CP decomposition in the backbone blocks. End-to-end training of this model was conducted following the process specified for HOP Stage1. Training for latter stages of HOP is ongoing, results to be provided once available.

Achieved model complexity in kMAC/pixel: 14.8 (frame-wise); 16.7 (block-wise of block 256x256); 18.7 (block-wise of block 128x128).

Comparing to NNVC-5.0 HOP Anchor (NN Intra and LOP filter enabled), the proposed filter (quantized model) provides: {0.08%, -5.41%, -5.83%} in AI test configuration. For the joint YCbCr metric, the BD-rate gain is -1.1%.

Comparing to NNVC-5.0 (NN tools OFF), the proposed filter (quantized model) provides: {-4.42%, -10.55%, -11.17%} in AI test configuration.

Comparing to NNVC-5.0 (NN tools OFF), the proposed filter (quantized model) with NN Intra enabled provides: {-7.73%, -13.64%, -14.28%} in AI test configuration.

It is proposed to study the proposed method in the next round of EE.

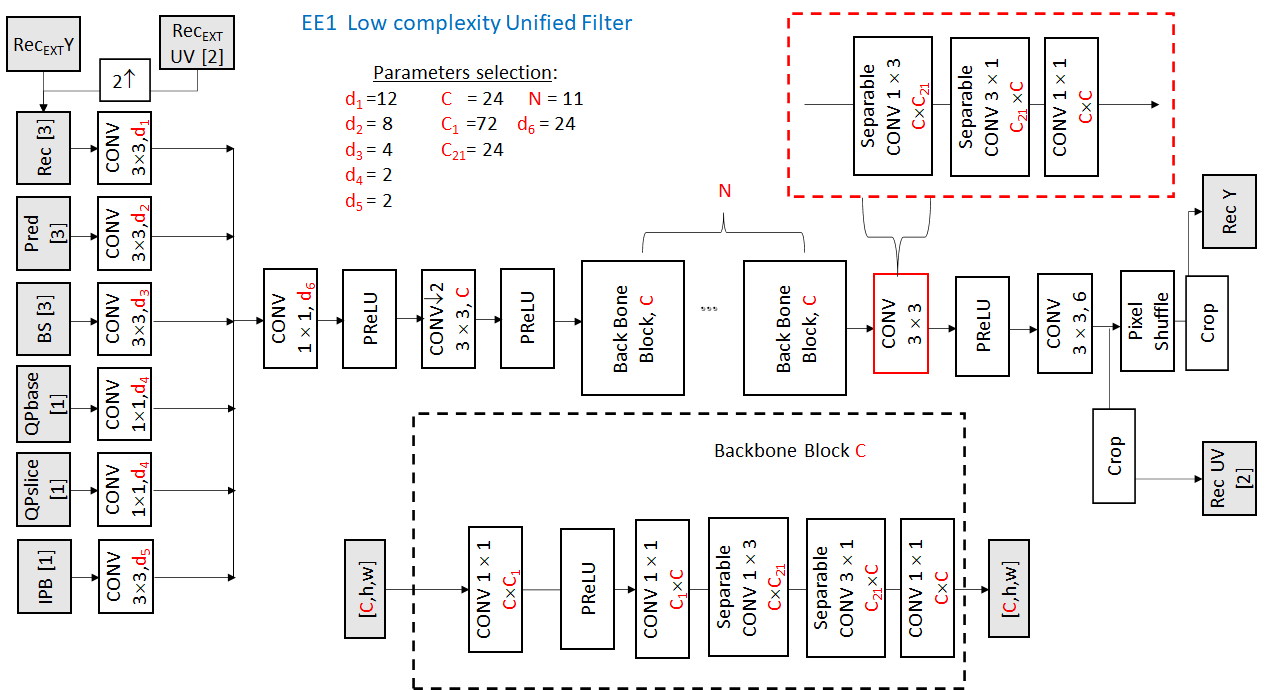


Figure 2. The proposed filter architecture. In the blocks marked with “Separable”, depth-wise convolution is performed.

The information for the inference and the training stage is presented in Table 1 and 2, respectively.

Table1. Network information for the inference

|  |  |  |  |
| --- | --- | --- | --- |
| **Network Information in Inference Stage** | | | |
|  | HW environment: | | |
| GPU Type | | N/A |
| Framework: | | SADL |
| Number of GPUs per Task | | 0 |
| Total Parameter Number | | 56544 /model |
| Parameter Precision (Bits) | | 32 or 16 |
| Memory Parameter (MB) | |  |
| Multiply Accumulate (kMAC/pixel) block (frame) based | | 18.7 (14.8 frame-wise) |
| Total Conv. Layers | | 67 |
| Total FC Layers | | 0 |
| Total Memory (MB) | Float |  |
| Int16 | 0.125/model |
| Batch size: | | 1 |
| Patch size | | 128128, 256256 |
| Changes to network configuration or weights required to generate rate points | |  |
| Peak Memory Usage | |  |
| Other information: | |  |

Table 2. Network information for the training

|  |  |  |
| --- | --- | --- |
| **Network Information in Training Stage** | | |
| Mandatory | GPU Type | GPU: NVIDIA Tesla V100-SXM2-32GB |
| Framework: | PyTorch v1.8 |
| Number of GPUs per Task | 1 |
|  |  |
| Epoch: | 130 |
| Batch size: | 64 |
| Training time: | ~240h |
| Training data information: | DIV2K |
| Training configurations for generating compressed training data (if different to VTM CTC): | QP {22, 27, 32, 37, 42} |
|  | Loss function: | L1, L2 |
| Optional | Number of iterations |  |
| Patch size | 144x144 |
| Learning rate: | 1e-4 |
| Optimizer: | ADAM |
| Preprocessing: |  |
| Other information: |  |

Backbone blocks not exactly identical to HOP model, for the benefit of complexity reduction.

Many more epochs were used in training than for HOP model.

It was requested to upload the modified software to EE repository.

It was suggested to merge convolutions at the input, as there is no non-linearity.

### EE1 related contributions: Neural network-based video coding (5)

Contributions in this area were discussed at XXXX–XXXX on XXday XX July 2023 (chaired by XX).

(include from BoG report JVET-AE0282)

[JVET-AE0072](https://jvet-experts.org/doc_end_user/current_document.php?id=13020) [AHG11] On residual adjustments for NNLF [Z. Dai, Y. Yu, H. Yu, D. Wang (OPPO)]

Reviewed in BoG

[JVET-AE0232](https://jvet-experts.org/doc_end_user/current_document.php?id=13195) Crosscheck of JVET-AE0072 ([AHG11] On residual adjustments for NNLF) [T. Shao (Dolby)] [late]

[JVET-AE0093](https://jvet-experts.org/doc_end_user/current_document.php?id=13041) AHG11: Content-adaptive neural network loop-filter [R. Yang, M. Santamaria, F. Cricri, H. Zhang, J. Lainema, R. G. Youvalari, M. M. Hannuksela (Nokia)]

Reviewed in BoG

[JVET-AE0161](https://jvet-experts.org/doc_end_user/current_document.php?id=13124) AHG11: Input and output rotation of model for NNVC in-loop filter [R. Chang, L. Wang, X. Xu, S. Liu (Tencent)]

Reviewed in BoG

[JVET-AE0276](https://jvet-experts.org/doc_end_user/current_document.php?id=13239) Crosscheck of JVET-AE0161 (AHG11: Input and output rotation of model for NNVC in-loop filter) [Z. Xie (OPPO)] [late]

[JVET-AE0171](https://jvet-experts.org/doc_end_user/current_document.php?id=13134) AHG11: Neural network-based in-loop filter with layer normalization [Y. Li, K. Zhang, L. Zhang (Bytedance)]

Reviewed in BoG

[JVET-AE0238](https://jvet-experts.org/doc_end_user/current_document.php?id=13201) AHG11: A unified design of NN-based loop-filters at low and high operation points [J. Li, Y. Li, C. Lin, K. Zhang, L. Zhang (Bytedance)] [late]

To be reviewed in BoG unless more study is needed before reviewing (very late upload)

[JVET-AE0281](https://jvet-experts.org/doc_end_user/current_document.php?id=13244) Unified LOP filter design, training procedure and filter usage [D. Rusanovskyy, Y. Li, M. Karczewicz (Qualcomm), T. Shao, P.Yin, S. McCarthy (Dolby), J. N. Shingala, A. Shyam(Ittiam), J. Li, Y. Li, C. Lin, K. Zhang, L. Zhang (Bytedance)] [late]

Reviewed in BoG

### Improvements of NNVC software beyond EE1 (3)

Contributions in this area were discussed at XXXX–XXXX on XXday XX July 2023 (chaired by XX).

(include from BoG report JVET-AE0282)

[JVET-AE0119](https://jvet-experts.org/doc_end_user/current_document.php?id=13067) AHG14: The extension of SADL library [W. Bao, H. Wang, Z. Chen (Wuhan Univ.)]

Reviewed in BoG

[JVET-AE0162](https://jvet-experts.org/doc_end_user/current_document.php?id=13125) AHG11/AHG14: Fix MS-SSIM calculation for SR [R. Chang, L. Wang, X. Xu, S. Liu (Tencent)]

Reviewed in BoG

[JVET-AE0230](https://jvet-experts.org/doc_end_user/current_document.php?id=13193) Crosscheck of JVET-AE0162(AHG11/AHG14: Fix MS-SSIM calculation for SR) [Z. Xie (OPPO)] [late]

[JVET-AE0194](https://jvet-experts.org/doc_end_user/current_document.php?id=13157) AhG14: SADL update [F. Galpin, T. Dumas, P. Bordes, E. François (InterDigital)] [late]

Reviewed in BoG

### Other aspects of neural network-based video coding (0)

Section kept as a template for future use.

## AHG6/AHG12: Enhanced compression beyond VVC capability (76)

### Summary and BoG reports

Contributions in this area were discussed at 1515–1815 on Tuesday 11 July 2023 and at 0830–1045 on Wednesday 12 July 2023 (chaired by JRO).

[JVET-AE0024](https://jvet-experts.org/doc_end_user/current_document.php?id=13092) EE2: Summary report of exploration experiment on enhanced compression beyond VVC capability [V. Seregin, J. Chen, R. Chernyak, K. Naser, J. Ström, F. Wang, M. Winken, X. Xiu, K. Zhang (EE coordinators)]

This document provides a summary report of Exploration Experiment on Enhanced Compression beyond VVC capability. The tests are categorized as partitioning, intra prediction, inter prediction, transform and coefficient coding, and in-loop filtering.

The software basis for this EE is ECM-9.0, released at <https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tags/ECM-9.0>. ECM-9.0 is used as an anchor in the tests.

Software for EE tests is released in the corresponding branches at <https://vcgit.hhi.fraunhofer.de/ecm/jvet-ad-ee2/ECM/-/branches>.

Test results can be found in input JVET contributions, cross-check results are uploaded to <https://vcgit.hhi.fraunhofer.de/ecm/jvet-ad-ee2/simulation-results> if cross-check reports are not submitted as they are optional for EE tests.

**List of tests**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Tests** | **Tester** | **Cross-checker** |
| **1 Partitioning** | | | |
| 1.1a | Partitioning prediction | Canon  G. Laroche  [JVET-AE0132](https://jvet-experts.org/doc_end_user/current_document.php?id=13095) | InterDigital  S. Puri |
| 1.1b | Partitioning prediction with unconditional MTT depth increment | Canon  G. Laroche  [JVET-AE0132](https://jvet-experts.org/doc_end_user/current_document.php?id=13095) | InterDigital  S. Puri |
| 1.1c | Encoder only partitioning prediction | Canon  G. Laroche  [JVET-AE0132](https://jvet-experts.org/doc_end_user/current_document.php?id=13095) | InterDigital  F. Le Léannec |
| **2 Intra prediction** | | | |
| 2.1a | Block vector guided CCCM with IBC BV | Nokia  R. Youvalari | B<>com  M. Abdoli |
| 2.1b | Block vector guided CCCM with IBC and IntraTMP BV | Nokia  R. Youvalari | B<>com  M. Abdoli |
| 2.2a | Extended IBC-GPM with two IBC predictions | Kwai C. Ma | withdrawn |
| 2.2b | Bi-predictive IBC-GPM | KDDI Y. Kidani  [JVET-AE0169](https://jvet-experts.org/doc_end_user/current_document.php?id=13132) | withdrawn |
| 2.2c | Bi-predictive IBC-GPM | Kwai C. Ma  KDDI Y. Kidani  [JVET-AE0169](https://jvet-experts.org/doc_end_user/current_document.php?id=13132) | Bytedance  W. Yin  [JVET-AE0197](https://jvet-experts.org/doc_end_user/current_document.php?id=13160) |
| 2.3a | IBC BVP-merge | KDDI Y. Kidani  [JVET-AE0169](https://jvet-experts.org/doc_end_user/current_document.php?id=13132) |  |
| 2.3b | Test 2.3a + bi-predictive IBC merge without IBC-GPM | KDDI Y. Kidani  [JVET-AE0169](https://jvet-experts.org/doc_end_user/current_document.php?id=13132) | Ericsson  R. Yu  [JVET-AE0205](https://jvet-experts.org/doc_end_user/current_document.php?id=13168)  Bytedance  W. Yin  [JVET-AE0197](https://jvet-experts.org/doc_end_user/current_document.php?id=13160) |
| 2.3c | Bi-predictive IBC merge without + Test 2.3a + Test 2.2c | KDDI Y. Kidani  Kwai C. Ma | withdrawn |
| 2.4a | IBC MBVD list derivation | Qualcomm  Z. Zhang  [JVET-AE0169](https://jvet-experts.org/doc_end_user/current_document.php?id=13132) | Ofinno  D. Ruiz Coll |
| 2.4b | Test 2.4a + Test 2.3a | Qualcomm  Z. Zhang  KDDI Y. Kidani | withdrawn |
| 2.4c | Test 2.4a + Test 2.3b + Test 2.2c | Qualcomm  Z. Zhang  KDDI Y. Kidani  Kwai C. Ma  [JVET-AE0169](https://jvet-experts.org/doc_end_user/current_document.php?id=13132) | Ericsson  R. Yu  [JVET-AE0205](https://jvet-experts.org/doc_end_user/current_document.php?id=13168) |
| 2.5a | Filtered IBC | Kwai H.-J. Jhu  [JVET-AE0159](https://jvet-experts.org/doc_end_user/current_document.php?id=13122) | withdrawn |
| 2.5b | Filtering IBC predicted blocks | Qualcomm B. Ray  [JVET-AE0159](https://jvet-experts.org/doc_end_user/current_document.php?id=13122) | withdrawn |
| 2.5c | Test 2.5a + Test 2.5b | Kwai H.-J. Jhu  Qualcomm B. Ray  [JVET-AE0159](https://jvet-experts.org/doc_end_user/current_document.php?id=13122) | OPPO  Y. Yu  L. Zhang  [JVET-AE0288](https://jvet-experts.org/doc_end_user/current_document.php?id=13191) |
| 2.6a | IBC with non-adjacent spatial candidates | Kwai  C. Ma | withdrawn |
| 2.6b | Non-adjacent spatial candidates for IBC | Bytedance  Y. Wang | withdrawn |
| 2.6c | IBC with non-adjacent spatial candidates (Test 2.6a + Test 2.6b) | Kwai  C. Ma    Bytedance  Y. Wang  [JVET-AE0094](https://jvet-experts.org/doc_end_user/current_document.php?id=13042) | Alibaba  J. Chen |
| 2.7 | Cross-component merge mode with temporal candidates | MediaTek H.-Y. Tseng | InterDigital P. Bordes |
| 2.8 | An extrapolation filter-based intra prediction mode | OPPO  L. Xu | Kwai  H.-J. Jhu  [JVET-AE0217](https://jvet-experts.org/doc_end_user/current_document.php?id=13180) |
| 2.9 | Extended search areas for IntraTMP mode | OPPO  Y. Yu  Xidian  Y. Ma,  H.Zhang  Kwai  X. Xiu  [JVET-AE0077](https://jvet-experts.org/doc_end_user/current_document.php?id=13025) | InterDigital  K. Naser  [JVET-AE0210](https://jvet-experts.org/doc_end_user/current_document.php?id=13173)  Qualcomm  P.-H Lin  [JVET-AE0216](https://jvet-experts.org/doc_end_user/current_document.php?id=13179)  Alibaba  X. Li  JVET-AE0245 |
| 2.9b | Extended search areas for IntraTMP mode with scan order #2 | OPPO  Y. Yu  Xidian  Y. Ma,  H.Zhang  Kwai  X. Xiu | withdrawn |
| 2.9c | IntraTMP mode with partial extended search areas with scan order #1 | OPPO  Y. Yu  Xidian  Y. Ma  H. Zhang  Kwai  X. Xiu | withdrawn |
| 2.9d | IntraTMP mode with partial extended search areas with scan order #2 | OPPO  Y. Yu  Xidian  Y. Ma  H. Zhang  Kwai  X. Xiu | withdrawn |
| 2.10a | IBC-LIC extension without large block-size constraint | OPPO  Z. Xie  [JVET-AE0078](https://jvet-experts.org/doc_end_user/current_document.php?id=13026) | Bytedance  W. Yin  [JVET-AE0157](https://jvet-experts.org/doc_end_user/current_document.php?id=13120) |
| 2.10b | ECM IBC-LIC without large block-size constraint | OPPO  Z. Xie  [JVET-AE0078](https://jvet-experts.org/doc_end_user/current_document.php?id=13026) | Bytedance  W. Yin  [JVET-AE0157](https://jvet-experts.org/doc_end_user/current_document.php?id=13120) |
| 2.11a | Harmonization between IBC HMVP and IBC-LIC | Bytedance  N. Zhang  [JVET-AE0084](https://jvet-experts.org/doc_end_user/current_document.php?id=13032) | Kwai C. Ma  [JVET-AE0222](https://jvet-experts.org/doc_end_user/current_document.php?id=13185) |
| 2.11b | Test 2.11a + Test 2.10 | Bytedance  N. Zhang  OPPO  Z. Xie  [JVET-AE0084](https://jvet-experts.org/doc_end_user/current_document.php?id=13032) | Kwai C. Ma  [JVET-AE0222](https://jvet-experts.org/doc_end_user/current_document.php?id=13185) |
| **3** **Inter prediction** | | | |
| 3.1a | Cross-component residual model | Nokia  P. Astola  [JVET-AE0059](https://jvet-experts.org/doc_end_user/current_document.php?id=13007) | InterDigital  F. Le Léannec  Ittiam J. Raj |
| 3.1b | Cross-component residual model with complexity reductions | Nokia  P. Astola  [JVET-AE0059](https://jvet-experts.org/doc_end_user/current_document.php?id=13007) | InterDigital  F. Le Léannec  Ittiam J. Raj |
| 3.2 | Bi-predictive GPM | Ericsson  R. Yu  [JVET-AE0046](https://jvet-experts.org/doc_end_user/current_document.php?id=12994) | LGE  Y. Ahn  [JVET-AE0199](https://jvet-experts.org/doc_end_user/current_document.php?id=13162)  KDDI Y. Kidani |
| 3.3a | Additional TM refinement for bi-prediction | Bytedance  [Y. Wang](mailto:wangyang.cs@bytedance.com)  [JVET-AE0087](https://jvet-experts.org/doc_end_user/current_document.php?id=13035) | Alibaba  J. Chen  [JVET-AE0239](https://jvet-experts.org/doc_end_user/current_document.php?id=13202) |
| 3.3b | 16-point diamond search pattern for TM | Bytedance  [Y. Wang](mailto:wangyang.cs@bytedance.com)  [JVET-AE0087](https://jvet-experts.org/doc_end_user/current_document.php?id=13035) | Alibaba  J. Chen  [JVET-AE0239](https://jvet-experts.org/doc_end_user/current_document.php?id=13202) |
| 3.3c | Enabling TM for bi-prediction under DMVR condition | Bytedance  [Y. Wang](mailto:wangyang.cs@bytedance.com)  [JVET-AE0087](https://jvet-experts.org/doc_end_user/current_document.php?id=13035) | Alibaba  J. Chen  [JVET-AE0239](https://jvet-experts.org/doc_end_user/current_document.php?id=13202) |
| 3.3d | Test 3.3a + Test 3.3c | Bytedance  [Y. Wang](mailto:wangyang.cs@bytedance.com)  [JVET-AE0087](https://jvet-experts.org/doc_end_user/current_document.php?id=13035) | Alibaba  J. Chen  [JVET-AE0239](https://jvet-experts.org/doc_end_user/current_document.php?id=13202) |
| 3.3e | Test 3.3a + Test 3.3b + Test 3.3c | Bytedance  [Y. Wang](mailto:wangyang.cs@bytedance.com)  [JVET-AE0087](https://jvet-experts.org/doc_end_user/current_document.php?id=13035) | Alibaba  J. Chen  [JVET-AE0239](https://jvet-experts.org/doc_end_user/current_document.php?id=13202) |
| 3.4 | HPel flag and BCW weight usage in OBMC | InterDigital  [A. Robert](mailto:antoine.robert@interdigital.com)  [JVET-AE0196](https://jvet-experts.org/doc_end_user/current_document.php?id=13159) | withdrawn |
| 3.5 | Iterative BDOF pass in multi-pass DMVR | Bytedance  M. Salehifar  Alibaba  J. Chen  [JVET-AE0065](https://jvet-experts.org/doc_end_user/current_document.php?id=13013) | vivo  Z. Lv  [JVET-AE0200](https://jvet-experts.org/doc_end_user/current_document.php?id=13163) |
| 3.6 | Affine AMVP mode with one MVD | Qualcomm H. Huang | withdrawn |
| 3.7a | RPR with new filters, scale factor 1.25x | Sharp [J. Samuelsson-Allendes](mailto:samuelssonj@sharplabs.com)  [JVET-AE0150](https://jvet-experts.org/doc_end_user/current_document.php?id=13113) | Qualcomm  Z. Zhang  [JVET-AE0204](https://jvet-experts.org/doc_end_user/current_document.php?id=13167) |
| 3.7b | RPR with new filters, scale factor 1.33x | Sharp [J. Samuelsson-Allendes](mailto:samuelssonj@sharplabs.com)  [JVET-AE0150](https://jvet-experts.org/doc_end_user/current_document.php?id=13113) | Qualcomm  Z. Zhang  [JVET-AE0204](https://jvet-experts.org/doc_end_user/current_document.php?id=13167) |
| 3.8 | Combination of Test 3.3e and Test 3.5 | Bytedance  M. Salehifar  Alibaba  J. Chen  [JVET-AE0091](https://jvet-experts.org/doc_end_user/current_document.php?id=13039) | Kwai  C. Ma  [JVET-AE0223](https://jvet-experts.org/doc_end_user/current_document.php?id=13186) |
| **4 Transform and coefficient coding** | | | |
| 4.1 | Shifting quantizer center | InterDigital  M. Balcilar  [JVET-AE0125](https://jvet-experts.org/doc_end_user/current_document.php?id=13079) | Qualcomm  M. Coban  [JVET-AE0149](https://jvet-experts.org/doc_end_user/current_document.php?id=13112) |
| 4.2 | Large NSPT | LGE  M. Koo  [JVET-AE0086](https://jvet-experts.org/doc_end_user/current_document.php?id=13034) | Qualcomm  P. Garus  [JVET-AE0118](https://jvet-experts.org/doc_end_user/current_document.php?id=13066) |
| 4.3 | Context modelling for transform coefficients for LFNST/NSPT | Qualcomm  P. Nikitin  [JVET-AE0102](https://jvet-experts.org/doc_end_user/current_document.php?id=13050) | IRT b-com M.Abdoli  [JVET-AE0225](https://jvet-experts.org/doc_end_user/current_document.php?id=13188) |
| 4.4a | InterMTS is enabled for IBC-coded blocks in AMVP mode. | Qualcomm  P. Garus  [JVET-AE0116](https://jvet-experts.org/doc_end_user/current_document.php?id=13064) | InterDigital K. Naser  [JVET-AE0211](https://jvet-experts.org/doc_end_user/current_document.php?id=13174) |
| 4.4b | Test 4.4a + IntraTMP using interMTS instead of intraMTS kernels | Qualcomm  P. Garus  [JVET-AE0116](https://jvet-experts.org/doc_end_user/current_document.php?id=13064) | InterDigital K. Naser  [JVET-AE0211](https://jvet-experts.org/doc_end_user/current_document.php?id=13174) |
| 4.4c | IntraMTS disabled for IntraTMP | Qualcomm  P. Garus  [JVET-AE0116](https://jvet-experts.org/doc_end_user/current_document.php?id=13064) | InterDigital K. Naser  [JVET-AE0211](https://jvet-experts.org/doc_end_user/current_document.php?id=13174) |
| **5 In-loop filtering** | | | |
| 5.1a | CCSAO with temporal history offset | Kwai  C.-W. Kuo  [JVET-AE0151](https://jvet-experts.org/doc_end_user/current_document.php?id=13114) | Qualcomm  N. Hu  [JVET-AE0209](https://jvet-experts.org/doc_end_user/current_document.php?id=13172) |
| 5.1b | Test 5.1a + extended edge classifier | Kwai  C.-W. Kuo  [JVET-AE0151](https://jvet-experts.org/doc_end_user/current_document.php?id=13114) | Qualcomm  N. Hu  [JVET-AE0209](https://jvet-experts.org/doc_end_user/current_document.php?id=13172) |
| 5.2a | Applying fixed filters to samples before DBF | Qualcomm  N. Hu  [JVET-AE0139](https://jvet-experts.org/doc_end_user/current_document.php?id=13102) | Kwai  C.-W. Kuo  [JVET-AE0193](https://jvet-experts.org/doc_end_user/current_document.php?id=13156) |
| 5.2b | Test 5.2a + extended classifiers for fixed filters | Qualcomm  N. Hu  [JVET-AE0139](https://jvet-experts.org/doc_end_user/current_document.php?id=13102) | Kwai  C.-W. Kuo  [JVET-AE0193](https://jvet-experts.org/doc_end_user/current_document.php?id=13156) |
| 5.2c | Test 5.2b + applying the second fixed filter to outputs of the first fixed filter | Qualcomm  N. Hu  [JVET-AE0139](https://jvet-experts.org/doc_end_user/current_document.php?id=13102) | Kwai  C.-W. Kuo  [JVET-AE0193](https://jvet-experts.org/doc_end_user/current_document.php?id=13156) |
| 5.3 | Combination of Test 5.1b and Test 5.2c | Kwai  C.-W. Kuo  Qualcomm  N. Hu  [JVET-AE0152](https://jvet-experts.org/doc_end_user/current_document.php?id=13115) | Alibaba  J. Chen |

**Category 1: Partitioning prediction**

**Test 1.1:** Partitioning prediction **(**[**JVET-AE0132**](https://jvet-experts.org/doc_end_user/documents/30_Antalya/wg11/JVET-AD0147-v1.zip)**)**

In this test, a temporal partitioning is introduced, where for each block, the allowed partitioning splits are predicted according to the minimum QT/MTT split and the average QT/MTT split obtained from a temporal area:

* if the current QT depth is less than the temporal minimum QT depth minus 1, only the QT split is allowed,
* if the current QT depth is less than the temporal average QT depth minus 1, no split, QT and TT splits are allowed, and BT split is allowed if TT is selected in a parent node,
* if the temporal maximum MTT depth and less than the maximum MTT depth of the current block, the depth is decreased except if the current QP is less than the temporal QP,
* if the temporal maximum MTT depth is large than the maximum MTT depth of the current block and if the current depth is equal to the current QT depth, the maximum MTT depth is incremented,
* if the maximum QT depth of the current frame is larger than the maximum multi-tree depth and if the temporal maximum MTT depth is equal to the maximum MTT depth of the current frame, and if the temporal QT depth is less than the QT depth of the current frame, the maximum MTT depth of the current block is decreased.

The modifications on the maximum MTT depth are enabled only if the palette mode is disabled.

The temporal area is a collocated area from the same reference frame used for the temporal motion predictor.

JVET-AD0147 has memory requirement analysis where the largest memory which is required to store temporal depths for class A is estimated as 40.5 KB.

Test 1.1a: Partitioning prediction (as described above)

Test 1.1b: Test 1.1a where MTT depth increment is enabled unconditionally. It affects classes A and B, class A1 encoder runtime reaches 105.8%.

Test 1.1c: Use partition prediction as encoder only optimization.





Questions:

* Is the partitioning only from the reference frame? A: Yes.
* What is the additional memory requirement? A: Approx. 40 Kbyte per reference picture in 4K.
* Does CABAC parsing require the temporal partitioning info? A; Yes.

It was commented that the memory requirement is significantly less than for motion vectors.

The cross-checkers pointed that the implicit splitting performed at the picture boundary is useless (no gain when removing it, but more complicated).

It was commented that for LB there is no gain (or small losses).

Even though there was no strong objection, from the questions raised the implementation of the idea appears not mature. The gain is relatively low, even though it could be attractive in terms of run time, it introduces additional temporal dependencies that might be undesirable.

No action.

**Category 2: Intra prediction**

**Test 2.1: Block vector guided CCCM (**[**JVET-AE0100**](https://jvet-experts.org/doc_end_user/current_document.php?id=13048)**)**

In this test, co-located luma BV is used to determine the reference area for calculating the CCCM parameters. Then the reference area in luma and corresponding area in chroma channel is used to calculate the CCCM parameters. The prediction uses the calculated model parameters and co-located luma samples to do the CCCM prediction as follows:

predChromaVal = c0C + c1N + c2S + c3E + c4W + c5P(C) + c6P(N) + c7P(S) + c8P(W) + c9P(E)+ c10B,

where spatial components are as shown in the next figure, the nonlinear term P is represented as power of two of the corresponding luma sample, and B is the bias term.

A picture containing text, shoji, crossword puzzle

Description automatically generated

The mode is enabled only in intra slices and is controlled by SPS flag.

Test 2.1a: the mode only uses BVs of the IBC coded blocks from co-located luma area.

Test 2.1b: the mode can use BVs of both IBC and IntraTMP coded blocks from co-located luma area.

**Test 2.2: Bi-predictive GPM (**[**JVET-AE0169**](https://jvet-experts.org/doc_end_user/current_document.php?id=13132)**)**

In ECM-9.0, IBC GPM uses uni prediction for one partition and intra mode for the other one. In the test, the core GPM design is kept unchanged which uses 48 GPM modes and IBC merge candidate list, while it enables both partitions being predicted using IBC with different BVs.

Two flags are signalled to indicate the prediction modes of two partitions, the first flag indicates whether the first partition is intra predicted, and if not then the second flag is signalled to indicate whether intra prediction is used for the second partition.

The method is applied to SCC only.

**Test 2.3: IBC BVP-merge and bi-predictive IBC merge (**[**JVET-AE0169**](https://jvet-experts.org/doc_end_user/current_document.php?id=13132)**)**

In Test 2.3a, IBC-BVP-merge, which is similar to AMVP-merge, derives one BV from IBC block vector prediction (BVP) and the second BV from IBC merge to form bi-prediction for IBC. Two different indices for the IBC BVP and the IBC merge candidates are signalled.

In Test 2.3b, bi-predictive IBC merge is introduced, and it is enabled together with the existed in ECM MBVD and uni-merge (currently disabled by the encoder configuration for non-SCC classes). In bi-predictive IBC merge, two BVs from the existing IBC merge candidate list are derived, utilizing two different indices, which are signalled. Bi-predictive IBC merge is applied to IBC regular merge and IBC MBVD. Bi-predictive IBC merge, IBC MBVD, and IBC uni-merge are enabled for non-SCC classes and is tested together with Test 2.3a.

**Test 2.4: IBC MBVD list derivation (**[**JVET-AE0169**](https://jvet-experts.org/doc_end_user/current_document.php?id=13132)**)**

In the test 2.4a, adaptive BVD offsets along MVBD directions and enabled for IBC MBVD mode. The MBVD candidates search is a two-step process, which starts with checking template SAD costs of offsets added to BVP along each direction with the interval of 1-pel. The second step of the search checks template SAD costs with 1/4-pel interval for the candidates around the selected candidates from the first step. For the integer MBVD (when existed in ECM ph\_fpel\_mbvd\_enabled\_flag is 0), those intervals are multiplied by 4. The candidates with the lowest TM cost are included into the final MBVD list.

Test 2.4c is a combination of Test 2.4a, Test 2.3b, and Test 2.2c.

**Test 2.5: Filtered IBC (**[**JVET-AE0159**](https://jvet-experts.org/doc_end_user/current_document.php?id=13122)**)**

In the test, additional filtered IBC mode is introduced, where a filter is applied to IBC predictor, which is derived by minimizing MSE between current and reference template.

Output of the filter is calculated as follows:

predLumaVal = c0C + c1N + c2S + c3E + c4W + c5P + c6B

The nonlinear term P is represented as power of two of the center sample C and scaled to the sample value range of the content:

P = ( C\*C + midVal ) >> bitDepth

The bias term B represents a scalar offset between the input and output and is set to middle luma value (512 for 10-bit content).

This filtered mode is used as an additional mode for non-merge IBC blocks, and it is not used together with IBC-LIC, IBC-CIIP or RR-IBC. For IBC merge modes, this filtering mode is inherited when merge mode list is constructed.

In Test 2.5a, the mode flag is signalled conditioned on the IBC-LIC flag, so when the IBC-LIC flag is true, the flag is signalled and used to indicate whether the tested mode is applied to the current block or not.

In Test 2.5b, the mode flag is signalled before IBC-LIC flag, and the model does not have non-linear term.

In Test 2.5c, the model of Test 2.5a and signalling of Test 2.5b are used, different aspects of encoder design is chosen from both tests considering the best trade-off.

**Test 2.6: IBC with non-adjacent spatial candidates (**[**JVET-AE0094**](https://jvet-experts.org/doc_end_user/current_document.php?id=13042)**)**

In the test, several new candidates obtained from the BVs of non-adjacent spatial neighboring blocks are added to the candidate lists of IBC merge modes and IBC AMVP. A pattern similar to the non-adjacent spatial candidates of regular inter is used to obtain non-adjacent BVs shown in the below figure. The obtained BVs are inserted between the adjacent spatial candidates and the HBVP candidates for both IBC merge and IBC AMVP. Additionally, the restriction that spatial candidates are not used to construct IBC merge candidate list for a 4×4 CU is removed.

A grid with blue squares and black text

Description automatically generated

In Test 2.6a, the same reference area as for non-adjacent regular merge is reused for the IBC, where up to 4 rows/columns of spatial neighboring blocks (in terms of the CU height/width) of the CU can be used as reference for IBC AMVP, and up to 7 rows/columns of the spatial neighboring blocks (in terms of the CU height/width) of the CU are used for IBC merge.

In Test 2.6b, for both IBC AMVP and IBC merge, the spatial neighboring blocks located up to 4 rows/columns of spatial neighboring blocks (in terms of the CU height/width) of the CU can be used as reference. Additionally, the restriction that spatial candidates are not used for the IBC merge of a 4×4 CU in ECM-9.0 is removed.

Test 2.6c is the combination of Test 2.6a and Test 2.6b, where the sampling method in Test 2.6a is applied. Additionally, the removal of the restriction that spatial candidates cannot be used for IBC merge of a 4x4 CU from Test 2.6b is used.

**Test 2.7: Cross-component merge mode with temporal candidates (**[**JVET-AE0043**](https://jvet-experts.org/doc_end_user/current_document.php?id=12991)**)**

In the current CCP merge mode, a CCP merge candidate list contains spatial adjacent, spatial non-adjacent, and history-based candidates coded in CCLM, MMLM, CCCM, GLM, chroma fusion, and CCP merge modes. After including these candidates, default models can be included to fill the remaining empty positions in the merge list if necessary. To remove redundant CCP models in the merge list, pruning operations are applied. The candidates in the list are reordered based on the SAD costs obtained using the neighboring template of the current block.

In the test, two types of candidates, namely temporal candidates and shifted temporal candidates, are additionally included in the merge list for non-intra slices. Temporal candidates are added to the merge list after the spatial adjacent candidates. Shifted temporal candidates are added after the history-based candidates.

CCP merge mode is allowed to be applied in non-intra slices for chroma blocks with sizes less than or equal to 16.

Temporal candidates

Temporal candidates are selected from the collocated picture, the positions and inclusion order of the temporal candidates, which are the same as those for temporal candidates in inter merge mode of ECM-9.0, is shown in Figure 3.

A screenshot of a game

Description automatically generated

Figure 3. Positions of the temporal candidates

An inclusion order of the temporal candidates is C01 🡪 C02 🡪 … 🡪 C010. If C0i is outside of the picture/slice boundary and C1i is inside of the picture/slice boundary, C1i will be used instead of C0i, where 1 ≤ i ≤ 10. Otherwise, the next inclusion position is checked.

Shifted temporal candidates

Shifted temporal candidates are also selected from the collocated picture. As depicted in Figure 4, the position of the collocated block is shifted by a selected neighboring motion vector. Consequently, the positions of C0i and C1i are shifted by the same neighboring motion vector. The inclusion order of shifted temporal candidates is the same as that of temporal candidates.

The adjacent motion vector is selected from the motion vectors of the neighboring blocks ordered as follows: L0B1 🡪 L1B1 🡪 L0A1 🡪 L1A1 🡪 L0B0 🡪 L1B0 🡪 L0A0 🡪 L1A0 🡪 L0B2 🡪 L1B2. The first motion vector, which uses the collocated picture as the reference picture, is selected. If no such motion vector is found, no shifted temporal candidate is added.

A screenshot of a game

Description automatically generated **A green square with blue text and red arrows

Description automatically generated**

Figure 4. Positions of the shifted candidates (left) and selecting the neighboring motion vector (right)

**Test 2.8: Extrapolation filter-based intra prediction mode (**[**JVET-AE0076**](https://jvet-experts.org/doc_end_user/current_document.php?id=13024)**)**

In the test, extrapolation filter-based intra prediction is introduced with 3 types of reconstructed areas and 3 filter shapes as shown in Figure 5, a choice of reconstructed area and filter shape is signalled.

The size of reconstructed area depends on the min(blockWidth, blockHeight) and the selected filter shape. For example, when the current block is an 8x16 block and the selected filter shape is 4x4. The aboveSize of reconstructed area is min(8, 16) + 4 – 1 = 11, and the leftSize of reconstructed area is min(8, 16) + 4 – 1 = 11.

A black screen with white squares

Description automatically generated

A screenshot of a computer game

Description automatically generated

Figure 5. Three types of reconstructed areas (top) and three filter shapes (bottom)

The selected filter moves in the selected reconstructed area with a one-sample step to collect input samples and output samples, then the filter coefficients are derived using CCCM solver.

In the current block, the recursive predictor is derived from the top-left to the bottom-right position by a diagonal prediction order as shown in Figure 6, where the predicted results of the previous diagonal are used.

A diagram of a graph

Description automatically generated

Figure 6. Example of generating predictions for different positions in the current block by a diagonal order

To reduce the prediction error, the min and max values from the neighboring reconstructed area are applied to restrict the output range of each predicted value.

The predicted samples are calculated as follows,

where is the predicted value at (x, y) in the current block. and offset are derived in the selected reconstructed area, is the coefficient of the derived filter, is reconstructed or predicted value used for prediction in the current position.

The tested mode is applied in I-slices only and is only used for the block sizes less than or equal to 32x32.

**Test 2.9: Extended search areas for IntraTMP mode (**[**JVET-AE0077**](https://jvet-experts.org/doc_end_user/current_document.php?id=13025)**)**

In the test, IntraTMP search areas are extended by including the bottom-left R5 and top-right R6 areas relative to the current block as they may also be available as shown in Figure 7. In ECM-9.0, the region scan order is R4, R1, R2, and R3. In the tested method, the region scan order is R4, R5, R6, R1, R2, and R3.

A diagram of a diagram

Description automatically generated

Figure 7. Extended IntraTMP search areas (R5 and R6)

In addition, the current ECM-9.0 always searches all available positions within the current CTU while it may be beyond the specified search range or width/height. In the test, all search areas are restricted to be within the specified search range or width/height specified by (*searchRangeWidth*, *searchRangeHeight*) as illustrated in Figure 8. The maximum horizontal and vertical search sizes, i.e., *searchRangeWidth* and *searchRangeHeight*, are set to *max(5W, 64*) and *max(5H, 64)* respectively. It is not allowed to search any positions outside the specified search range even those available positions are within the current CTU.

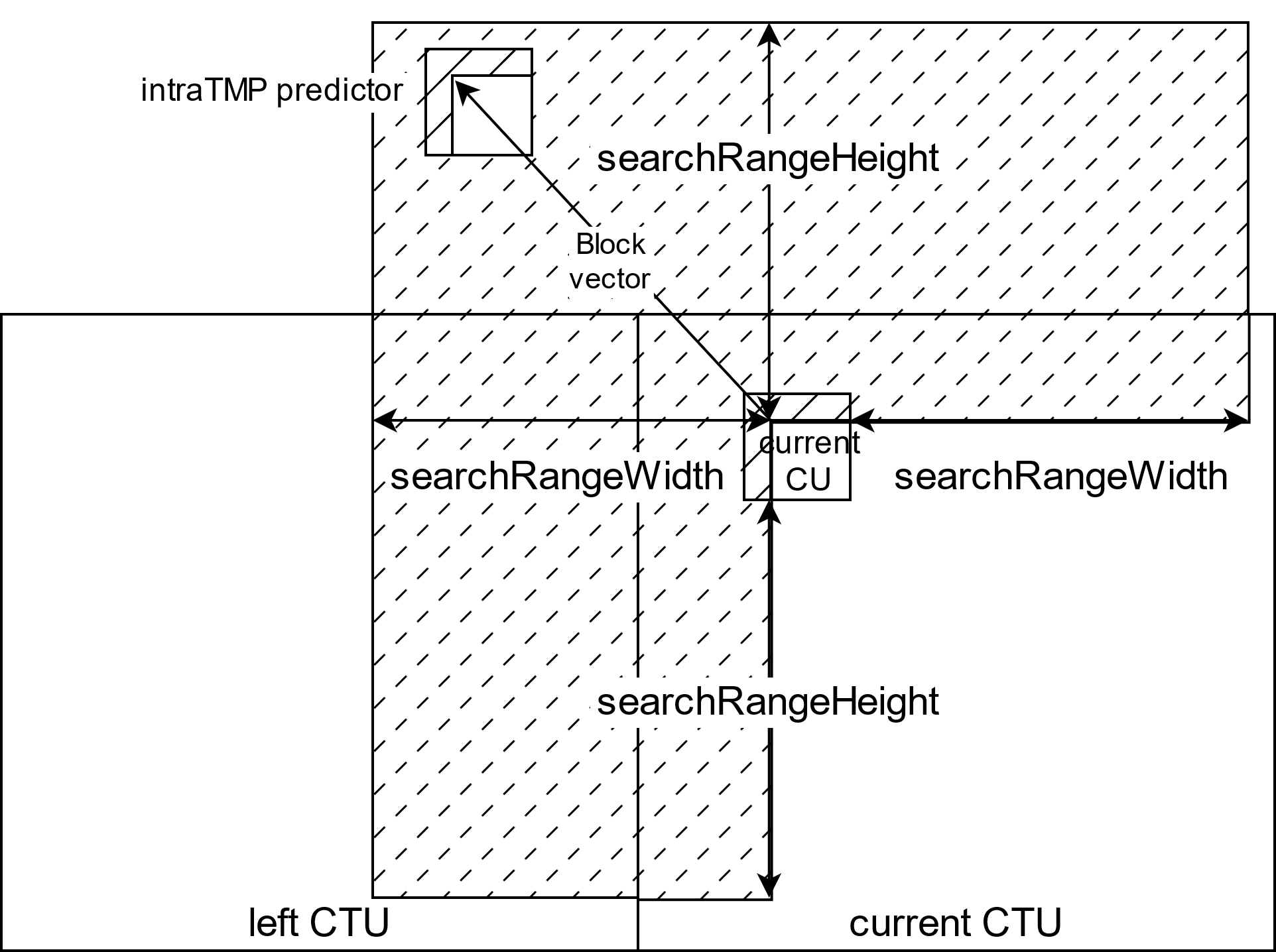


Figure 8. Search areas restriction

**Test 2.10: IBC-LIC extension (**[**JVET-AE0078**](https://jvet-experts.org/doc_end_user/current_document.php?id=13026)**)**

In Test 2.10a, IBC-LIC is extended by including 3 additional modes:

* Use only left or use only above LIC template in addition to the current L-shape template,
* Extend the MMLM to IBC-LIC, which allows IBC-LIC to have two linear models in one CU for L-shape template.

Large block-size constraint for ECM IBC-LIC, tested separately in Test 2.10b, is additionally removed in Test 2.10a.In Test 2.10b, large block-size constraint for ECM IBC-LIC is removed, so IBC-LIC can be applied to the CU whose block size is larger than or equal to 32.

**Test 2.11: Harmonization of IBC HMVP and IBC-LIC (**[**JVET-AE0084**](https://jvet-experts.org/doc_end_user/current_document.php?id=13032)**)**

In ECM-9.0, the inter LIC flag can be inherited from a HMVP candidate. However, the IBC-LIC flag is not inherited when the motion information from an IBC HMVP candidate is used.

In Test 2.11a, IBC-LIC flag is inherited from an IBC HMVP candidate, which is similar to the inter LIC case.

Test 2.11b: Test 2.11a + Test 2.10a

Results for AI/RA



Test 2.1 is a chroma tool, main benefit on screen content. The gain is small, but it has no impact on run time, and implementation is straightforward according to cross-checkers. Adoption was also supported by independent experts.

Decision: Adopt JVET-AE0100, test 2.1b

Tests 2.2..2.4 target several aspects of improving IBC in context with GPM and bi-prediction, 2.3b also provides gain on camera-captured content (but had been reported previously to not give gain in case of screen content). In the combination 2.4c (2.2c+2.3b+2.4a), 2.3b is disabled for screen content, and 2.2c is disabled for camera-captured content, whereas 2.4a is enabled for both. gives best gain overall, including screen content. However, the results show that 2.4a does not have benefit for camera capured content when combined with 2.3b, whereas for screen content the gain from 2.2c and 2.4a is somewhat additive.

Decision: Adopt JVET-AE0169 Test 2.2c, in CTC enabled only for screen content

Decision: Adopt JVET-AE0169 Test 2.3b, in CTC enabled only for camera-captured content

Decision: Adopt JVET-AE0169 Test 2.4a, in CTC enabled only for screen content

For 2.5c, a discussion was performed about the benefit of a combination test when the individual results of the elements being combined are not available. Proponents are strongly discouraged to withdraw proposal elements when still including them in a combination, as this will not allow to judge the individual benefits.

Individual results of 2.5a and 2.5b were presented in an update of the summary report on Wednesday July 12.

Gains appear to be somewhat additive, but encoding run time is significantly higher in 2.5b than it is in 2.5c. According to proponents, an encoder with less RDO checks was used in 2.5c. This is also the likely reason for losing a slight bit rate reduction in 2.5c, compared to 2.5b.

It was asked whether the encoder search strategy of 2.5c is the same as 2.5a, such that it could be confirmed that the additional benefit of 2.5c comes by the signalling method of 2.5b, and its non-linear filter term? According to proponents, main gain comes from signalling.

Results were matched according to cross-checker (verbally reported in session, document to be uploaded).

Decision: Adopt JVET-AE0159 Test 2.5c (enabled only for screen content)

Same situation was initially found with 2.6c.I Individual results of 2.6a and 2.6b were presented in an update of the summary report on Wednesday July 12.

Results indicate that the combination of 2.6c provides more gain than the individual parts 2.6a and 2.6b. Gains cannot be expected to be additive, as the basic approach is identical. It can however be concluded that the element of 2.6b (removing the restriction on 4x4 blocks) is also beneficial in combination with 2.6a. The proposal was supported by independent experts (including cross-checkers).

Decision: Adopt JVET-AE0094 Test 2.6c (enabled only for screen content)

Test 2.7 is targeting improvement of chroma, by adding additional temporal candidates in cross-component merge (which had been adopted in the last meeting). Candidates are stored in an 8x8 grid.

It was asked what the benefit of shifted candidates is. According to proponents, it had been 10% of the total gain in the original contribution. It was argued that this element of the proposal might be difficult to implement in hardware. It was however asserted by JVET that this is not of highest important in this stage of exploration, and technology investigated in EE should normally be adopted as proposed.

Decision: Adopt JVET-AE0043 test 2.7.

Test 2.8: It was asked how the mode is used when reconstructed area is not available. It is disabled. It was further pointed out that the gain seems to be very small for small blocks, where the pixel dependency might be critical in real-world implementation. The cross-checker confirms that such dependency is probably present, but would nevertheless support this tools at this stage of exploration.

2.5% increase of encoder run time is pointed out to be not a good tradeoff for 0.16% bit rate reduction in AI. For RA, the gain is less, as it is only applied in I slices currently.

Further study recommended (in EE) to reduce encoder run time, and also apply in inter slices.

Reduction of pixel dependency also would be important, but probably is more difficult to achieve.

Test 2.9: Increase in decoder run time by 2.5% by additional search operations. Encoder run time also slightly increased. One cross-checker points out that the increase of decoder run time may be caused by pixel-wise availability checks. It was also pointed out that this is straightforward extension of search range relative to ECM 9, and the tradeoff between gain and run time is still in an attractive range.

Decision: Adopt JVET-AE 0077 test 2.9

Tests 2.10/2.11 are only beneficial for screen content. From comparing 2.10a/b, it cannot be concluded that removing the large block size constraint is beneficial, as there are no results for the IBC-LIC extension with the large block size constraint of current ECM (such a test had originally been planned, but was not conducted). Proponents are requested to present results of a variant of 2.10a with the constraint, and also in combination 2.11.b to assess whether the gain is additive with 2.11a. Revisit.

Test 2.11a is a minor change which gives some benefit in class F without impact on run time. However, it should also be enabled for camera-captured content to avoid another high-level flag (according to proponents, no losses would be expected). Therefore, the new combination test 2.11b should be conducted such that the method 2.11a is also used for camera-catured content. Revisit.

**Inter prediction**

**Test 3.1: Cross-component residual model for inter prediction (**[**JVET-AE0059**](https://jvet-experts.org/doc_end_user/current_document.php?id=13007)**)**

Cross-component residual model to predict chroma samples from reconstructed luma samples when the block uses inter prediction or intra block copy (IBC) is tested. As illustrated in Figure 9, the cross-component filters are derived using the prediction blocks of luma and chroma. The derived filters are applied to the reconstructed luma block and blended with the prediction blocks of chroma to produce the final chroma prediction blocks. In the blending process the filtered reconstructed luma blocks use blending weight of 0.75 and chroma prediction blocks use blending weight of 0.25.

A diagram of a computer

Description automatically generated

Figure 9. Cross-component residual model

Model uses 8-tap filter consisting of 6 spatial luma samples shown in Figure 10, a nonlinear term, and a bias term as follows:

predChromaVal = c0 L0+ c1L1 + c2L2 + c3L3 + c4L4 + c5L5 + c6 nonlinear((L0+L3+1) >> 1) + c7 B,

A white grid with black letters and numbers

Description automatically generated

Figure 10. 2 Luma samples L0,..,L5 in relation to the chroma sample C.

The filter coefficients are derived using ECM’s division-free Gaussian elimination method and the necessary offsets are applied to samples prior to filter derivation.

For filter coefficient derivation at most 256 chroma samples are used.

The mode flag is only signalled if the TU’s luma Cbf is non-zero and the CU’s predMode is either MODE\_INTER or MODE\_IBC.

Test 3.1a: Cross-component residual model, where offsets for division free operation are obtained by averaging the samples in the block.

Test 3.1b: Cross-component residual model, where offsets for division free operation are obtained by averaging the four points correspond to the top-left, top-right, bottom-left and bottom-right corners of the blocks.

**Test 3.2: Bi-predictive GPM (**[**JVET-AE0046**](https://jvet-experts.org/doc_end_user/current_document.php?id=12994)**)**

In the test, the existing uni-predictive GPM is extended to allow usage of bi-predictive motion vectors for generating motion compensated prediction samples for inter partitions. The method consists of the following elements:

The first element conditionally invokes the existing extraction process that extracts uni-predictive motion vectors from the initial list. The extraction process is invoked only for small blocks 8x8, 16x8 and 8x16. For other larger blocks, the extraction process is bypassed, so the initial list (which may contain merged Bi-MVs) is directly used as the final GPM merge list. The generation of the initial list is the same as before (i.e., the normal merge list generation without any candidate reordering) except that when generating the initial list for larger blocks (i.e., blocks with the extraction process bypassed), the motion vector difference threshold for controlling whether a candidate can be added into the initial list is increased to be one full sample distance.

The second element modifies GPM-MMVD to support bi-predictive motion vector as the base vector. For low-delay pictures, the signalled MVD is applied on top of the L0 and L1 motion vector as in the existing merge MMVD design. For non-low-delay pictures, the bi-predictive motion vector is converted into a uni-predictive motion vector first and then the MVD is applied on top.

The third element modifies GPM-TM to also support bi-predictive motion vectors.

The last element is to enable the BDOF based mv refinement as in the multi-pass DMVR on top of the associated bi-predictive motion vectors for each inter partition.

**Test 3.3: High-Accuracy template matching (**[**JVET-AE0087**](https://jvet-experts.org/doc_end_user/current_document.php?id=13035)**)**

In ECM, template matching is a decoder-side MV derivation method to refine the motion information of the current CU by finding the closest match between a template (i.e., top and/or left neighbouring blocks of the current CU) in the current picture and a block (i.e., same size to the template) in a reference picture.

Diagram

Description automatically generated with medium confidence

When TM is used for bi-prediction, the following steps are applied:

1. The initial motion vector of list 0 (MV0) is refined to derive a refined MV (MV’0) and a TM cost C0 is calculated;
2. The initial motion vector of list 1 (MV1) is refined to derive a refined MV (MV’1) and a TM cost C1 is calculated;
3. If C0 is larger than C1, MV’1 is used to derive a further refined MV (MV’’0) by refining MV’0; Otherwise, MV’0 is used to derive a further refined MV (MV’’1) by refining MV’1 .

In high-accuracy template matching, three aspects are tested:

* In Test 3.3a, an additional refinement is applied to TM for bi-prediction, when MV’0 is refined in step 3), MV’’0 is used to derive MV’’1 by refining MV’1; Otherwise, MV’’1 is used to derive MV’’0 by refining MV’0.
* In Test 3.3b, the diamond search pattern used in TM is modified from 8-point to 16-point.

图标

描述已自动生成

* In Test 3.3c, TM for bi-prediction is enabled when DMVR condition is satisfied.

Test 3.3d: Test 3.3a + Test 3.3c.

Test 3.3e: Test 3.3a + Test 3.3b + Test 3.3c.

**Test 3.4: OBMC with HPel flag and BCW weights (**[**JVET-AE0196**](https://jvet-experts.org/doc_end_user/current_document.php?id=13159)**)**

In the current OBMC design, the motion compensation of the neighboring blocks and subblocks performed by OBMC only considers the motion vectors and reference pictures.

In the test, HPel flags and BCW weights of the neighboring blocks and subblocks are considered in the OBMC. When processing the top and left borders, the neighboring BCW weights and HPel flags are used in addition to the motion vectors and reference pictures. The neighboring HPel flag allows setting the correct AMVR index. For the internal subblock boundaries, the BCW and AMVR indexes of the current CU are also used.

It has to be noticed that the integration of adopted contribution JVET-AD0213 (Bi-prediction LIC) turns out to already include main aspect of this test (BCW weights in OBMC) in the ECM-9.0 which was not the case in ECM-8.0.

**Test 3.5: Iterative BDOF pass in multi-pass DMVR (**[**JVET-AE0065**](https://jvet-experts.org/doc_end_user/current_document.php?id=13013)**)**

In ECM, a multi-pass DMVR has three passes of motion refinement. In the first pass, bilateral matching (BM) is applied to the coding block (CB) to get the block level refined MV pair. In the second pass, the block is divided into 16×16 subblocks and on top of block level refined MV pair obtained in the first pass, BM is applied on each 16×16 subblock to get a 16×16 subblock level refined MV pair. In the third pass, BDOF based MV refinement is applied on the 4×4 or 8×8 subblocks depending on the CB size.

In the test, the current multi-pass DMVR is extended by adding another BDOF-based motion refinement pass as the 4th pass of the multi-pass DMVR (i.e., the 2nd round of BDOF-based DMVR).

Also, the subblock size of the proposed round of the BDOF-based DMVR is adaptively selected depending on the area (i.e., width×height) of the coding block. For blocks smaller than 1024 pixels, the subblock size of 4×4, and otherwise 8×8 is used.

**Test 3.7: RPR with new filters for scale factor 1.25x and 1.33x (**[**JVET-AE0150**](https://jvet-experts.org/doc_end_user/current_document.php?id=13113)**)**

In the test, additional set of 8-bit precision RPR filters consisting of three tables (respectively 12-tap for luma, 6-tap for chroma, and 10-tap for affine) is introduced with the target to improve the performance for scaling ratios 1.25x and 1.33x. These new set of filters is used for scaling ratios in between 1.1x and 1.35x.

In the tests, two PSNR values are used:

* PSNR1: it is measured on the decoded picture size by calculating MSE for each picture, accumulating MSEs for all pictures and calculating average PSNR from the accumulated MSE.
* PSNR2: it is measured after upsampling of the decoded picture if the decoded picture size is in smaller resolution, the existed ECM interpolation filters are used for the upsampling.

**Test 3.8: Combination of Test 3.3e and Test 3.5 (**[**JVET-AE0091**](https://jvet-experts.org/doc_end_user/current_document.php?id=13039)**)**

Results for RA/LB



Test 3.1: Increase of encoder run time should be expected in test 3.1b (RA results not complete yet).

It was commented that the method should better be called “cross component chroma prediction in inter”, as it is not the residual that is predicted.

It was also commented that a blending with 0.75/0.25 weights was newly introduced, which according to proponents effects better performance in blending. It was requested to provide quantitative information about the gain by blending

Results were presented Wednesday 8:30. Test 3.1b has slightly less encoder run time (which is explainable by having less checks), and bit rate reduction is almost identical. This was also confirmed by cross-checker.

It was asked why a different cross-component model was used than the one in ECM. It was answered by proponent that downsampling is avoided. The unification might be desirable, but is not of prior importance at this moment.

Decision: Adopt JVET-AE0059 Test 3.1b.

Test 3.2: 0.2% bit rate reduction in RA with slight increase in encoding time. According to one cross-checker it may increase the worst-case memory access, which may however not be too severe, as the proposal is not used in case of small block size, and multi-hypothesis is also used in other tools of ECM. Several experts (including cross-checkers) supported the proposal. It was however suggested to modify the software such that the memory for motion vector storage is reduced in the encoder.

Decision: Adopt JVET-AE0046 Test 3.2.

Test 3.3x, 3.5, and combination 3.8: Test 3.3c seems to have no benefit standalone, but according to proponents has some benefit in combination with 3.3a. In combination, gains of 3.3e and 3.5 are somewhat additive in RA, whereas only 3.3e contributes gain LB (BDOF of 3.5 not applicable in LB). Several experts supported the combination, and no objection was raised.

Decision: Adopt JVET-AE0091 Test 3.8

Test 3.4 does not show benefit in compression, might be due to new adoptions in ECM 9. No action.

Test 3.7 is about RPR (non-CTC test conditions deduced from RPR specific VTM test conditions which only used 1.5x and 2x). The results indicate that improvement is possible by using different filters for small factors of resolution change. Though it is not of prior importance for the exploration, it may be interesting to provide such implementation in the software for experimentation of interested parties.

Decision(SW): Adopt JVET-AE0150 (not in CTC, not in ECM description)

***Transforms and coefficient coding***

**Test 4.1: Shifting quantizer center (**[**JVET-AE0125**](https://jvet-experts.org/doc_end_user/current_document.php?id=13079)**)**

In the test, a quantization offset is added to the quantized level, the offset is quantization index dependent, a look-up table is used to derive the offset .

where x is the dequantized coefficient, y and y’ are the quantization indices, Q-1 is the dequantization operation, T is a look-up table

.

**Test 4.2: Large NSPT (**[**JVET-AE0086**](https://jvet-experts.org/doc_end_user/current_document.php?id=13034)**)**

The large NSPT kernels are tested for 4x32/32x4 and 8x32/32x8 blocks, of which kernel matrix dimensions are 20x128 and 24x256. Therefore, 20 and 24 transform coefficients are generated by applying the two types of kernel matrices, respectively, which are placed from DC position following scan order. The remaining 108 and 232 positions in each transform block are zeroed-out, respectively.

Large NSPT kernels are applied in the same way as for other block sizes 4x4, 4x8/8x4, 4x16/16x4, 8x8, and 8x16/16x8 and those NSPT kernels are not changed.

**Test 4.3: Context modeling for transform coefficients for LFNST/NSPT (**[**JVET-AE0102**](https://jvet-experts.org/doc_end_user/current_document.php?id=13050)**)**

In ECM, the causal 2D neighborhood of coefficients is used to model the context to parse the sig\_coeff\_flag, gt1\_flag, gt2\_flag as shown in Figure 11. For LFNST, DCT-II coefficients are placed into the coefficient block using diagonal reordering.

A number on a black background

Description automatically generated

Figure 11. Context modeling for LFNST coefficients in ECM

In the test, when LFNST/NSPT is applied, the previous 5 coefficients in the coding order are used for context derivation instead of 2D neighborhood.

A number on a black background

Description automatically generated

Figure 12. Context modeling for LFNST coefficients in TEST4.3

In ECM, the lfnstIdx is signalled after all the coefficients in a CU. In the test, lfnstIdx is signalled after all last\_sig\_coeff\_pos syntax elements in a CU since lfsntIdx is required for parsing the transform coefficients.

**Test 4.4: InterMTS for IBC and IntraTMP (**[**JVET-AE0116**](https://jvet-experts.org/doc_end_user/current_document.php?id=13064)**)**

InterMTS enabled for IBC and IntraTMP is tested. The following tests are performed:

Test 4.4a: InterMTS is enabled for IBC-coded blocks in AMVP mode. MTS is not enabled for IBC merge.

Test 4.4b: InterMTS is enabled for IBC-coded blocks in AMVP mode as in Test 4.4a. In addition, IntraTMP is utilizing InterMTS instead of IntraMTS. Consequently, the number of MTS candidates for IntraTMP has been fixed to 4 instead of adaptively setting it to 1, 4 or 6 as in IntraMTS.

Test 4.4c: IntraMTS is disabled for IntraTMP coded blocks.

Results for AI/RA (LFNST not enabled in LB)



Test 4.1: Shifting quantizer center provides 0.1% gain without any complexity impact.

Decision: Adopt JVET-AE0125 Test 4.1

Test 4.2: Reasonable tradeoff runtime vs. bit rate reduction, additional memory consumption of larger kernels not relevant in cntext of exploration.

Decision: Adopt JVET-AE0086 Test 4.2

Test 4.3: It was commented that the usage of immediately preceding coefficients might have some latency impact (VVC did not use coefficients from same diagonal for that reason), but for the purpose of exploration several experts (including cross-checkers) supported the proposal.

Decision: Adopt JVET-AE0102 Test 4.3

Test 4.4x: Enabling MTS for IBC does not provide relevant gain to justify the increase of encoding time (caused by additional RDO checks). No action.

***Loop filtering***

**Test 5.1: CCSAO with extended edge classifiers and history offsets (**[**JVET-AE0151**](https://jvet-experts.org/doc_end_user/current_document.php?id=13114)**)**

In ECM-9.0, CCSAO uses band and edge classifiers switched at CTB level. Classifier parameters/offsets are derived at encoder and signaled for each slice independently.

In Test 5.1a, to reduce the signaling overhead, CCSAO inheritance scheme is introduced, where the offsets/parameters of some coded pictures are stored at both encoder and decoder which are allowed to be used as the CCSAO offsets/classifiers of future pictures. An index is signaled in SH to indicate which candidate in the CCSAO parameter storage is selected for the current slice. The candidates in the CCSAO parameter storage are updated in a FIFO manner and refreshed at IDR pictures.

In Test 5.1b, on top of Test 5.1a, beside the existing CCSAO edge classification, one new edge classification is added, which is a subset of the original one with less edge range divisions.

where is calculated by comparing a sample difference in one direction with a threshold .

Additionally, the component used for edge classification can be selected from one of all three components. Same to the existing CCSAO design, the selected edge classifier and edge component are decided by encoder and signaled in SH.

**Test 5.2: Improved fixed filters for ALF (**[**JVET-AE0139**](https://jvet-experts.org/doc_end_user/current_document.php?id=13102)**)**

In ECM-9.0, two Laplacian-based classifiers (one for each fixed filter) are applied to a 2x2 block. In each classifier, activity and directionality values are derived based on vertical, horizontal, and diagonal gradients using a window surrounding each 2x2 block. Then a class index is determined based on the activity and directionality values. Two 13x13 diamond shaped fixed filters are selected from the two filter sets by using the derived two class indices. Then the two fixed filters ( with ) are applied to the ALF input samples. Finally, a signalled filter is applied to the ALF input samples, samples before the deblocking filter (DBF), outputs of the two fixed filters, output of a gaussian filter and the residual data.

In Test 5.2a, both fixed filters are applied to samples before DBF and ALF input, where additional diamond 9x9 filter is used for the samples before DBF. The shape of the first fixed filter applied to the ALF input samples is reduced from 13x13 to 9x9, and the shape of the second fixed filter, which is 13x13, applied to ALF input is unchanged as shown in the below table.

|  |  |  |  |
| --- | --- | --- | --- |
|  | ECM-9.0 | Proposed method | |
|  | ALF input | Samples before DBF | ALF input |
| Fixed filer | 13x13 | 9x9 | 9x9 |
| Fixed filer | 13x13 | 9x9 | 13x13 |

In Test 5.2b, on top of Test 5.2a, the classifiers of the fixed filters are extended. For each 2x2 block, the mean value of a surrounding window is calculated. Then, for each sample of this window, the difference between the sample value and the mean value is calculated. A scaling factor is determined based on the activity value derived from a Laplacian classifier. The square root of the sum of the squared differences is further quantized to by a scaling factor. The value of is an integer between 0 and 7, inclusively. With *i*=0, 1, let denote the classifier from the classifier of i-th fixed filter in ECM-9.0. Then the proposed class index is derived as

.

The total number of the fixed filters is not changed.

In Test 5.2c, on top of Test 5.2b, fixed filter is applied to outputs of (instead of ALF input) and samples before DBF.

**Test 5.3: Combination of Test 5.1b and Test 5.2c (**[**JVET-AE0152**](https://jvet-experts.org/doc_end_user/current_document.php?id=13115)**)**



One expert commented that the changes in 5.1a and 5.1b are straightforward and supported it.

It was report that an additional buffer of approcimately 5Kbyte is necessary for storage.

It was confirmed that the data flow in 5.2x is not changed relative to the current ECM.

It can be expected that the gains of CCSAO modifications and ALF modifications are non-overlapping, and it was also reported that preliminary results of 5.3 indicate additive gains.

Decision: Adopt JVET-AE0151 Test 5.1b

Decision: Adopt JVET-AE0139 Test 5.2c

It was commented that the results of 5.3 (combination test) are expected to be provided for information.

### EE2 contributions: Enhanced compression beyond VVC capability (24)

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-AE0043](https://jvet-experts.org/doc_end_user/current_document.php?id=12991) EE2-2.7: Cross-component merge mode with temporal candidates [H.-Y. Tseng, C.-M. Tsai, C.-Y. Chuang, C.-W. Hsu, C.-Y. Chen, T.-D. Chuang, O. Chubach, Y.-W. Chen, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-AE0220](https://jvet-experts.org/doc_end_user/current_document.php?id=13183) Crosscheck of JVET-AE0043 (EE2-2.7: Cross-component merge mode with temporal candidates) [P. Bordes (InterDigital)] [late]

[JVET-AE0046](https://jvet-experts.org/doc_end_user/current_document.php?id=12994) EE2-3.2: Bi-predictive GPM [R. Yu, P. Wennersten, J. Enhorn, K. Andersson (Ericsson)]

[JVET-AE0199](https://jvet-experts.org/doc_end_user/current_document.php?id=13162) Crosscheck of JVET-AE0046 (EE2-3.2: Bi-predictive GPM) [Y. Ahn (LGE)] [late]

[JVET-AE0249](https://jvet-experts.org/doc_end_user/current_document.php?id=13212) Crosscheck of JVET-AE0046 (EE2-3.2: Bi-predictive GPM) [Y. Kidani, K. Kawamura (KDDI)] [late]

[JVET-AE0059](https://jvet-experts.org/doc_end_user/current_document.php?id=13007) EE2-3.1: Cross-component residual model (CCRM) for inter prediction [P. Astola, J. Lainema (Nokia)]

[JVET-AE0236](https://jvet-experts.org/doc_end_user/current_document.php?id=13199) Cross-check of JVET-AE0059 “EE2-3.1: Cross-component residual model (CCRM) for inter prediction” [F. Le Léannec (InterDigital)] [late]

[JVET-AE0253](https://jvet-experts.org/doc_end_user/current_document.php?id=13216) Cross-check of JVET-AE0059 (EE2-3.1: Cross-component residual model (CCRM) for inter prediction) [J. Arumugam (Ittiam)] [late]

[JVET-AE0065](https://jvet-experts.org/doc_end_user/current_document.php?id=13013) EE2-3.5: Iterative BDOF pass in multi-pass DMVR [M. Salehifar, Y. He, K. Zhang, H. Liu, L. Zhang (Bytedance), J. Chen, R.-L. Liao, X. Li, Y. Ye (Alibaba)]

[JVET-AE0200](https://jvet-experts.org/doc_end_user/current_document.php?id=13163) Crosscheck of JVET-AE0065 (EE2-3.5: Iterative BDOF pass in multi-pass DMVR) [Z. Lv (vivo)] [late]

[JVET-AE0076](https://jvet-experts.org/doc_end_user/current_document.php?id=13024) EE2-2.8: An extrapolation filter-based intra prediction mode [L. Xu, Y. Yu, H. Yu, D. Wang (OPPO)]

[JVET-AE0217](https://jvet-experts.org/doc_end_user/current_document.php?id=13180) Crosscheck of JVET-AE0076 (EE2-2.8: An extrapolation filter-based intra prediction mode) [H.-J. Jhu (Kwai)] [late]

[JVET-AE0077](https://jvet-experts.org/doc_end_user/current_document.php?id=13025) EE2-2.9: Extended search areas for IntraTMP mode [Y. Yu, L. Zhang, F. Wang, J. Gan, H. Yu, L. Xu, Z. Xie, D. Wang (OPPO), Y. Ma, H. Zhang, J. Huo, F. Yang, X. Gao (Xidian University), X. Xiu, N. Yan, C. Ma, H. Jhu, C. Kuo, W. Chen (Kwai)]

[JVET-AE0210](https://jvet-experts.org/doc_end_user/current_document.php?id=13173) crosscheck of JVET-AE0077 EE2-2.9: Extended search areas for IntraTMP mode [K. Naser (InterDigital)] [late]

[JVET-AE0216](https://jvet-experts.org/doc_end_user/current_document.php?id=13179) Crosscheck of JVET-AE0077 (EE2-2.9: Extended search areas for IntraTMP mode) [P.-H Lin (Qualcomm)] [late]

[JVET-AE0245](https://jvet-experts.org/doc_end_user/current_document.php?id=13208) Crosscheck of JVET-AE0077 (EE2-2.9: Extended search areas for IntraTMP mode) [X. Li (Alibaba)] [late]

[JVET-AE0078](https://jvet-experts.org/doc_end_user/current_document.php?id=13026) EE2-2.10: IBC-LIC extension [Z. Xie, Y. Yu, H. Yu, D. Wang (OPPO)]

[JVET-AE0157](https://jvet-experts.org/doc_end_user/current_document.php?id=13120) Crosscheck of JVET-AE0078(EE2-2.10: IBC-LIC extension) [W. Yin (Bytedance)] [late]

[JVET-AE0084](https://jvet-experts.org/doc_end_user/current_document.php?id=13032) EE2-2.11: Harmonization of IBC HMVP and IBC-LIC [N. Zhang, K. Zhang, L. Zhang (Bytedance), Z. Xie, Y. Yu, H. Yu, D. Wang (OPPO)]

[JVET-AE0222](https://jvet-experts.org/doc_end_user/current_document.php?id=13185) Crosscheck of JVET-AE0084 (EE2-2.11: Harmonization of IBC HMVP and IBC-LIC) [C. Ma (Kwai)] [late]

[JVET-AE0086](https://jvet-experts.org/doc_end_user/current_document.php?id=13034) EE2-4.2: Large NSPT [M. Koo, J. Zhao, J. Lim, S. Kim (LGE)]

[JVET-AE0118](https://jvet-experts.org/doc_end_user/current_document.php?id=13066) Crosscheck of JVET-AE0086 (EE2-4.2: Large NSPT) [P. Garus (Qualcomm)] [late]

[JVET-AE0087](https://jvet-experts.org/doc_end_user/current_document.php?id=13035) EE2-3.3: High-Accuracy template matching [Y. Wang, K. Zhang, L. Zhang (Bytedance)]

[JVET-AE0239](https://jvet-experts.org/doc_end_user/current_document.php?id=13202) Crosscheck of JVET-AE0087 (EE2-3.3: High-Accuracy template matching) [J. Chen (Alibaba)] [late] [miss]

[JVET-AE0091](https://jvet-experts.org/doc_end_user/current_document.php?id=13039) EE2-3.8: Combination of Test 3.3e and Test 3.5 [M. Salehifar, Y. Wang, Y. He, K. Zhang, H. Liu, L. Zhang (Bytedance), J. Chen, R.-L. Liao, X. Li, Y. Ye (Alibaba)]

[JVET-AE0223](https://jvet-experts.org/doc_end_user/current_document.php?id=13186) Crosscheck of JVET-AE0091 (EE2-3.8: Combination of Test 3.3e and Test 3.5) [C. Ma (Kwai)] [late]

[JVET-AE0094](https://jvet-experts.org/doc_end_user/current_document.php?id=13042) EE2-2.6: IBC with non-adjacent spatial candidates [C. Ma, X. Xiu, W. Chen, H.-H. Jhu, C.-W. Kuo, N. Yan, X. Wang (Kwai), Y. Wang, K. Zhang, L. Zhang (Bytedance)]

[JVET-AE0240](https://jvet-experts.org/doc_end_user/current_document.php?id=13203) Crosscheck of JVET-AE0094 (EE2-2.6c: IBC with non-adjacent spatial candidates) [J. Chen (Alibaba)] [late] [miss]

[JVET-AE0100](https://jvet-experts.org/doc_end_user/current_document.php?id=13048) EE2-2.1: Block vector guided CCCM [R. G. Youvalari, D. Bugdayci Sansli, P. Astola, J. Lainema (Nokia)]

[JVET-AE0226](https://jvet-experts.org/doc_end_user/current_document.php?id=13189) Crosscheck of JVET-AE0100 (EE2-2.1: Block vector guided CCCM) [M. Abdoli (IRT b-com)] [late]

[JVET-AE0102](https://jvet-experts.org/doc_end_user/current_document.php?id=13050) EE2-4.3: Context modeling for transform coefficients for LFNST/NSPT [P. Nikitin, M. Coban, C. S. Coban, B. Ray, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-AE0225](https://jvet-experts.org/doc_end_user/current_document.php?id=13188) Crosscheck of JVET-AE0102 (EE2-4.3: Context modelling for transform coefficients for LFNST/NSPT) [M. Abdoli (IRT b-com)] [late]

[JVET-AE0116](https://jvet-experts.org/doc_end_user/current_document.php?id=13064) EE2-4.4: InterMTS for IBC and IntraTMP [P. Garus, M. Coban, B. Ray, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-AE0211](https://jvet-experts.org/doc_end_user/current_document.php?id=13174) crosscheck of JVET-AE0116 EE2-4.4: InterMTS for IBC and IntraTMP [K. Naser (InterDigital)] [late]

[JVET-AE0125](https://jvet-experts.org/doc_end_user/current_document.php?id=13079) EE2-4.1: Shifting quantization center [M. Balcilar, K. Naser, F. Galpin, F. Le Léannec (InterDigital)]

[JVET-AE0149](https://jvet-experts.org/doc_end_user/current_document.php?id=13112) Crosscheck of JVET-AE0125 (EE2-4.1 Shifting Quantizer Center) [M. Coban (Qualcomm)] [late]

[JVET-AE0132](https://jvet-experts.org/doc_end_user/current_document.php?id=13095) EE2-1.1: Partitioning prediction [G. Laroche, P. Onno (Canon)]

[JVET-AE0235](https://jvet-experts.org/doc_end_user/current_document.php?id=13198) Cross-check of JVET-AE0132 “EE2-1.1: Partitioning prediction” [R. Utida, F. Le Léannec (InterDigital)] [late]

[JVET-AE0139](https://jvet-experts.org/doc_end_user/current_document.php?id=13102) EE2-5.2: Improved fixed filters for ALF [M. Karczewicz, N. Hu, H. Wang, V. Seregin (Qualcomm)]

[JVET-AE0193](https://jvet-experts.org/doc_end_user/current_document.php?id=13156) Crosscheck of JVET-AE0139 (EE2-5.2: Improved fixed filters for ALF) [C.-W. Kuo (Kwai)] [late]

[JVET-AE0150](https://jvet-experts.org/doc_end_user/current_document.php?id=13113) EE2-3.7 RPR with new filters for scale factor 1.25x and 1.33x [J. Samuelsson-Allendes, S. Deshpande (Sharp)]

[JVET-AE0204](https://jvet-experts.org/doc_end_user/current_document.php?id=13167) Crosscheck of JVET-AE0150 (EE2-3.7 RPR with new filters for scale factor 1.25x and 1.33x) [Z. Zhang (Qualcomm)] [late]

[JVET-AE0151](https://jvet-experts.org/doc_end_user/current_document.php?id=13114) EE2-5.1: CCSAO with extended edge classifiers and history offsets [C.-W. Kuo, X. Xiu, W. Chen, H.-J. Jhu, N. Yan, C. Ma, X. Wang (Kwai)]

[JVET-AE0209](https://jvet-experts.org/doc_end_user/current_document.php?id=13172) Crosscheck of JVET-AE0151 (EE2-5.1: CCSAO with extended edge classifiers and history offsets) [N. Hu (Qualcomm)] [late]

[JVET-AE0152](https://jvet-experts.org/doc_end_user/current_document.php?id=13115) EE2-5.3: Combination of Test 5.1b and Test 5.2c [C.-W. Kuo, X. Xiu, W. Chen, H.-J. Jhu, N. Yan, C. Ma, X. Wang (Kwai), M. Karczewicz, N. Hu, H. Wang, V. Seregin (Qualcomm)]

[JVET-AE0241](https://jvet-experts.org/doc_end_user/current_document.php?id=13204) Crosscheck of JVET-AE0152 (EE2-5.3: Combination of Test 5.1b and Test 5.2c) [J. Chen (Alibaba)] [late] [miss]

[JVET-AE0159](https://jvet-experts.org/doc_end_user/current_document.php?id=13122) EE2-2.5: Filtered Intra Block Copy (FIBC) [H.-J. Jhu, X. Xiu, C.-W. Kuo, W. Chen, N. Yan, C. Ma, X. Wang (Kwai), B. Ray, M. Coban, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-AE0228](https://jvet-experts.org/doc_end_user/current_document.php?id=13191) Crosscheck of JVET-AE0159 (EE2-2.5: Filtered Intra Block Copy (FIBC) [L. Zhang, Y. Yu (OPPO)] [late]

[JVET-AE0169](https://jvet-experts.org/doc_end_user/current_document.php?id=13132) EE2-2.2/EE2-2.3/EE2-2.4: Bi-predictive IBC GPM, bi-predictive IBC, and IBC MBVD list derivation for camera captured and screen contents [Y. Kidani, H. Kato, K. Kawamura (KDDI), C. Ma, X. Xiu, W. Chen, H.-J. Jhu, C.-W. Kuo, N. Yan, X. Wang (Kwai), G. Verba, Z. Zhang, P. Nikitin, H. Huang, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-AE0197](https://jvet-experts.org/doc_end_user/current_document.php?id=13160) Crosscheck of JVET-AE0169 (EE2-2.2c: Bi-predictive IBC-GPM and EE2-2.3c: Combination of 2.3b and 2.2c)) [W. Yin (Bytedance)] [late]

[JVET-AE0205](https://jvet-experts.org/doc_end_user/current_document.php?id=13168) Crosscheck of JVET-AE0169 (Test EE2-2.3c and EE2-2.4c) [R. Yu (Ericsson)] [late]

[JVET-AE0233](https://jvet-experts.org/doc_end_user/current_document.php?id=13196) Crosscheck of JVET-AE0169 (Test EE2-2.4a and EE2-2.4c) [D. Ruiz Coll (Ofinno)] [late] [miss]

[JVET-AE0196](https://jvet-experts.org/doc_end_user/current_document.php?id=13159) EE2-3.4: OBMC with HPel flag and BCW weights [A. Robert, F. Le Léannec, F. Galpin, T. Poirier, Y. Chen (InterDigital)] [late]

### EE2 related contributions (2)

Contributions in this area were discussed at 1100–1125 on Wednesday 12 July 2023 (chaired by JRO).

[JVET-AE0176](https://jvet-experts.org/doc_end_user/current_document.php?id=13139) EE2-related: Enhancements on CCRM [Z. Deng, K. Zhang, L. Zhang (Bytedance)]

This proposal presents the results of a multi-model CCRM mode and a CCRM merge mode. In EE2-3.1a, a single-model CCRM is investigated. In this contribution, two methods are presented to further improve the coding efficiency of CCRM. In method #1, a multi-model CCRM mode is proposed. Samples are separated into two groups and each group can apply its own CCRM model. There is no flag signalling involved in the proposed method, and the final CCRM mode is implicitly determined by comparing decoder derived costs of the two CCRM modes with single or multiple models. In method #2, a CCRM merge mode is proposed, wherein the CCRM model is inherited from a candidate in the CCRM merge list. Simulation results based on ECM-9.0 are reported as below:

Method #1 :

RA { -0.20%, -1.66%, -1.89%, 101.8%, 100.0% }; LDB { -0.01%, -3.86%, -3.46%, 101.9%, 99.7% }

Method #2 :

RA {}; LDB {}

Simulation results versus EE2-3.1a are listed as below :

Method #1 :

RA { -0.03%, -0.07%, -0.07%, 100.4%, 100.1% }; LDB { -0.12%, 0.39%, 0.42%, 100.7%, 99.8% }

Method #2 :

RA {}; LDB {}

Relative to EE test 3.1a, small benefit in RA ; in LB, some gain in luma, but loss in chroma.

According to proponents of 3.1, it would be a straightforward add-on, and might be interesting to investigate on top of the adopted method 3.1b.

Investigate in EE.

[JVET-AE0186](https://jvet-experts.org/doc_end_user/current_document.php?id=13149) Non-EE2-2.1: Block Vector Guided Chroma Direct Mode [K. Naser, F. Le Léannec, P. Bordes, Y. Chen, G. Rath (InterDigital)] [late]

Chroma direct mode (DM) is a special intra mode for chroma component where the same prediction mode of the collocated luma block is used for the current chroma block. In ECM-9.0, it is asserted that when the collocated luma block is coded with IBC or IntraTMP, DM is set to planar. This contribution proposes to use a method similar to EE2-2.1, where the collocated luma block vectors are used to obtain its reference block intra mode. The following results are obtained:

AI -0.03% -0.15% -0.10% with EncT 100% and DecT 100%

Only co-located BV is used.

According to proponents, this should be independent from adoptions to ECM10, and also give some gain on camera-captured content. Several experts expressed interest to study this further.

It was commented that there is a drop in class D, and also some inconsistency in terms of chroma.

It was requested to provide per-sequence results, and RA results in xls spreadsheet.

Was considered candidate for EE, revisit after providing per-sequence results

### ECM modifications and software improvements beyond EE2 (45)

Contributions in this area were discussed at 1125–1300 and 1430–1940 on Wednesday 12 July 2023, at 1500–2000 on Thursday 13 July 2023, and … (chaired by JRO).

***Intra (18)***

[JVET-AE0056](https://jvet-experts.org/doc_end_user/current_document.php?id=13004) Non-EE2: Changes on TIMD and SGPM for reducing ECM decoder complexity [Z. Fan, Y. Yasugi, T. Ikai (Sharp)]

This proposal aims to reduce the complexity of the ECM decoder. There are two changes made in this proposal for ECM.

* The first change is replacing timdHor and timdVer with dimdHor and dimdVer in the SGPM candidate list. This allows for the SGPM method to be used without the need to run the additional template matching process of TIMD.
* The second change involves reducing the number of template matching / number of candidate in TIMD, reducing the time required for the TIMD method.

Based on these changes, and the experiment results are as follows:

AI : 0.04% 0.08% 0.10% 99% 95%  
RA: 0.05% -0.01% -0.03% 100% 101%

It was commented that reduction of decoder complexity is not of primary importance in this exploration, and the proposal would incur performance losses.

No action.

[JVET-AE0201](https://jvet-experts.org/doc_end_user/current_document.php?id=13164) Crosscheck of JVET-AE0056 (Non-EE2: Changes on TIMD and SGPM for reducing ECM decoder complexity) [F. Wang (OPPO)] [late]

[JVET-AE0071](https://jvet-experts.org/doc_end_user/current_document.php?id=13019) AHG12: DIMD Merge [S. Blasi, I. Zupancic, J. Lainema (Nokia)]

This contribution proposes to merge the Decoder-side Intra Mode Derivation (DIMD) information from neighbouring blocks to predict the current block. When neighbouring blocks encoded with DIMD are available, the DIMD histograms are combined to form a new DIMD Merge histogram for the current block. DIMD Merge modes and weights are computed based on this merged histogram.

It is reported that adding the proposed DIMD Merge to ECM in AI configuration provides -0.08%, 0.01% and 0.07% BD-rate impact for Y, U and V, respectively. In RA configuration the impact is reportedly x%, y%, z% for Y, U and V, respectively. The encoder and decoder runtimes are reportedly minimally changed with respect to ECM.

Gain is slightly higher for larger resolution (A classes)

Is it necessary to store histograms of neighboring PUs? Yes, up to 13. Methods to reduce this are under consideration according to proponents. It was recommended to investigate reduction of storage needs.

Usage of this mode is up to 45%.

Investigate in EE. Also analysis of memory needs is requested.

[JVET-AE0073](https://jvet-experts.org/doc_end_user/current_document.php?id=13021) Non-EE2: IBC-LIC Model Merge mode [L. Zhang, Y. Yu, H. Yu, D. Wang (OPPO)]

This contribution proposes an IBC-LIC model merge mode that inherits IBC-LIC model parameters from spatial adjacent and non-adjacent neighbors, history candidates, or default models. The proposed method is implemented on top of ECM-9.0 software, it shows coding gains as follows:

AI: { -0.56% Y, -0.50% U, -0.45% V, 101.9% EncT, 100.9% DecT } for class F

{ -0.17% Y, -0.16% U, -0.15% V, 101.9% EncT, 101.3% DecT } for class TGM

RA: { -0.39% Y, -0.37% U, -0.35% V, 100.6% EncT, 102.3% DecT } for class F

{ -0.04% Y, -0.08% U, -0.09% V, 101.1% EncT, 102.4% DecT } for class TGM

It was asked why the gain is lower in class TGM ? Probably, current IBC-LIC already works well in that class.

Maximum 12 candidates are used

It was pointed out that depending on the potential adoption of EE2 2.10/2.11, some gain might be lost. Proponents believe that the multiple models still would have benefit.

Proponents were asked to run tests in combination with the EE2 candidates and provide results. Could be candidate for EE if it still provides gain.

Revisit.

[JVET-AE0268](https://jvet-experts.org/doc_end_user/current_document.php?id=13231) Crosscheck of JVET-AE0073 (Non-EE2: IBC-LIC Model Merge mode) [Y. Wang (Bytedance)] [late]

[JVET-AE0075](https://jvet-experts.org/doc_end_user/current_document.php?id=13023) Non-EE2: On IntraTMP block vector [Y. Yu, L. Zhang, L. Xu, H. Yu, J. Gan, F. Wang, Z. Xie, D. Wang (OPPO)]

IntraTMP block vectors (BVs) are stored for coding future blocks. Currently, an intraTMP BV may have a quarter-pel resolution but it is always saved in a full-pel resolution after coding an intraTMP block. This contribution proposes to store IntraTMP BVs in their full resolution. In addition, it is also proposed that intraTMP BVs are stored for HMVP. On top of ECM-9.0, the simulation results are reported as follows:

AI: -0.07% (Y), -0.04% (U), -0.08% (V), 98.1% (EncT), 99.9% (DecT) for class F

-0.10% (Y), -0.11% (U), -0.14% (V), 99.0% (EncT), 99.1% (DecT) for class TGM

RA: -0.07% (Y), -0.05% (U), -0.11% (V), 101.8% (EncT), 99.9% (DecT) for class F

-0.03% (Y), 0.00% (U), -0.07% (V), 101.5% (EncT), 100.0% (DecT) for class TGM

Very small gain, and results reported only for screen content. It was pointed out that it might be more beneficial in case of camera-captured content, where also IBC merge will be enabled in next ECM.

Investigate in EE. The two aspects should be reported separately, and also results with camera content should be reported.

[JVET-AE0246](https://jvet-experts.org/doc_end_user/current_document.php?id=13209) Crosscheck of JVET-AE0075 (Non-EE2: On IntraTMP block vector) [X. Li (Alibaba)] [late]

[JVET-AE0202](https://jvet-experts.org/doc_end_user/current_document.php?id=13165) Crosscheck of JVET-AE0075 (Non-EE2: On IntraTMP block vector) [W. Lim, S.-C. Lim (ETRI)] [late]

[JVET-AE0085](https://jvet-experts.org/doc_end_user/current_document.php?id=13033) Non-EE2: Direct block vector (DBV) mode extension [M. Hong, J. Choi, N. Park, J. Lim, S. Kim (LGE)]

This contribution proposes direct block vector (DBV) mode extension. Firstly, the number of locations to find a block vector are extended from 5 to 17 keeping the existing search range. Also, DBV candidate list is composed of block vector of collocated luma blocks that is coded with IBC or IntraTMP mode, and DBV candidate list is sorted in ascending order of SAD cost. The best one of DBV candidate list is used to apply block prediction for chroma. Experimental results of the proposed method for class F and class TGM are reported as follows:

For AI configuration:

F: -0.08%/-0.19%/-0.09%, 100.1%(EncT), 100.1%(DecT),

TGM: -0.10%/-0.15%/-0.24%, 99.3%(EncT), 99.4%(DecT).

For RA configuration:

F: -0.10%/0.06%/-0.49%, 100.4%(EncT), 100.0%(DecT),

TGM: -0.03%/-0.05%/-0.02%, 99.7%(EncT), 99.6%(DecT).

The additional complexity is not justified by the extremely small gain (<0.1% for screen content only).

No action.

[JVET-AE0263](https://jvet-experts.org/doc_end_user/current_document.php?id=13226) Crosscheck of JVET-AE0085 (Non-EE2: Direct block vector (DBV) mode extension) [H. Huang (Qualcomm)] [late] [miss]

[JVET-AE0098](https://jvet-experts.org/doc_end_user/current_document.php?id=13046) AHG12: On the chroma DBV mode [H. Huang, H. Wang, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes to enable chroma DBV mode in single-tree configuration and to remove the template matching based refinement in chroma DBV mode. The simulation on top of ECM-9.0 reports the following results {Y, U, V, EncT, DecT}:

AI: -0.01%, -0.03, 0.02%, 100%, 95%

RA: not complete yet

AI (single tree): -0.19%, -0.24%, -0.19%, 100%, 100%

The results in the abstract are including removal of the refinement. Table 3 of the contribution includes results without removing the refinement, but not complete yet.

Single Tree would be non-CTC

Some concern is raised about removing the refinement, as in some classes losses occur, the overall gain is minor, and it is not critical in terms of complexity. This may not be worthwhile to consider.

The aspect of enabling chroma DBV in single-tree mode appears interesting, but this would only affect inter slices in current CTC (for which no results are available). Revisit after complete results (including those without removal of refinement) are available.

[JVET-AE0203](https://jvet-experts.org/doc_end_user/current_document.php?id=13166) Crosscheck of JVET-AE0098 (AHG12: On the chroma DBV mode) [L. Xu, Y. Yu (OPPO)] [late]

[JVET-AE0266](https://jvet-experts.org/doc_end_user/current_document.php?id=13229) Crosscheck of JVET-AE0098 (AHG12: On the chroma DBV mode) [M. Hong (LGE)] [late] [miss]

[JVET-AE0110](https://jvet-experts.org/doc_end_user/current_document.php?id=13058) Non-EE2: An improved method for IntraTMP fusion [S. Peng, D. Jiang, J. Lin, C. Zheng, K. Fu, P. Zhang (Dahua)]

This contribution proposes an improved method for IntraTMP fusion. In this method, the filtered matched blocks are used in the IntraTMP fusion process. The proposed method was implemented on top of ECM-9.0, it reportedly provides {Y, U, V} BD-reduction with encoder and decoder runtimes as follows:

AI: -0.05%, -0.01%, -0.02%, 104.4% (EncT), 102.8% (DecT)

No reasonable tradeoff for gain vs. computation time (both encoder and decoder), which is caused by substantial additional processing.

No action.

[JVET-AE0120](https://jvet-experts.org/doc_end_user/current_document.php?id=13068) Non-EE2: Intra Angular Prediction Extension [J. Fu, J. Zhang, C. Jia, S. Ma (PKU)] [late]

In ECM, current angular intra prediction is based on linear projection. 1 of 67 intra prediction modes is chosen to be the angle of projection, and the location of reference pixel are calculated, then the value of predicting pixel is calculated through 6-tap intra filter and the value of reference pixel.

In this contribution, it is proposed to extend the current linear projection function into quadratic projection function that takes two intra prediction modes. When two intra prediction modes are belonging to the set of vertical intra prediction mode, the location of reference pixel is inferred by the location of predicting pixel and the quadratic function generated by two intra prediction modes.

This document illustrates test results of the proposed method.

On top of the ECM-9.0, simulation result of the proposed method is reported as:

AI: -0.02%, -0.01%, 0.21%, 124.6%, 107.9%

Coding gain low, and not consistent for the classes. Run time much too large.

Interesting idea, but not relevant for action as long as the performance is not better.

[JVET-AE0124](https://jvet-experts.org/doc_end_user/current_document.php?id=13072) Non-EE2: Fixes related to Intra TMP [J.-L Lin, P.-H Lin, Y.-J Chang, V. Seregin, M. Karczewicz (Qualcomm)]

In ECM-9.0, it is asserted that the cost list for Intra TMP candidates is not updated correctly and the BV storage in Intra TMP is not aligned with the IBC. This contribution proposes to fix those issues, and the simulation results on top of ECM-9.0 are summarized as follows:

TEST1:

AI(Y/U/V): 0.00%/ 0.01 %/ 0.03 %, EncT 100 %, DecT 99 %

RA(Y/U/V): X.XX %/ X.XX %/ X.XX %, EncT XX %, DecT XX %

TEST2:

AI(Y/U/V): -0.01%/ 0.03 %/ 0.02 %, EncT 100 %, DecT 104\* %

RA(Y/U/V): X.XX %/ X.XX %/ X.XX %, EncT XX %, DecT XX %

TEST3:

AI(Y/U/V): 0.00%/ 0.03 %/ 0.04 %, EncT 101 %, DecT 103\* %

RA(Y/U/V): X.XX %/ X.XX %/ X.XX %, EncT XX %, DecT XX %

For “Test 1”, it is not obvious that a problem exists with the current implementation. The “Fix 1” even has a (very small) loss in chroma. No action on this.

“Test 2” is somewhat equivalent to an aspect of JVET-AE0075, which was also promoted to be investigated in EE. The results on camera-captured content indicate that with current ECM (not using IBC merge in camera-captured classes) no gain can be achieved, but that might change due to upcoming EE adoptions.

Decoding time is not measured reliable.

Investigate Test 2 in EE together with JVET.AE0075.

[JVET-AE0247](https://jvet-experts.org/doc_end_user/current_document.php?id=13210) Crosscheck of JVET-AE0124 (Non-EE2: Fixes related to Intra TMP) [X. Li (Alibaba)] [late]

[JVET-AE0258](https://jvet-experts.org/doc_end_user/current_document.php?id=13221) Crosscheck of JVET-AE0124 aspect 1 (Non-EE2: Fixes related to Intra TMP) [F. Wang, Y. Yu (OPPO)] [late]

[JVET-AE0130](https://jvet-experts.org/doc_end_user/current_document.php?id=13093) Non-EE2: DIMD with filtered template [C. Zhou, Z. Lv (vivo)]

This contribution proposes a template filtering method for DIMD mode. The template is filtered using a 3x3 filter operator before obtaining the gradient histogram. The 3x2 and 2x3 gradient operators are used instead when calculating a gradient. The experimental results are summarized as follows:

On top of ECM-9.0

AI: -0.09% (Y), 0.01% (U), 0.04% (V), 106% (EncT), 100% (DecT)

RA : -0.xx% (Y), 0.xx% (U), 0.xx% (V), 10x% (EncT), 10x% (DecT)

Large increase in encoding time (likely caused by one more RD check for the decision of enabling or not enabling the filtering). Without the enabling flag, loss would occur.

In the current version, the tradeoff between compression and run time is not reasonable. Investigate in EE, but the encoding time needs to be largely reduced, without any substantial loss in performance.

It would also be desirable to avoid the modification of gradient operators, e.g. by using unfiltered samples for gradient computation at poisitions where no filtered samples are available.

Investigate in EE.

[JVET-AE0242](https://jvet-experts.org/doc_end_user/current_document.php?id=13205) Crosscheck of JVET-AE0130 (Non-EE2: DIMD with filtered template) [J. Chen (Alibaba)] [late] [miss]

[JVET-AE0137](https://jvet-experts.org/doc_end_user/current_document.php?id=13100) Non-EE2: On TIMD fusion [P. Andrivon, M. Blestel (Ofinno)]

The contribution proposes to conditionally introduce in the TIMD fusion process a non-angular intra predictor (DC or Planar) additionally to the two selected TIMD modes. On top of ECM-9.0, experimental results reportedly show the following results:

AI: -0.04% / 0.06% / 0.04%, with 100% EncT and 101% DecT

Why is loss observed for class F? Could be caused by smoothing which is not good for screen content which has more sharp edges. It was also pointed out that in some cases losses occur in chroma.

From the preliminary results, it also appears that more cases of loss occur in RA.

Further study to make benefit more consistent over different classes.

[JVET-AE0270](https://jvet-experts.org/doc_end_user/current_document.php?id=13233) crosscheck of JVET-AE0137: Non-EE2: On TIMD fusion [K. Naser (InterDigital)] [late] [miss]

[JVET-AE0138](https://jvet-experts.org/doc_end_user/current_document.php?id=13101) Non-EE2: Reference sample interpolation for intra prediction [G. Rath, F. Le Leannec, F. Urban, F. Racape, T. Dumas (InterDigital)]

For intra prediction along angular directions, this contribution reports the results of using a simple 4-tap filter in the case of DIMD, TIMD, ISP, and wide angular modes instead of the existing 6-tap filters in ECM 9. The filtering can be implemented as a linear interpolation together with an update term, where the update term is excluded if the two nearest reference samples are sufficiently close. Besides BD-rate performance gain, the advantages of the proposed method are lower complexity, no memory storage of filter coefficients, and a uniform method for both Luma and Chroma components.

It is reported that, on top of ECM-9.0, the proposed interpolation together with unified PDPC yields -0.10%, -0.14%, -0.05% and -0.02%, -0.17%, -0.12% in AI and RA configurations respectively for encoding and decoding runtimes of 100% and 99%, and 100% and 99%, respectively.

Standalone gain of 4-tap filter would be 0.05% (with decreased runtime), unified PDPC gives 0.04% standalone (with increased run time).

It was questioned if the filter (which has some additional computation for adaptation) is really simpler than the current 6-tap cubic filter, implementation-wise. Further, it was asked why with the “unified PDPC” the filter is only used at the top side.

The 6-tap filter is not replaced everywhere in ECM. Therefore, yet another filter is introduced.

For RA, performance is not consistent over different sequences/classes. Also, a somewhat different PDPC is used there. Definitely, the same design should be used in AI and RA.

Various concerns were raised whether the design is appropriate.

Further study recommended for a more consistent design, and more consistent results over different sequences and configurations.

[JVET-AE0208](https://jvet-experts.org/doc_end_user/current_document.php?id=13171) Crosscheck of JVET-AE0138 (Non-EE2: Reference sample interpolation for intra prediction) [P. Nikitin (Qualcomm)] [late]

[JVET-AE0158](https://jvet-experts.org/doc_end_user/current_document.php?id=13121) AHG12: LUT-based angle calculation for DIMD [A. Aminlou, J. Lainema (Nokia)]

In the current version of DIMD in ECM9.0, for each neighboring sample of the block, a ratio value is calculated based on the gradient in horizontal and vertical directions, and then the intra mode is derived by searching the ratio value inside an angle table. The ratio has high dynamic range between 0 and 65536 (i.e., 16 bits), and searching the ratio value inside the angle table is a sequential process. This contribution proposes to calculate the ratio value in logarithm-like domain and replace the ratio searching with a look-up-table. The following BD-rate and runtime impact observed when it is implemented in ECM-9:

AI { 0.01%, 0.05%, 0.05%, 99%, 97%}

RA { } // RA simulations are still running.

Actual reduction in decoding time should be in the range of 0.5% (numbers above not reliable).

Not of primary importance to make DIMD more “implementation friendly” at the current stage. It is also pointed out that a major code cleanup of the overall design would be desirable.

Further study recommended.

[JVET-AE0166](https://jvet-experts.org/doc_end_user/current_document.php?id=13129) AHG12: IntraTMP and IBC search area harmonization [D. Ruiz Coll, B. Chen, P. Andrivon, M. Blestel (Ofinno)]

This contribution explores the search area harmonization of the IBC and IntraTMP tools. IBC and IntraTMP use the reconstructed samples of the current picture to select the prediction block for the current block. The IBC search area is fixed to a specific number of CTUs according to the current block locations, while the IntraTMP search area is based on the block size, which exceeds the IBC search region for some block sizes. In order to harmonize the decoded buffer for both tools, this contribution investigates two aspects, the constraint of the current IntraTMP search regions and the extent of such regions to the IBC search region.

On top of ECM-9.0, simulation results are reported as follows:

Test-1: IntraTMP search sub-regions constraint to IBC search area.

AI {-0.00% Y, -0.00% U, -0.00% V, 99.3% EncT, 99.1% DecT};

Test-2: Adjustment of the IntraTMP sub-regions to the IBC search area.

AI: Class F : {-0.44%, -0.57%, -0.38%, 102.4%, 109.5%},

AI: Class TGM : {-0.68%, -0.68 %, -0.69%, 102.5%, 114.5%};

Presentation deck to be provided.It was commented that “test 1” seems to have some relation to elements of EE2-2.9. Harmonization is not of prior importance at this moment, and the benefit in runtime is relatively small.

For test 2, it was commented that the tradeoff compression vs. encoder/decoder run time is not appropriate. In adoptions from EE, screen content gains around 0.5% usually came with almost no increase in run time.

Not attractive for EE, further study for better tradeoff.

[JVET-AE0269](https://jvet-experts.org/doc_end_user/current_document.php?id=13232) crosscheck of JVET-AE0166: AHG12: IntraTMP and IBC search area harmonization [K. Naser (InterDigital)] [late] [miss]

[JVET-AE0167](https://jvet-experts.org/doc_end_user/current_document.php?id=13130) AHG12: DBV improvement [L. Xu, Y. Yu, H. Yu, D. Wang (OPPO)]

When dual-tree block partition is enabled for an intra slice in ECM9, DBV is one of the chroma coding modes if one of the five corresponding luma positions is coded with either intraTMP or IBC. If the current luma block is coded with IBC or IntraTMP and DBV is not used, this contribution proposes to use one of the 5 prediction modes, i.e., DIMD, DC, Planar, Horizontal, and Vertical, without the DM flag to code the corresponding chroma block. On top of ECM-9.0, the simulation results are reported as follows:

AI: -0.02% (Y), -0.09% (U), -0.08% (V), xx% (EncT), xx% (DecT)

RA: -0.01% (Y), -0.13% (U), -0.01% (V), xx% (EncT), xx% (DecT)

Encoding/decoding times are not reliable,

Somewhat similar approach to JVET-AE0186, but that proposal showed slightly better (though still low) results. Proponents of JVET-AE0186 believe that the two proposals are not overlapping and might give additive gain.

Investigate in EE.

[JVET-AE0265](https://jvet-experts.org/doc_end_user/current_document.php?id=13228) Crosscheck of JVET-AE0167 (AHG12: DBV improvement) [H. Huang (Qualcomm)] [late]

[JVET-AE0184](https://jvet-experts.org/doc_end_user/current_document.php?id=13147) AHG12: TIMD with IntraTMP and IBC candidates [K. Naser, F. Le Léannec, P. Bordes, F. Galpin, A. Robert (InterDigital)] [late]

TIMD is a an intra tool that uses template analysis to generate the prediction signal. Specifically, a set of intra candidate modes are tested on the template and the template distance (SATD) is measured to deduce up to two best modes. This contribution proposes extending the candidate modes with IBC and IntraTMP candidates obtained from neighboring blocks. The proposed modification is tested on top of ECM-9.0 and the following results are obtained:

AI: -0.10% -0.13% -0.16% with EncT 101% and DecT 102%

IBC and IntraTMP are added as candidates into TIMD search, which is then performed as in current ECM. Up to 6 additional candidates.

Several experts expressed interest, as this is relatively straightforward and tradeoff is reasonable.

It was requested to provide per-sequence results, and RA results in xls spreadsheet.

Investigate in EE

[JVET-AE0185](https://jvet-experts.org/doc_end_user/current_document.php?id=13148) AHG12: IntraTMP with neighboring candidates [K. Naser, F. Le Léannec, A. Robert, F. Galpin (InterDigital)] [late]

In ECM-9.0, IntraTMP search is performed within predefined search area to obtain the best candidate block vectors. This contribution proposes to augment the TM search with the block vectors of neighboring PU’s that are coded with IBC or IntraTMP. The purpose is to improve the IntraTMP by further including probable candidates during the TM search. On top of ECM-9.0, the following results are obtained:

AI -0.09% 0.18% 0.17% 100% 100%

5 additional candidates in IntraTMP. Search area is not going beyond IBC search range. Only uses BVs that are already stored in current ECM.

Where does chroma loss for some cases come from? Was not analysed.

Tradeoff is reasonable.

It was requested to provide per-sequence results, and RA results in xls spreadsheet.

Was considered candidate for EE, revisit after providing per-sequence results, and an answer to the question on partial chroma losses.

[JVET-AE0198](https://jvet-experts.org/doc_end_user/current_document.php?id=13161) AHG12: Extended Search Region for IntraTMP [K. Naser, P. Bordes, F. Le Léannec, F. Galpin (InterDigital)] [late]

This contribution proposes an extension to the search range of IntraTMP tools. The extended search range is the same one as of IBC mode. It is based on adding new search region with specific subsampling factor to control the number of added SAD checks.

On top of ECM-9.0, simulation results are reported as follows:

AI:

CTC: -0.03% 0.24% 0.02% 102% 103%

classF: -0.40% -1.66% -0.28% 102% 103%

ClassTGM: -0.99% -1.38% -1.20% 103% 104%

It was commented that the tradeoff compression vs. encoder/decoder run time is not appropriate. In adoptions from EE, screen content gains around 0.5% usually came with almost no increase in run time.

Search region is not enlarged, but somewhat re-arranged in sub-regions. Similar concepts as in JVET-AE0166, but various aspects (subsampling, clipping) ar different.

Interesting gain for screen content, but no good tradeoff with encoder runtime (in particular for CTC)

Results somewhat inhomogeneous over classes, and chroma losses in several cases.

It was requested to provide per-sequence results, and RA results in xls spreadsheet.

Not attractive for EE, further study for better tradeoff.

[JVET-AE0214](https://jvet-experts.org/doc_end_user/current_document.php?id=13177) AHG12: on Intra CIIP [K. Naser, P. Bordes, F. Galpin, K. Reuzé (InterDigital)] [late]

IntraCIIP is a method proposed in JVET-AC0201 and studied in last EE under the test of EE2-1.18 (JVET-AD0177). This tool was not tested on top of the adoption of several IntraTMP improvement of EE2-1.20j (JVET-AD0086). It is reported that IntraCIIP provides further coding gain compared to last EE as it benefits from the overall improvement of IntraTMP. The following result are obtained on top of ECM-9.0

AI: -0.08% -0.02% 0.07% 100% 99%

It was commented that the contribution does practically not contain any description of the technology, just references to previous proposals. This makes it difficult to understand (this comment relates to various proposal docs, and some directives what a good proposal shall contain. should be added somewhere in the meeting notes regarding experiment rules, and/or in the document template – also for including results in Excel sheet).

It was requested to provide per-sequence results, and RA results in xls spreadsheet.

Reasonable tradeoff, would be interesting to investigate whether the gain is retained with ECM10.

It was further commented that there might be interdependency with JVET-AE0184, these should be tested in combination.

It was confirmed that the method is the same as with InterCIIP, logic could be re-used.

Investigate in EE.

***Inter (11)***

[JVET-AE0047](https://jvet-experts.org/doc_end_user/current_document.php?id=12995) AHG12: On GPM-MMVD [R. Yu (Ericsson)]

GPM-MMVD provides a way to adjust motion vectors for inter partitions of GPM blocks. Similarly to merge-MMVD, GPM-MMVD explicitly signals MVD which are applied on top of merged motion vectors. The MVD magnitudes that can be signaled in GPM-MMVD is limited within sets which always include fractional MVDs. It is proposed to let GPM-MMVD follow the merge-MMVD design to allow possibilities of disabling those fractional MVDs. The proposed modification is reported to be implemented in ECM-9.0 and tested for ClassA1 and ClassA2 under the ECM RA CTC. The BD-rate impact is reported to be -0.01% for ClassA1 and 0.01% for ClassA2.

In v2 of the contribution, additional results for disabling fractional MVD for GPM-MMVD on screen contents are provided.

It was commented that a per-sequence switch might give more gain. Currently, switch is only performed for A classes.

The cross-checker mentioned that there might be more advantage in the context of the newly adopted bi-predictive GPM.

Further study (not in EE from this meeting).

[JVET-AE0218](https://jvet-experts.org/doc_end_user/current_document.php?id=13181) Crosscheck of JVET-AE0047 (AHG12: On GPM-MMVD) [Z. Deng (Bytedance)] [late]

[JVET-AE0105](https://jvet-experts.org/doc_end_user/current_document.php?id=13053) Non-EE2: Local illumination compensation with multiple templates [Y. Wang, K. Zhang, Y. He, H. Liu, L. Zhang (Bytedance)]

This contribution presents a method of local illumination compensation with multiple templates. Apart from using a template comprising neighbouring samples both top to and left to the current block, two new LIC modes using top-only or left-only template are proposed to derive the parameters of LIC. On top of ECM-9.0, simulation results of the proposed method are reported (estimated) as below:

RA: {-0.09%%, -0.20%, -0.16%; 109.4%, 100.1%};

LDB\*: {-0.06%, -0.15%, -0.32%; 117.6%, 99.6%}.

Different shapes could be used for different motion vectors in same location.

Is there dependency of templates between adjacent blocks? Not currently exploited.

Possibility to reduce encoding time? No concrete ideas or preliminary experimentation yet.

Also applied to chroma.

Several experts expressed this is an interesting approach. Further study recommended on largely reducing encoding time while keeping the benefit in compression.

[JVET-AE0224](https://jvet-experts.org/doc_end_user/current_document.php?id=13187) Crosscheck of JVET-AE0105 (Non-EE2: Local illumination compensation with multiple templates) [C. Ma (Kwai)] [late] [miss]

[JVET-AE0108](https://jvet-experts.org/doc_end_user/current_document.php?id=13056) AHG12: DMVR with robust MV derivation [K. Andersson, R. Yu (Ericsson)]

In JVET-AD0045, it was reported that DMVR in VVC can sometimes produce dislocated subblocks with unreliable motion vectors which leads to subjective artifacts. It is found that multi-pass DMVR in ECM can also have such a problem. It is therefore proposed to include additional criterions on top of bilateral block matching cost calculation for cases where there is a risk to have an unreliable motion vector to make sure that selected motion vectors result in subblocks that are aligned with neighbouring subblocks. The proposed method inherits low spatial activity check, subblock motion vector difference check and boundary difference checks from JVET-AD0045. When there is a risk for unreliable motion vectors, a boundary distortion metric is added to the block matching cost to bias the selection of motion vectors to produce subblocks that are aligned with above and left neighbouring subblocks.

Objective performance (BDR) versus ECM-9.0 is reported as:

For CTC at normal QP range:

RA: Y: 0,00%, U -0,03%, V -0,01% Enc:100% Dec:101%

For CTC at higher QP range QP= 40, 43, 47, 50:

RA: Y -0,07%, U -0,06%, V -0,02% Enc: 100% Dec:101%

It is also asserted that the proposal can improve subjective quality.

It was asked if it would help to apply deblocking at DMVR subblock boundary? Was tried, but did not help. The DMVR outlier also influences the inner part of the block.

Threshold is QP adaptive.

Why is there some gain for high QP? Might be due to better consistency of MVs in combination with almost no residual updates.

The contribution solves a problem that is known from VTM and was resolved by an encoder modification (which reduces the compression performance). In VVC, it could not be done decoder-side, as the problem was only detected after finalization of the standard.

It is asserted that it would be valuable to solve the problem in ECM, where a decoder-side solution could be applied which comes without bit rate overhead.

Investigate in EE. Comparison could be made in terms of bit rate saving against an encoder-only solution (in ECM, similar to the one of VTM), and some subjective comparison should also be performed (possibly for subjective comparison also using non-CTC sequences, e.g. from verification tests).

[JVET-AE0109](https://jvet-experts.org/doc_end_user/current_document.php?id=13057) Non-EE2: LIC flag derivation of merge candidates with template costs [N. Zhang, K. Zhang, H. Liu, Y. Wang, L. Zhang (Bytedance)]

In ECM, the LIC flag is inherited for a merge candidate. It is proposed to derive the LIC flag of a merge candidate based on template costs. A SAD-based template cost, denoted as C0, and a Mean Removal SAD (MRSAD)-based template cost, denoted as C1, are calculated. The LIC flag is set to be false, if C0 < C1 and is set to be true, if C0 >= C1.

On top of the ECM9.0, simulation results of the proposed method are reported as below:

RA: -0.01%, -0.06%, -0.01%, 99%, 99%; LB: -0.12%,-0.12%,-0.26%,100%,101%.

Interesting gain for LB.

One expert mentions that the criterion is interesting. Why is the criterion not effective for RA?

It was pointed out that potentially a similar approach could be used for LIC flag generally, not only in merge case.

Study in EE.

[JVET-AE0117](https://jvet-experts.org/doc_end_user/current_document.php?id=13065) Non-EE2: Enhanced subblock-based motion compensation [L. Zhao, K. Zhang, L. Zhang (Bytedance)]

This proposal presents the combined solution which was initially proposed in JVET-AC0187 and JVET-AD0235. Three aspects are proposed to enhance the subblock-based motion compensation.

Aspect #1: TM-based subblock motion refinement.

Aspect #2: Interweaved affine prediction.

Aspect #3: RMVF candidate derivation with multiple CUs as input.

It is reported that on top of ECM-9.0, simulation results of the proposed method are as below:

RA: -0.10 % (Y), -0.11 % (U), -0.01 % (V), 102.3% (EncT), 100.8% (DecT)

LB: -0.05% (Y), -0.05% (U), -0.15 % (V), 106.8% (EncT), 102.7% (DecT)

Better tradeoff than previous proposal, but in particular for LB not yet attractive

What is the impact on the decoder complexity? Probably, subblock processing in generating the prediction becomes more complex.

What is the benefit of the three aspects, and what is their complexity impact?

Only applied to first candidate, out of complexity considerations

Investigate in EE, but beyond answering the questions above, the encoder run time should be further reduced to make it attractive.

[JVET-AE0129](https://jvet-experts.org/doc_end_user/current_document.php?id=13088) AHG12 Template-based CIIP weight derivation [M.-H. Jia, Y.-L. Hu, S.-W. Xie, Y. Gao, C. Huang (ZTE)]

This contribution proposes to derive the inter weight and intra weight of CIIP based on template matching cost. Given a CIIP coded CU, the TM cost values are calculated with different CIIP weights setting. Then, the CIIP weight with minimum TM cost value is selected to fusion intra prediction and inter prediction of CIIP. It is reported that on top of ECM-9.0, the overall coding performance impact are as follows:

* LDB: xx%, xx%, xx%
* RA: xx%, xx%, xx%

Incomplete results, classes B/C/D come with bit rate losses.

Further study.

[JVET-AE0136](https://jvet-experts.org/doc_end_user/current_document.php?id=13099) Non-EE2: Fix on TM-based reordering for affine MMVD mode [D. Kim, W. Lim, J. Kim, S.-C. Lim, J. S. Choi (ETRI)]

This contribution addresses a bug in the implementation of template matching (TM)-based reordering for affine merge with motion vector difference (MMVD) mode in ECM-9.0. In the affine MMVD mode, the refinement positions for the base candidate motion are reordered based on template costs calculated using the sum of absolute difference (SAD). However, during the reordering process, there is an issue where the incorrect reference picture index is used when calculating the picture order count (POC) of the reference picture from list 1, leading to inaccuracies in template SAD cost calculations. To resolve this issue, the proposed fix involves replacing the reference picture index of list 0 with the correct reference picture index of list 1 in "sortAffineMergeCandidates" function. Experimental results reportedly shows that the proposed fix has a negligible impact on coding performance in the RA case. Therefore, it is recommended to incorporate the proposed fix in the next version of ECM.

Decision(BF/SW): Adopt JVET-AE0136

[JVET-AE0213](https://jvet-experts.org/doc_end_user/current_document.php?id=13176) Crosscheck of JVET-AE0136 (Non-EE2: Fix on TM-based reordering for affine MMVD mode) [M. Salehifar (Bytedance)] [late]

[JVET-AE0140](https://jvet-experts.org/doc_end_user/current_document.php?id=13103) Non-EE2: LIC extensions [A. Filippov, V. Rufitskiy, K. Suverov (Ofinno)] [late]

The contribution proposes an additional method of filtering (interpolated) reference samples for inter predicted blocks. This method is complementary to LIC that utilizes 2-parametric linear model whereas the proposed technique involves high-order models like done in CCCM. On top of ECM-9.0, the following BD-rate gain and run-time results are reportedly obtained for the proposed method {Y, U, V, EncT, DecT}: {-0.XX%, -0.XX%, -0.XX%, 10X.X%, 10X.X%} and {-0.XX%, -0.XX%, -0.XX%, 10X.X%, 10X.X%} for RA and LDB configurations, respectively.

Preliminary results (RA) indicate -0.16%, -0.24%, -0.18% with 102.2% encoder and 100.6% decoder

Higher-order models are more than an extension, mixture of explicit and implicit signalling, for the latter a “probe samples template” is introduced including also one more line/column of reference samples.

Up to 9 models (not all may be used).

More detail about models: Non-linear/linear/offset terms?

Very interesting technology, more detail requested, further study recommended.

Revisit: Investigate in EE provided that more detail is provided in an update of the contribution, and it could more clearly be specified which aspects to be tested.

[JVET-AE0148](https://jvet-experts.org/doc_end_user/current_document.php?id=13111) Non-EE2: Affine subblock BDOF refinement [Z. Zhang, H. Huang, J.-L Lin, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes to apply BDOF to refine subblock MV for an affine coded block. When an affine coded block meets the BDOF enabling condition and it is decided to do subblock MC, BDOF is applied to refine each subblock MV and do sample adjustments. The refined MVs are stored for future block MV prediction. The proposed method is also applied to SbTMVP mode.

It also proposes the AMVP-Merge mode for affine block. When a block is decided to be AMVP-Merge mode, a flag is signalled to further indicate the block is an affine block or not.

The proposed method was implemented on top of ECM-9.0 and the test results are as follows:

BDOF affine subblock MV refinement:

RA : -0.14% (Y), -0.17% (U), -0.07% (V), 103.0% (EncT), 104.4% (DecT)

BDOF affine subblock MV refinement and AMVP-Merge mode for affine:

RA : -0.19% (Y), -0.21% (U), -0.13% (V), 107.4% (EncT), 104.3% (DecT)

PROF will not be applied in case when BDOF is applied to affine subblocks. Also not applied in case of pixel-based MC.

Blocksize cannot be lower than 4x4.

It was asked if the method including AMVP merge for affine is simpler for decoder?

It was suggested to test AMVP merge for affine separately (without BDOF for affine)

Can encoding time be reduced? There are some options, e.g. storage of computed MVs, skipping checks in certain cases.

Investigate in EE.

[JVET-AE0168](https://jvet-experts.org/doc_end_user/current_document.php?id=13131) Non-EE2: Sample-based BDOF for Chroma [C.-C. Chen, H. Huang, Z. Zhang, C. S. Coban, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposed to extend the notion of sample-based bi-directional optical flow (BDOF) to refining bi-prediction samples of chroma components. The same process of sample-based BDOF in ECM is applied to each chroma sample in a coding unit, except that its per-sample motion refinement is inferred from its co-located luma components in the same coding unit. In other words, the per-sample motion refinement of a luma sample is shared across all color components. Then, each bi-prediction chroma sample is adjusted using an offset value computed by the same BDOF function of gradient vector and motion refinement. The experimental results on top of ECM-9.0 are reported, as follows:

- Class B: (Y) 0.01%, (U) -0.13%, (V) -0.03%, (Enc) 100.3%, (Dec) 100.2%

- Class C: (Y) -0.01%, (U) -0.15%, (V) -0.13%, (Enc) 99.6%, (Dec) 99.3%

- Class D: (Y) 0.00%, (U) -0.45%, (V) -0.35%, (Enc) 100.3%, (Dec) 99.9%

- Class F: (Y) -0.02%, (U) -0.11%, (V) -0.09%, (Enc) 100.3%, (Dec) 100.0%

- Class TGM: (Y) 0.01%, (U) -0.08%, (V) -0.08%, (Enc) 100.4%, (Dec) 99.5%

Results A1/A2 not available yet.

It is claimed that the gain comes for free, but other experts expressed opinion that additional processing (or logic in hardware) would be necessary. Surprising that decoder run time is decreased, may be unreliable.

Opinion was expressed that the gain is too low to be considered. Apparently, BDOF does not have a significant benefit on chroma.

Class D has higher gain – why? Perhaps chroma has more structure in low resolution than in high resolution. Also most gain from basketballpass.

No action.

[JVET-AE0177](https://jvet-experts.org/doc_end_user/current_document.php?id=13140) Non-EE2: Local illumination compensation with slope adjustment [Y. Wang, K. Zhang, Y. He, H. Liu, L. Zhang (Bytedance)]

This contribution presents a method of local illumination compensation with slope adjustment, in which an adjustment parameter is used to modify parameters of LIC. On top of ECM-9.0, simulation results of the proposed method are reported (estimated) as below:

RA\*: {-0.10%, -0.22%, -0.17%; 112.9%, 100.3%};

LDB\*: {-0.10%, -0.18%, -0.31%; 120.8%, 100.4%}.

Slope adjustment similar to CCLM

Several experts this is interesting and straightforward, but tradeoff is not reasonable

Further study recommended on largely reducing encoding time while keeping the benefit in compression, and potentially further improvements in compression.

[JVET-AE0248](https://jvet-experts.org/doc_end_user/current_document.php?id=13211) Crosscheck of JVET-AE0177 (Non-EE2: Local illumination compensation with slope adjustment) [X. Li (Alibaba)] [late]

***Cross Component Prediction (5)***

[JVET-AE0074](https://jvet-experts.org/doc_end_user/current_document.php?id=13022) Non-EE2: Improvement on cross-component prediction merge mode [H. Huang, Y. Yu, H. Yu, D. Wang (OPPO)]

This contribution presents improvement methods for the cross-component prediction (CCP) merge mode. The CCP merge mode is proposed to be applied to the linear-model part of the chroma fusion mode. The cross-component model parameters applied to the linear-model predictor of chroma fusion mode are inherited from the spatial adjacent, spatial non-adjacent, history-based, or default CCP models. In addition, a filter flag for local boosting is also proposed to be inherited in the CCP merge mode. Therefore, the prediction samples generated by the CCP merge mode may also be filtered with neighboring samples.

On top of ECM-9.0, the simulation results of the proposed method are reported as follows:

AI: {-0.01%, -0.34%, -0.33%, 100.9%, 99.9%}; RA:{ -0.xx%, -0.xx%, -0.xx%, xxx%, xxx%}

The two aspects should be investigated separately.

There are similar aspects in JVET-AE0170

Further study recommended if more gain can be achieved, potentially also by combination.

[JVET-AE0254](https://jvet-experts.org/doc_end_user/current_document.php?id=13217) Crosscheck of JVET-AC0074 (Non-EE2: Improvement on cross-component prediction merge mode) [X. Li (Alibaba)] [late]

[JVET-AE0260](https://jvet-experts.org/doc_end_user/current_document.php?id=13223) Crosscheck of JVET-AE0074 (Non-EE2: Improvement on cross-component prediction merge mode) [K. Zhang (Bytedance)] [late] [miss]

[JVET-AE0097](https://jvet-experts.org/doc_end_user/current_document.php?id=13045) AHG12: On the cross-component merge mode [H. Huang, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes to remove the template based offset update process for the cross-component merge candidate.

On top of ECM-9.0, the average BD-rate impact, encoder and decoder run times are summarized as follows {Y, U, V, EncT, DecT}:

AI: -0.01%, 0.01, 0.00%, 98.8%, 97.0%; class TGM -0.39%, -1.01%, -1.02%, 99.4%, 99.7%

RA: -0.02%, 0.00%, -0.22%, 99.8%, 99.9%

Run time numbers are likely unreliable.

The results are confirmed by two cross-checks, and the change is straightforward due to the cross-checkers. It was further mentioned that when the template based offset update process was adopted, it had only been tested on AI and not on screen content and not on RA, and in the meantime some other elements in cross-component processing had been added.

Decision: Adopt JVET-AE0097.

[JVET-AE0188](https://jvet-experts.org/doc_end_user/current_document.php?id=13151) Crosscheck of JVET-AE0097 (AHG12: On the cross-component merge mode) [K. Zhang (Bytedance)] [late] [miss]

[JVET-AE0243](https://jvet-experts.org/doc_end_user/current_document.php?id=13206) Crosscheck of JVET-AE0097 (AHG12: On the cross-component merge mode) [C.-M. Tsai (MediaTek)] [late]

[JVET-AE0115](https://jvet-experts.org/doc_end_user/current_document.php?id=13063) Non-EE2: Unified intra CC-models parameters precision P. Bordes, K. Naser, F. Galpin, F. Le Léannec (InterDigital)]

In ECM-9.0, several intra coding tools use cross-component (CC) prediction. However, some of these tools use 64-bits precision for storage and applying CC-models, while others use 32-bits. It is proposed to unify the CC-model parameters to use 32-bits precision. It is reported that the proposed change does not impact the RD performance of ECM-9.0: AI 0.00% -0.01% 0.01%, RA xx% xx% xx%.

It was asked which would be the actual reduction in memory usage?

It was asked what the actual dynamic range of the parameters is? Is 32 bits necessary? Is it necessary to specify a shifting additionally, depending on which the value range of a given parameter is.

Several experts supported the proposal, being straightforward to implement.

It was further argued that a more clean solution would be to compute the parameters in 32 bits from the beginning, rather than leaving the 64 bit computation unchanged and convert them into 32 in a second step, as proposed here.

Without doubt, this would be necessary for a standard, but then a more consistent overall design would be necessary. Not urgent in exploration stage.

[JVET-AE0244](https://jvet-experts.org/doc_end_user/current_document.php?id=13207) Crosscheck of JVET-AE0115 (Non-EE2: Unified intra CC-models parameters precision) [C.-Y. Chuang (MediaTek)] [late]

[JVET-AE0170](https://jvet-experts.org/doc_end_user/current_document.php?id=13133) Non-EE2: Enhancements on CCP merge [K. Zhang, Z. Deng, L. Zhang (Bytedance)]

This proposal presents the results of the enhanced CCP merge mode. The fusion prediction can be generated by the CCP-merge prediction and the MM-CCCM prediction as a weighted sum. Besides, CCP merge mode is harmonized with LB-CCP by including the LB-CCP flag in the CCP merge candidate. Simulation results based on ECM-9.0 are reported as below:

AI {-0.03%, -0.43%, -0.45%, 101%, 100%}; RA {}.

The two aspects should be studied separately.

There are similar aspects in JVET-AE0074

Further study recommended if more gain can be achieved, potentially also by combination.

[JVET-AE0261](https://jvet-experts.org/doc_end_user/current_document.php?id=13224) Crosscheck of JVET-AE0170 (Non-EE2: Enhancements on CCP merge) [H. Huang (OPPO)] [late] [miss]

[JVET-AE0178](https://jvet-experts.org/doc_end_user/current_document.php?id=13141) Non-EE2: Cross-component prediction merge mode for chroma inter coding [M.-S. Chiang, H.-Y. Tseng, C.-M. Tsai, C.-Y. Chuang, C.-W. Hsu, C.-Y. Chen, T.-D. Chuang, O. Chubach, Y.-W. Chen, Y.-W. Huang, S.-M. Lei (MediaTek)]

In ECM-9.0, when the current chroma intra block is coded in cross-component prediction (CCP) merge mode, the cross-component model parameters of this chroma block can be inherited from previously coded blocks, as described in JVET-AD0188. This contribution extends the CCP merge mode to inter coding blocks, where the final chroma inter prediction combines motion-compensation predicted signals and cross-component predicted signals derived using an inherited CCP model.

On top of ECM-9.0, the simulation results of the proposed method are reported below:  
RA: {Y BD-rate = 0.xx%, Cb BD-rate = -0.xx%, Cr BD-rate = -0.xx%, EncT = 10x%, DecT = 10x%}

Some preliminary results from powerpoint deck:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Over ECM-9.0** | | | | | **RA** | | | | |
| Y | | Cb | | Cr | | EncT | | DecT | |
| **Class A1** | -0.48% | | -1.97% | | -2.64% | | 107.1% | | 103.1% |
| **Class A2** | | | | | | | | | |
| **Class B** | 0.04% | | -1.22% | | -1.07% | | 104.7% | | 100.6% |
| **Class C** | 0.04% | | -0.73% | | -0.74% | | 105.7% | | 99.7% |

The method uses the CCP merge mechanism from intra and applies it to inter coded blocks, and allows weighted combination with MC prediction. A temporal candidate is added to the merge list.

High gain is reported for class A, significantly higher for Campfire.

It was commented that some of the gain may be reduced due to adoption of EE2-3.1b, which also takes benefit from cross-component in intra.

Tradeoff runtime/compression not attractive in the current version.

Several experts supported investigation in EE, possibly in combination with changes made to CCP merge list construction in ECM10, and reduction of encoder runtime. Test impact of adding the temporal candidate to the merge list.

***In-Loop Filters (3)***

[JVET-AE0044](https://jvet-experts.org/doc_end_user/current_document.php?id=12992) AHG12: Dynamic Scaling of Bilateral Filter (BIF) [V. Shchukin, P. Wennersten, J. Ström (Ericsson)]

This contribution proposes three additional techniques for the Bilateral Filter (BIF):

1. The TU scale factor depends on the TU shape size.
2. The TU scale factor depends on the mean absolute difference (MAD) of the TU.
3. The BIF LUTs are interpolated.

The proposed algorithm is implemented on the top of ECM-9.0. According to simulation results under common test conditions (CTC), BD-rate PSNR impact is reported to be -0.14%/-0.11% for AI/RA for Luma, the average decoding run time is reported to be 100.5%/99.8% for AI/RA. For AI, the average encoding run time is reported to be 101.4%. For RA, the encoding run time averaged over class B and class C is reported to be 100.5%.

Some complexity increase, but OK for exploration

Several experts supported this as an interesting approach

Investigate in EE.

[JVET-AE0190](https://jvet-experts.org/doc_end_user/current_document.php?id=13153) Crosscheck of JVET-AE0044 (AHG12: Dynamic Scaling of Bilateral Filter) [W. Yin (Bytedance)] [late]

[JVET-AE0121](https://jvet-experts.org/doc_end_user/current_document.php?id=13069) Non-EE2: Luma Residual Taps in Chroma-ALF and CCALF [W. Yin, K. Zhang, Z. Deng, L. Zhang (Bytedance)]

In the current adaptive loop filter (ALF) design of ECM-9.0, luma residual values are stored and used in the Luma-ALF. In this contribution, luma residual values are further introduced into Chroma-ALF and CCALF.

On top of ECM-9.0, simulation results of the proposed method are reported as below:

RA: \*0.02%, -0.56%, -0.67%, 99.8%, 100.0%.

LB: %, %, %, %, %.

Preliminary results (not class B yet) indicate higher chroma gain for LB than in RA.

It was found that no benefit (some chroma losses) would be present in AI. The proponents suggest this might be caused by different characteristics of intra residual. In case of luma, the benefit was also lower for AI, but no loss.

Several experts supported study in EE. Separate results should be provided for chroma ALF and CCALF, and also the AI configuration should be studied.

Single tree might be studied in AI, in case that the separate tree is the cause of losses.

[JVET-AE0231](https://jvet-experts.org/doc_end_user/current_document.php?id=13194) Crosscheck of JVET-AE0121 (Non-EE2: Luma Residual Taps in Chroma-ALF and CCALF) [Z. Xie (OPPO)] [late]

[JVET-AE0131](https://jvet-experts.org/doc_end_user/current_document.php?id=13094) Non-EE2: Variance based classification for in-loop filtering [W. Yin, K. Zhang, L. Zhang (Bytedance)]

In this contribution, a variance-based classification for in-loop filtering is proposed. In the proposed method, variance is computed and used as an additional classification rule for BF and ALF. The existing classification methods in BF and ALF are kept unchanged.

On top of ECM-9.0, simulation results of the proposed methods are reported as below:

AI: \*-0.02%, -0.02%, 0.03%, 100.0%, 100.1%.

RA: \*-0.04%, -0.12%, 0.07%, 100.0%, 99.8%.

LB: %, %, %, %, %.

How was the weight derived? Boundary info from partitioning, dependency on variance and combination with existing classification are only vaguely described and the proponents were requested to provide an update of the contribution.

Two filters were modified, but gain is not significant enough to consider the proposal.

Further study if more benefit can be achieved.

Cross-checker found increase of decoding time, sees potential for improvement in gain.

[JVET-AE0207](https://jvet-experts.org/doc_end_user/current_document.php?id=13170) Crosscheck of JVET-AE0131 (Non-EE2: Variance based Classification for In-loop Filtering) [V. Shchukin (Ericsson)] [late]

***Entropy coding and transform coefficient coding (2)***

[JVET-AE0058](https://jvet-experts.org/doc_end_user/current_document.php?id=13006) AHG12: Spatial CABAC tuning [J. Lainema, A. Aminlou, P. Astola, R. G. Youvalari, D. Bugdayci Sansli (Nokia)]

This contribution asserts that the current CABAC context update mechanism may lead to “overfitting” the probability estimates to the statistical properties of the CTU that is preceding the current CTU in scan order. To help overcome such bias, it is proposed to update the CABAC contexts using syntax element values retrieved from the above CTU when starting to process a new CTU. A limited number of bins (e.g., 512) of syntax elements belonging to the bottom CUs of the above CTU are fed into the context update mechanism of the shorter window probability estimator to achieve this.

It is reported that the proposed approach has the following BD-rate and runtime impact when implemented in ECM-9:

AI { -0.04%, 0.00%, -0.01%, 99%, 97%}

RA { -0.06%, -0.18%, -0.02%, 100%, 100%}

It is further reported that an alternative configuration with larger maximum number of stored bins (1024 instead of 512) per CTU and updating both shorter and longer window probability estimators has the following BD-rate and runtime impact:

AI { -0.05%, -0.03%, -0.04%, 99%, 100%}

RA { -0.07%, -0.18%, -0.07%, 101%, 100%}

Only cases are stored where bins are actually appearing in bitstream. Indices need to be stored along with the bin.

How much storage is needed at maximum? Roughly 6000 bits per CTU, for entire row of CTUs.

Would this cause some latency problem? Maximum of 32 bins per context needs to be done sequentially, update can be done in parallel for different contexts. Limitation of maximum number might be necessary. Analysis would be desirable how many updates are actually performed at certain bit rates.

Would it be effective in case of mixed content, where characteristics change? Statistically likely yes, according to proponents.

Investigate in EE. Also report about aspects mentioned above.

[JVET-AE0147](https://jvet-experts.org/doc_end_user/current_document.php?id=13110) AHG12: Updating context model parameters for ECM [R.-L. Liao, Y. Ye, J. Chen, X. Li (Alibaba)]

This contribution proposes to update context initialization parameters for I-slice. Specifically, the initial probabilities, window sizes and adaptive weights for I-slice context models are retrained and updated. It is reported that on top of ECM-9.0, the overall coding performance impact for {Y, U, V} is {-0.10%, 0.07%, 0.15%} for AI, {-0.01%, -0.01%, 0.03%} for RA and {0.05%, -0.22%, 0.32%} for LB.

The training script was used that was once used during in VVC development, and adapted to ECM accordingly. It was reported that currently, the script is not properly working for PB slices, likely due to temporal inheritance. Contexts are also more complex in ECM, so the modifications necessary may not be trivial.

It was questioned whether the gain in intra is justified, while loss occurs in LB. It is also observed that the gain is quite sequence dependent, some sequences with high.

It was reported that in case of VTM, also other content such as HDR was used.

Several experts expressed that it would be beneficial to update initialization regularly. However, the proposed approach does not seem to do it in an optimum way, and in particular not for PB slices yet.

Study in EE, provided that the modified script could be shared and if possible further improved.

Further study recommended for the application to PB slices.

[JVET-AE0252](https://jvet-experts.org/doc_end_user/current_document.php?id=13215) Crosscheck of JVET-AE0147 (AHG12: Updating context model parameters for ECM) [Z. Deng (Bytedance)] [late]

***Partitioning (1)***

[JVET-AE0082](https://jvet-experts.org/doc_end_user/current_document.php?id=13030) AHG12: Non-square quadtree partitioning [Y. Ahn, J. Nam, N. Park, J. Lim, S. Kim (LGE)]

In the current VVC and ECM, quadtree partitioning is applied to only square blocks while binary and ternary tree partitioning are allowed to be applied to both square and non-square blocks. This contribution allows quadtree partitioning is also applied to non-square blocks of which shapes are 2N×N and N×2N. The proposed partitioning method was implemented on top of ECM-9.0 and the summarized test results are shown as follows:

AI: -0.17%/-0.29%/-0.26% (Y/U/V), 137% (EncT), 100% (DecT)

RA: -0.67%/-0.92%/-0.82% (Y/U/V), 167% (EncT), 100% (DecT)

LDB: -0.31%/-0.90%/-0.54% (Y/U/V), 144% (EncT), 99% (DecT)

Minimum non-square QT sizes are 8x16, 16x8

No new block sizes that are non-existing.

As there are no new partition options, to show benefit, it should be compared with encoders using very similar encoding times. It was suggested to release in ECM the constraint for maximum MTT depth, which would increase the encoding time of ECM and would make the comparison more fair. However, to be attractive for adoption, the encoding time shall be in the ballpark of current ECM.

Currently, no fast decisions in the code.

Investigate in EE. See notes above about comparison.

[JVET-AE0206](https://jvet-experts.org/doc_end_user/current_document.php?id=13169) Crosscheck of JVET-AE0082 (AHG12: Non-square quadtree partitioning) [R. Yu (Ericsson)] [late]

***GDR (2)***

[JVET-AE0145](https://jvet-experts.org/doc_end_user/current_document.php?id=13108) AHG 12: Flexible GDR [L. Wang, S. Hong, K. Panusopone (Nokia)]

[JVET-AE0267](https://jvet-experts.org/doc_end_user/current_document.php?id=13230) Crosscheck of JVET-AE0145 (AHG 12: Flexible GDR) [J. Enhorn (Ericsson)] [late]

[JVET-AE0146](https://jvet-experts.org/doc_end_user/current_document.php?id=13109) AHG 12: Fixes H-CCP Table for CUs in Refreshed Areas of GDR/Recovering Pictures [S. Hong, L. Wang, K. Panusopone (Nokia)]

[JVET-AE0273](https://jvet-experts.org/doc_end_user/current_document.php?id=13236) Crosscheck of JVET-AE0146 (AHG 12: Fixes H-CCP Table for CUs in Refreshed Areas of GDR/Recovering Pictures) [L. Zhao (Bytedance)] [late] [miss]

***RPR (2)***

[JVET-AE0103](https://jvet-experts.org/doc_end_user/current_document.php?id=13051) Weighted Edge Enhancement Filtering for Picture Upscaling and RPR [T. Claßen, M. Wien (RWTH Aachen University)]

[JVET-AE0153](https://jvet-experts.org/doc_end_user/current_document.php?id=13116) Non-EE2: Enabling template-based inter tools for scaled reference pictures in the RPR [X. Xiu, H.-J. Jhu, C.-W. Kuo, C. Ma, N. Yan, W. Chen, X. Wang (Kwai)]

[JVET-AE0259](https://jvet-experts.org/doc_end_user/current_document.php?id=13222) Crosscheck of JVET-AE0153 (Non-EE2: Enabling template-based inter tools for scaled reference pictures in the RPR) [Z. Zhang (Qualcomm)] [late]

### CTC for EE2/ECM and ECM improvements (6)

Contributions in this area were discussed at XXXX–XXXX on XXday XX July 2023 (chaired by XX).

[JVET-AE0055](https://jvet-experts.org/doc_end_user/current_document.php?id=13003) AHG12: Sign prediction parameter configuration for low-delay conditions [Y. Yasugi, T. Ikai (Sharp)]

[JVET-AE0271](https://jvet-experts.org/doc_end_user/current_document.php?id=13234) Crosscheck of JVET-AE0055 (AHG12: Sign prediction parameter configuration for low-delay conditions) [Y.-J. Chang (Qualcomm)] [late]

[JVET-AE0133](https://jvet-experts.org/doc_end_user/current_document.php?id=13096) AhG12: On EE2 Common test conditions change based on EE2 Test 1.1 [G. Laroche, P. Onno (Canon)]

v1 of document does not come with any results – placeholder status?

[JVET-AE0163](https://jvet-experts.org/doc_end_user/current_document.php?id=13126) AHG12: modified CTC proposal for low-delay configuration [S. Puri, C. Bonnineau, F. Le Léannec, T. Poirier, E. François (InterDigital)]

[JVET-AE0264](https://jvet-experts.org/doc_end_user/current_document.php?id=13227) Crosscheck of JVET-AE0163 (AHG12: Modified CTC proposal for low-delay configuration) [H. Huang (Qualcomm)] [late] [miss]

[JVET-AE0180](https://jvet-experts.org/doc_end_user/current_document.php?id=13143) On ECM SW memory consumption [R. Chernyak, S. Liu (Tencent), Y. Yasugi, T. Ikai (Sharp)]

[JVET-AE0192](https://jvet-experts.org/doc_end_user/current_document.php?id=13155) Crosscheck JVET-AE0180 (On ECM SW memory consumption) [C.-W. Kuo, X. Xiu (Kwai)] [late]

[JVET-AE0257](https://jvet-experts.org/doc_end_user/current_document.php?id=13220) Non-EE2: Peak Memory Reduction for ECM [W. Yin, K. Zhang, L. Zhang (Bytedance)] [late]

[JVET-AE0274](https://jvet-experts.org/doc_end_user/current_document.php?id=13237) Crosscheck of JVET-AE0257 (Non-EE2: Peak Memory Reduction for ECM) [Y. Yasugi (Sharp)] [late] [miss]

[JVET-AE0278](https://jvet-experts.org/doc_end_user/current_document.php?id=13241) AHG6/Non-EE2: Suggestions on memory reduction for ECM encoder [N. Hu, V. Seregin, M. Karczewicz (Qualcomm)] [late]

# High-level syntax (HLS) and related proposals (41)

## AHG9: SEI messages on neural-network post filter (28)

Contributions in this area were discussed in BoG JVET-AE0272 (those related to v3 VVC/VSEI), and otherwise at 0830–1215 on Friday 14 July 2023 (chaired by JRO).

[JVET-AE018](https://www.jvet-experts.org/doc_end_user/current_document.php?id=13150)7 AHG9: A summary of SEI proposals on NNPF [Y.-K. Wang (Bytedance)] [late]

[JVET-AE0048](https://jvet-experts.org/doc_end_user/current_document.php?id=12996) AHG9: Miscellaneous VSEI changes on neural-network post-filter SEI messages [M. M. Hannuksela, F. Cricri, M. Santamaria (Nokia)]

[JVET-AE0049](https://jvet-experts.org/doc_end_user/current_document.php?id=12997) AHG9: Miscellaneous VVC changes on neural-network post-filter SEI messages [M. M. Hannuksela, F. Cricri, M. Santamaria (Nokia)]

[JVET-AE0050](https://jvet-experts.org/doc_end_user/current_document.php?id=12998) AHG9: On NNPF input picture selection [M. M. Hannuksela, F. Cricri (Nokia)]

[JVET-AE0051](https://jvet-experts.org/doc_end_user/current_document.php?id=12999) AHG9: On persistent NNPF activation [M. M. Hannuksela, F. Cricri (Nokia)]

[JVET-AE0052](https://jvet-experts.org/doc_end_user/current_document.php?id=13000) AHG9: NNPF cascades and alternatives [M. M. Hannuksela, F. Cricri (Nokia)] [late]

Discussed on Friday July 14.

This input document proposes the following aspects to be included in a working draft of a VSEI amendment:

1. NNPF group characteristics (NNPFGC) SEI message, which can be used to define either a cascade of NNPFs or indicate a set of alternative NNPFs or alternative NNPF cascades.
2. NNPF group activation (NNPFGA) SEI message, which is used for activating an NNPF cascade.
3. An update of the general post-processing filtering process using NNPFs to take into account the proposed NNPFGC and NNPFGA SEI messages.

NNPF group characteristics (NNPFGC) SEI message contains:

* 1. An identifier of the NNPF group
  2. A grouping type of the NNPF group, where 0 indicates a cascade of NNPFs and 1 indicates alternative NNPFs or alternative NNPF cascades
  3. The purpose of the NNPF group, the presence of which is conditioned by the grouping type indicating a cascade of NNPFs.

Note: When the grouping type indicates alternative filters, their purpose is available in the NNPFC SEI messages defining the filters.

* 1. The number of members in the NNPF group, where each member can be an NNPF or an NNPF group of a cascade of NNPFs
  2. The identifiers of the NNPFs or NNPF groups belonging to this NNPF group
  3. A flag indicating if the complexity information for the NNPF group is present
  4. Complexity information specified like in the NNPFC SEI message

NNPFGA SEI message contains:

1. nnpfga\_target\_id, nnpfga\_cancel\_flag, and nnpfga\_persistence\_flag specified like nnpfa\_target\_id, nnpfa\_cancel\_flag, and nnpfa\_persistence\_flag, respectively, but applying to an NNPF group rather than to a single NNPF.
2. The number of filters in the NNPF group
3. For each filter in the NNPF group, the NNPFGA SEI message contains:
   * 1. nnpfga\_target\_base\_flag[ i ] specified for the i-th NNPF in the NNPF group like nnpfa\_target\_base\_flag
     2. An indication if the input pictures for the i-th NNPF in the NNPF group are formed from the list of candidate input pictures without skipping. If not, the following is indicated:
        1. The number of input pictures for the i-th NNPF in the NNPF group
        2. The picture count that is skipped in the list of candidate input pictures when selecting input pictures for the i-th NNPF in the NNPF group
4. nnpfga\_num\_output\_entries[ i ] and nnpfga\_output\_flag[ i ][ j ] specified for the i-th NNPF in the NNPF group like nnpfa\_num\_output\_entries and nnpfa\_output\_flag[ j ]

The general post-processing filtering process using NNPFs is proposed to be updated as follows:

* 1. It is clarified that the bitstream may be filtered multiple times to produce different results.
  2. Selecting between alternatives indicated in an NNPFGC SEI message is added.
  3. Applying an NNPF cascade as indicated in NNPFGC and NNPFGA SEI messages is added.

Relation with SEI processing order SEI? For NNPF, specific information is needed, for example, about the complexity, input/output signalling of temporal picture positions. Another aspect is that the NNPFGC can offer several alternative groups, and the receiver selects one.

It was pointed out that the example of hierarchical temporal upsampling in a cascade may not be used, as no new information is added to the intermediate picture. This example is the only one from the contribution which requires signalling of picture position. There could be other examples where several pictures are used as input to a subsequent neural network.

It was asked if the needed functionality (signalling of temporal position of input/output, alternative groups, operation) could also be achieved by extending the concept of SEI processing order SEI, rather than defining it specifically for the NNPF case.

[JVET-AE0053](https://jvet-experts.org/doc_end_user/current_document.php?id=13001) AHG8/AHG9: Neural-network post-filter regions SEI message [T. Chujoh, Y. Yasugi, T. Ikai (Sharp)]

Discussed on Friday July 14.

The neural-network post-filter characteristic and activation SEI messages process picture by picture. In several past meetings, there were several proposals to activate regions in the picture, and it has been recommended to use the annotated regions SEI message, which carries parameters that identify annotated regions using bounding boxes representing the size and location of identified objects. However, information on using regions and labels specified in the annotated regions SEI message for post-processing is not defined. Therefore, this contribution considers a minimum specification of how to post-process regions and introduces a new SEI message indicating an ID for the annotated regions SEI message. This proposal is expected to apply to the VSEI (V4) and is useful not only for ROI but also encoders and receiving systems for machine analysis of coded video content.

It was commented that region based activation of NNPF is a useful functionality. However, the combination with annotated regions SEI message may not be the best way to implement this, as its original purpose is different (more in the semantics of content, rather than processing). Furthermore, such an SEI message should support multiple regions, and also using different NNPF for different regions.

The functionality of region based processing is not supported in v3, and would be interesting for a future version. Further study recommended.

[JVET-AE0060](https://jvet-experts.org/doc_end_user/current_document.php?id=13008) [AHG9] Comments on NNPFC [S. Deshpande (Sharp)]

JVET-AE0060 item 2 was further discussed on Friday July 14.

Some modifications are proposed to the Neural-Network Post-Filter Characteristics SEI Message. Following is proposed:

* Proposal 1: (discussed in BoG)
* Proposal 2: It is proposed to add decolourization (converting colour video to monochrome) as a new nnpfc\_purpose.

It was asked why this would not be done at the encoder side

It would also be possible to use colourization, which also has three components as input, and produces an output which has zero chroma.

No evidence that this is needed.

[JVET-AE0061](https://jvet-experts.org/doc_end_user/current_document.php?id=13009) [AHG9] On NNPFC Application Purpose [S. Deshpande (Sharp)]

Item 1 was discussed on Friday July 14.

Some modifications are proposed to the Neural-Network Post-Filter Characteristics SEI Message related to the application determined purpose. Following is proposed:

* Proposal 1: It is proposed to signal a new syntax element nnpfc\_application\_purpose\_tag\_uri when nnpfc\_purpose is equal to 0 (i.e. when the NNPF may be used as determined by the application) to provide more details about the application-specific purpose.

It was commented that the presence of the **nnpfc\_application\_purpose\_tag\_uri** syntax element should be gated by a flag, to allow future extensions.

Decision: Include in “TuC for VSEI”, with additional flag.

[JVET-AE0062](https://jvet-experts.org/doc_end_user/current_document.php?id=13010) [AHG9] On NNPF for Deinterlacing [A. Sidiya, S. Deshpande (Sharp)]

Discussed on Friday July 14.

It is proposed to allow specifying deinterlacing as a new NNPF purpose. For this following is proposed:

1. It is proposed to signal a flag that indicates the number of output pictures for each interlaced input picture to the NNPF.
2. It is proposed to signal if the input needs to be flipped so that a neural network trained on top field first can work on bottom field first content and vice-versa.

It was commented that this would be beneficial for legacy AVC content, which is valuable. In HEVC and VVC, deinterlacing would mostly be done before encoding.

It was commented that rotating would be inappropriate for the top-/bottom-field first problem, as also within the lines the scan from left to right would be converted right-to-left. It might be more general to use different networks for the two cases.

It was asked whether more than a pair of fields can be used as input. This is the case, as in the current NNPF for frame interpolation. It was reported that the rotation method also works in case of more than 2 fields as input.

It was commented that the necessity of having this is not clear, as de-interlacing is typically done in a display when the decoder output is in interlaced formats. What would be use cases that a content provider would send information about a recommended de-interlacing algorithm?

Further study.

[JVET-AE0063](https://jvet-experts.org/doc_end_user/current_document.php?id=13011) [AHG9] On operations for multiple NNPFs [L. Chen, O. Chubach, Y.-W. Huang, S. Lei (MediaTek)]

Discussed on Friday July 14.

This contribution proposes an SEI message indicating multiple NNPFs as a group/list and related operations for a picture in response to the four questions on item 1)c in JVET-AD0362. The related operations for the indicated multiple NNPFs consist of four types of cascade, alternate, coherence and optional.

(coherence=parallel)

Conceptually similar to JVET-AE0052. Parallel is not explicitly in 0052. Another aspect is that the encoder can take provision that only a single filter is activated at one time. No specific group activation SEI is needed in JVET-AE0063

Parallel operation might, for example, be useful to generate attention mechanism input for a subsequent network. This is something that would not appear in other SEI messages.

Use cases for parallel operation need more study, including the question how many could be operated in parallel.

Further study required if it is possible to bring those aspects into the SEI processing order SEI message (in VVC), or if this is so specific to NNPF that it would be a burden for implementers who only want to use SEI processing order for simpler cases of SEI messages.

It was agreed to produce an output document “Technology under consideration for future extensions of VSEI”. Proponents of JVET-AE0052 and JVET-AE0063 should combine the aspects from both contributions, and propose a text part.

It was also commented that a software implementation demonstrating different options would be desirable.

[JVET-AE0068](https://jvet-experts.org/doc_end_user/current_document.php?id=13016) AHG9: On extensibility of purpose syntax element in NNPFC SEI message [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

[JVET-AE0069](https://jvet-experts.org/doc_end_user/current_document.php?id=13017) AHG9: On the signalling of output pictures in NNPFA SEI message [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

[JVET-AE0070](https://jvet-experts.org/doc_end_user/current_document.php?id=13018) AHG9: On input pictures that are not present in the bitstream for NNPF SEI messages [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

[JVET-AE0101](https://jvet-experts.org/doc_end_user/current_document.php?id=13049) [AHG2][AHG9] Neural network post filter and phase indication SEI messages for AVC and HEVC [T. Ikai, T. Chujoh (Sharp), Y.-K. Wang, J. Xu, W. Jia (Bytedance)]

Discussed on Friday July 14.

Neural network post filter (NNPF) SEI messages, the NNPF characteristic SEI and the NNPF activation SEI, and phase indication SEI messages been developed and incorporated into the VVC draft text and VSEI draft text. On the other hand, AVC and HEVC have not defined those SEI messages in the specifications while they have been widely deployed in various applications [1] and those SEI messages are applicable if the specifications have the corresponding payload type numbers and interface definitions using the normative references. It is proposed to support the NNPF and phase indication SEI messages in the AVC and HEVC specifications.

A new edition of H.265 is planned to be submitted for ITU consent, but it would be premature to include by last minute, and the corresponding edition 5 of 23008-2 is already under FDIS ballot.

According to opening remarks, another edition of H.265 could be submitted in April 2024, and the three SEI messages (NNPFC, NNPFA, phase indication) should be enabled in that. It was agreed to implement this by establishing a link to VSEI (v3). Another edition of 23008-2 could be initiated for 2024 once it becomes clear that the 5th ed. would be published in 2023. At this moment, an amendment (CDAM1) on top of the 5th edition could be started, which would contain the colour descriptors from JVET-AD1008, and the proposed text from JVET-AE0101 (possibly some refinements). Decision: This was agreed.

According to opening remarks, another edition of H.264 could be submitted in April 2024. Furthermore, it is planned to start working on the 11th edition of 14496-10, with CD at this meeting. This should include various corrigenda items from JVET-AD1004, the colour descriptors from JVET-AD1008, and the proposed text from JVET-AE0101 (possibly some refinements). This was

Further study should be performed if other SEI messages existing in HEVC and VSEI should also be included in AVC. Decision: This was agreed.

[JVET-AE0106](https://jvet-experts.org/doc_end_user/current_document.php?id=13054) AHG9: On missing value ranges for some syntax elements in the NNPFC SEI message [C. Lin, Y.-K. Wang, J. Xu, W. Jia, J. Li, Y. Li, K. Zhang, L. Zhang (Bytedance)] [late]

[JVET-AE0113](https://jvet-experts.org/doc_end_user/current_document.php?id=13061) AHG9: Extendibility and code word length of nnpfc\_purpose [R. Sjöberg, M. Pettersson (Ericsson)]

[JVET-AE0126](https://jvet-experts.org/doc_end_user/current_document.php?id=13081) AHG9: NNPF cleanup and editorial changes for VSEI [Y.-K. Wang, W. Jia, J. Xu, C. Lin (Bytedance)]

[JVET-AE0127](https://jvet-experts.org/doc_end_user/current_document.php?id=13083) AHG9: NNPF editorial changes for VVC [Y.-K. Wang, W. Jia, J. Xu (Bytedance)]

[JVET-AE0128](https://jvet-experts.org/doc_end_user/current_document.php?id=13084) AHG9: On NNPFC extensibility and base filter signalling [Y.-K. Wang (Bytedance)]

[JVET-AE0134](https://jvet-experts.org/doc_end_user/current_document.php?id=13097) AHG9: Align the design of NNPF with multiple input pictures to NNPF including picture rate upsampling [J. Xu, Y.-K. Wang (Bytedance)]

[JVET-AE0135](https://jvet-experts.org/doc_end_user/current_document.php?id=13098) AHG9: On NNPF picture rate upsampling constraints [J. Xu, Y.-K. Wang (Bytedance)]

[JVET-AE0141](https://jvet-experts.org/doc_end_user/current_document.php?id=13104) AHG9: Fix a bug in NNPFC SEI message for colourization [J. Xu, Y.-K. Wang (Bytedance)]

[JVET-AE0142](https://jvet-experts.org/doc_end_user/current_document.php?id=13105) AHG9: On derivation of NNPF input pictures and the value of nnpfc\_purpose [W. Jia, Y.-K. Wang, J. Xu, L. Zhang (Bytedance)]

[JVET-AE0173](https://jvet-experts.org/doc_end_user/current_document.php?id=13136) [AHG9] Clarification and improvements of signaling of NNPF update [Y. Lim (Samsung)]

[JVET-AE0175](https://jvet-experts.org/doc_end_user/current_document.php?id=13138) [AHG9] Editorial improvements of nnpfc\_mode\_idc [Y. Lim (Samsung)]

[JVET-AE0189](https://jvet-experts.org/doc_end_user/current_document.php?id=13152) AHG9: On the design of nnpfa\_target\_base\_flag in NNPFA SEI message [Hendry (LGE)] [late]

[JVET-AE0279](https://jvet-experts.org/doc_end_user/current_document.php?id=13242) AHG9: New nnpfc\_mode\_idc for reuse of NNPF definition for new purpose [Y. Lim (Samsung)] [late]

Presented on Friday July 14

An NNPF can be associated with a single purpose and ID within a particular CLVS regardless of its capability. Sometime it might be efficient and useful to reuse NNPF data with new purpose and ID. This contribution proposes a new nnpfc\_mode\_idc for referencing an NNPF already defined with a purpose and ID for new purpose and ID without sending the data again.

(related to discussion regarding disallowing change of purpose, and “purpose” vs. “capability” as per BoG report JVET-AE0172 v2)

Several experts see some inconsistency w.r.t. the concept of updating vs. base.

Further study.

## AHG9: SEI messages on topics other than NNPF (13)

Contributions in this area were discussed at 1215–XXXX on Friday 14 July 2023 (chaired by JRO).

[JVET-AE0054](https://jvet-experts.org/doc_end_user/current_document.php?id=13002) AHG2/AHG9: Editorial improvements of annotated regions SEI message [T. Chujoh, T. Aono, T. Ikai (Sharp)]

Addressed by the BoG on SEI messages for VSEI v3 and VVC v3.

[JVET-AE0064](https://jvet-experts.org/doc_end_user/current_document.php?id=13012) AHG8/AHG9: Signalling encoder preprocessing and human / machine viewing indications [C. Kim, D. Gwak, Hendry, J. Lim, S. Kim (LGE), M. M. Hannuksela, F. Cricri, H. Zhang (Nokia)]

[JVET-AE0066](https://jvet-experts.org/doc_end_user/current_document.php?id=13014) AHG9: Multiplane Image Information SEI [T. Lu, P. Yin, G. Su, D. Lee, T. Huang, S. McCarthy, W. Husak, G. J. Sullivan (Dolby)]

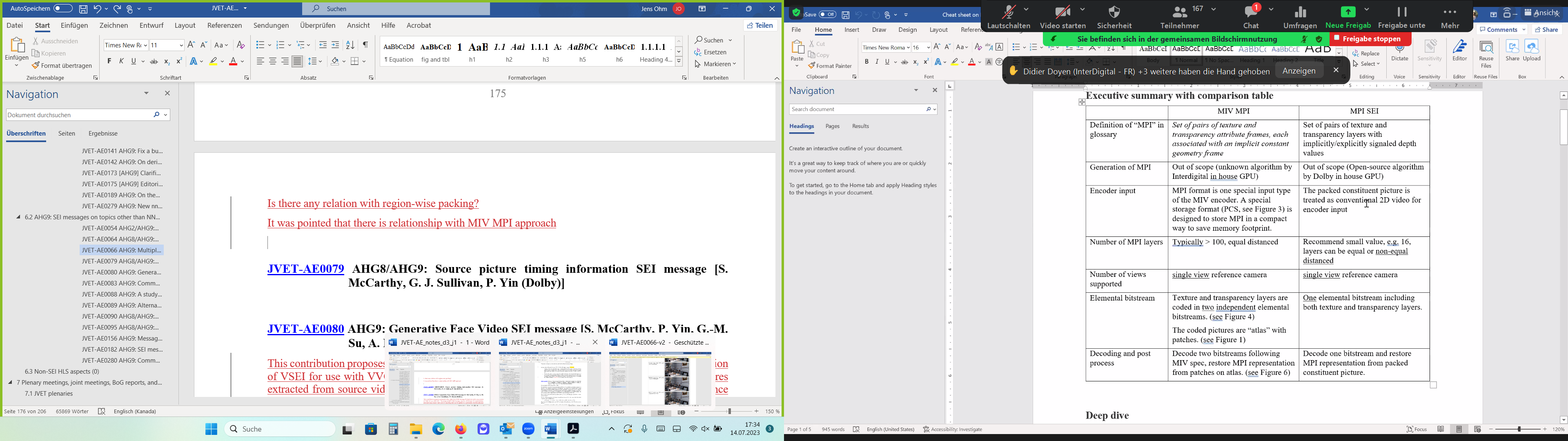
This contribution proposes a multiplane image information (MPII) SEI message be included in the next version of VSEI. The proposed MPII SEI message is intended to enable virtual reality/immersive experiences on existing consumer devices using existing video codecs such as HEVC and VVC that were developed for conventional two-dimensional video.

Up to 16 planes

Alpha map plus video

Is there any relation with region-wise packing?

It was asked if there is relationship with MIV MPI approach? The following table was shown on this aspect:



It was requested to put this information into an update of the contribution.

Application scope? Mainly consumer

Following comments were made in detail on syntax (YK Wang):

* Is view id useful in single layer bitstream
* Usage of ue(v)
* Usage of cancel flag
* Some details

Otherwise, editorial quality sufficient.

Communication with WG 4 should be sought on this (joint meeting). MIV supports both single-view and mult-view MPI.

Revisit.

[JVET-AE0079](https://jvet-experts.org/doc_end_user/current_document.php?id=13027) AHG8/AHG9: Source picture timing information SEI message [S. McCarthy, G. J. Sullivan, P. Yin (Dolby)]

[JVET-AE0080](https://jvet-experts.org/doc_end_user/current_document.php?id=13028) AHG9: Generative Face Video SEI message [S. McCarthy, P. Yin, G.-M. Su, A. K. Choudhury, W. Husak (Dolby)]

This contribution proposes a generative face video (GFV) SEI message for inclusion in the next version of VSEI for use with VVC. The proposed GFV SEI message conveys information for facial features extracted from source video prior to encoding that can be applied to a previously decoded reference picture to reconstruct a face picture at the current time. The proposed GFV SEI message also specifies a neural network, for example, a generative adversarial network (GAN), that could be used to generate face pictures based on the face feature information conveyed in the GFV SEI message. As noted in previous related JVET contributions JVET-AC0088 and JVET-AD0051, the GFV SEI message is intended to enable face picture reconstruction processes to use VVC-coded pictures as reference pictures. It is also noted that extracting face features on the encoder-side can reduce receiver-side complexity. The present contribution proposes additions and refinements to previously proposed syntax and semantics to address issues noted by experts during previous JVET meetings. The proposed GFV SEI message is intended for ultra low bitrate face video compression applications such as video conferencing, live entertainment, and face animation.

When a receiver ignores the SEI message, the base image would be shown.

It was asked whether this would not rather be something to standardize as a specific standard rather than SEI message. The benefit of the technology is well recognized, very low bit rate for video conferencing.

One expert mentions that this technology has a very limited scope. The proponents say this is intentional.

Is it an advantage to couple this to VVC/HEVC? If somebody is interested only in this functionality, they might not want to implement a full VVC decoder.

Would it not be better to define a profile if the intent is limited scope?

The relationship with VPC is also pointed out, but this is more generic, not only using VVC as base codec. According to last meeting’s report, WG 7 also seems to have some plans about Avatar coding. Unclear if that is based on neural networks or old MPEG-4 technology. Joint meeting?

Usage of other legacy codecs is also important in the opinion of another expert.

Complexity issue? It was pointed out that real-time implementations on smartphone exist, which also fulfills real-time requirements (see https://arxiv.org/abs/2012.00328)

One expert suggested that an exploration experiment might be started.

Interoperability between encoder and decoder could be important in this context – can this be provided by an SEI message? This role could be taken over by application standards who could define a specific encoder/decoder pair.

[JVET-AE0083](https://jvet-experts.org/doc_end_user/current_document.php?id=13031) AHG9: Common SEI Message of Generative Face Video [B. Chen, J. Chen, Y. Ye (Alibaba), S. Wang (CityU)]

At the 29th JVET meeting, JVET-AC0088 proposed a generative face video SEI message to allow VVC-coded pictures to be used as base pictures to drive a deep generative network to generate additional images, therefore enabling ultra low rate face video compression. At the 30th JVET meeting, JVET-AD0051 generalized the generative face video SEI message to include syntax elements that correspond to a variety of facial representations, such as 2D keypoints, 2D landmarks, 3D keypoints, facial semantics and other formats.

Compared with previous contributions, in this contribution the following changes were made

1. refine the syntax and semantics of the proposed SEI message
2. define the interface between the decoder (which decodes the base pictures and parses the SEI message) and the generative neural network
3. provide further information on available software implementations.

Powerpoint presentation to be provided.

Dummy picture: Not sent, decoder uses previous base picture

Mechanism to signal kind of algorithm unclear

“undefined matrix” vague

It was questioned whether an SEI message is an appropriate mechanism to generate an output video.

A software implementation is available and was demonstrated.

[JVET-AE0088](https://jvet-experts.org/doc_end_user/current_document.php?id=13036) AHG9: A study on Generative Face Video SEI Message [H.-B. Teo, J.-Y Thong, K. Jayashree, C.-S. Lim, K. Abe (Panasonic)]

A generative face video SEI message was presented in JVET-AC0088 at the 29th JVET meeting for compression of talking-face videos at ultra low bitrate, where at the 30th JVET meeting JVET-AD0051 further expanded its scope to encompass multiple facial representations. This contribution demonstrates the effectiveness of 3-D facial landmarks for generating talking-face videos. This document further highlights the framework’s independence of number of landmarks used for generation, and its interoperability of landmark representations across various systems.

[JVET-AE0089](https://jvet-experts.org/doc_end_user/current_document.php?id=13037) AHG9: Alternative Output Timing Hint SEI [H.-B. Teo, J. Gao, C.-S. Lim, K. Abe (Panasonic)]

[JVET-AE0090](https://jvet-experts.org/doc_end_user/current_document.php?id=13038) AHG8/AHG9: On machine vision indication [J. Gao, H.-B. Teo, C.-S. Lim, K. Abe (Panasonic)]

[JVET-AE0095](https://jvet-experts.org/doc_end_user/current_document.php?id=13043) AHG8/AHG9: proposed changes to the candidate new object mask information SEI message [P. de Lagrange, E. François, D. Doyen (InterDigital), J. Chen, S. Wang, Y. Ye (Alibaba)] [late]

[JVET-AE0156](https://jvet-experts.org/doc_end_user/current_document.php?id=13119) AHG9: Message wrapping and importance indication for the SEI processing order SEI message [G. J. Sullivan, S. McCarthy, P. Yin (Dolby Labs)]

[JVET-AE0182](https://jvet-experts.org/doc_end_user/current_document.php?id=13145) AHG9: SEI message extension of VVC for object-wave compression and computer-generated hologram use [K. Nonaka, R. Koiso, H. Kojima, K. Kawamura, H. Kato (KDDI)]

[JVET-AE0280](https://jvet-experts.org/doc_end_user/current_document.php?id=13243) AHG9: Common text for proposed generative face video SEI message [B. Chen, J. Chen, Y. Ye (Alibaba), S. Wang (CityU), S. McCarthy, P. Yin, G.-M. Su, A. K. Choudhury, W. Husak (Dolby)] [late]

This contribution proposes common text based on JVET-AE0080 and JVET-AE0083 for a generative face video (GFV) SEI message for inclusion in the next version of VSEI. The proposed GFV SEI message conveys information for facial features extracted from source video prior to encoding that can be applied to a previously decoded reference picture to reconstruct a face picture at the current time. The proposed GFV SEI message is intended for ultra low bitrate face video compression applications such as video conferencing, live entertainment, and face animation.

The application of translating a real person into an avatar does not require precise encoder/decoder match, synthesis models would translate the parametric description into the avatar’s movements. Ultimately, a decoder would need to be able to interpret different combinations of parameters.

The application of “video realistic” generation of video would require precise match of encoder/decoder. The examples shown in the demo are not fully reconstructing all movements such as eye closing (at 2 kbit/s), but could be adjusted with increased number of landmarks, etc.

Request for joint meeting with WG2/WG7/VCEG on Monday.

Revisit.

## Non-SEI HLS aspects (0)

Section kept as a template for future use.

# Plenary meetings, joint meetings, BoG reports, and liaison communications

## JVET plenaries

No intermediate plenaries were held, as document review and decisions were made in single-track mode at this meeting (with some BoG activity as noted). Further detail on scheduling is recorded in section 2.15.

Communication and coordination items were discussed on XXday July XX at XXXX.

* Initial joint meetings planned were as follows:
  + …
* Session planning (participants were encouraged to see the posted online calendar)
* Output document / deliverables planning
* Liaison …

General plenary wrap-up discussions are recorded under sections 8, 9, and 10.

## Information sharing meetings

Information sharing sessions with other WGs and AGs of the MPEG community were held on Monday 17 July 0900–1130, Wednesday 19 July 0900–1030, and Friday 21 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 with MPEG … on …

This joint session was held on XXday XX July at XXXX-XXXX.

…

## BoGs (3)

The following break-out groups were established at this meeting to conduct discussion and develop recommendations on particular subjects.

[JVET-AE0272](https://jvet-experts.org/doc_end_user/current_document.php?id=13235) BoG report on SEI messages for VSEI v3 and VVC v3 [G. J. Sullivan (BoG coordinator)]

This document is a break-out-group report on SEI messages for VSEI v3 and VVC v3. The BoG met from 15:15 to 18:30 on Tuesday 11 July 2023 and from 08:30 to 12:30 and 14:30 to 18:30 on Wednesday 12 July 2023 in Salle 3 of the Geneva CICG facility.

The BoG reviewed the relevant proposals for NNPF SEI messages based primarily on the summary of NNPF-related SEI messages provided in JVET-AE0187-v3, which formed the original basis of the structure of this report.

No specific plans were recommended for topics that were agreed not to be within the scope of VSEI v3 and VVC v3 (except to encourage progress toward future support for VSEI v3 SEI messages for NNPF and filtering phase indication in HEVC and AVC).

Actions recommended by the BoG are highlighted in this report. Five items are marked to revisit for further discussion; these five are suggested to be relatively minor and may not require a further BoG session (if the BoG scope is not expanded to consider additional topics).

The first version of the BoG report was presented in JVET plenary on Thursday July 13 at 0830. The changes made during discussion in the JVET plenary are reflected in version 2 of JVET-AE0272.

All BoG recommendations as per v2 were confirmed by JVET plenary.

[JVET-AE0282](https://jvet-experts.org/doc_end_user/current_document.php?id=13245) BoG report on NNVC work planning [[E. Alshina](mailto:elena.alshina@huawei.com), [F. Galpin](mailto:franck.galpin@interdigital.com) (BoG coordinators)]

Mandate of this BoG is preparing the next EE1 on NNVC and review documents (from section 5.1 JVET meeting notes) relevant to EE1 planning.

BoG met twice on July, 13 (Thursday) 14:30, and on July, 14 (Friday) 09:00. Up to 81 experts attended second BoG discussion in person or remotely.

In total eight contributions have been reviewed on July 13 ([JVET-AE0072](https://jvet-experts.org/doc_end_user/current_document.php?id=13020), [JVET-AE0093](https://jvet-experts.org/doc_end_user/current_document.php?id=13041), [JVET-AE0161](https://jvet-experts.org/doc_end_user/current_document.php?id=13124), [JVET-AE0238](https://jvet-experts.org/doc_end_user/current_document.php?id=13201), [JVET-AE0171](https://jvet-experts.org/doc_end_user/current_document.php?id=13134), [JVET-AE0160](https://jvet-experts.org/doc_end_user/current_document.php?id=13123), [JVET-AE0255](https://jvet-experts.org/doc_end_user/current_document.php?id=13218), [JVET-AE0164](https://jvet-experts.org/doc_end_user/current_document.php?id=13127)).

Five more contributions have been reviewed on July 14 ([JVET-AE0119](https://jvet-experts.org/doc_end_user/current_document.php?id=13067), [JVET-AE0162](https://jvet-experts.org/doc_end_user/current_document.php?id=13125) , [JVET-AE0230](https://jvet-experts.org/doc_end_user/current_document.php?id=13193), [JVET-AE0194](https://jvet-experts.org/doc_end_user/current_document.php?id=13157), JVET-AE0281.

BoG made following recommendations:

1. Adopt JVET-AE0191 to NNVC-6.0 as first generation of unified filter design for high operation point HOP.1 (disabled by default). Keep study of alternative HOP designs as EE1 tests. No joint training for HOP expected this time. The strongest candidate for joint HOP training to be decided based on EE1 results during JVET-AF meeting.
2. Use existing low operation point filter in NNVC-5.0 as first generation of LOP (LOP.1).
3. AhG11, EE1 (and non-EE1) proposals which study filter usage aspects (such as residual adjustment, input/output rotation, overfitting with signalling), or additional functionalities (such as considering filter in RDO, extra models dealing with input from different POC) must use HOP.1 or/and LOP.1 as base.
4. All EE1 tests which were not conducted due to delay of HOP.1 to be included into next EE1. Test description to be provided to EE1 coordinators before end of the day Monday, July 17.
5. Document with training instructions, which allow reproduction of HOP.1, starting from Stage I, or Stage II, or Stage III, with the link to intermediate models needed for reproducing the training, to be provided by proponent of JVET-AE0191 and corresponding text to be included to the next EE1 description (later can be considered for NNVC CTC).
6. Conduct AhG11/AhG14/EE1 teleconference on August 2, reviewing HOP.1 after Stage III training, model quantization and test (all-intra and random access cfg) are completed. JVET-AE0191 authors requested to report Stage III results (float and integer 16) before August, 2. The best performing model will be selected during teleconference. Publish test results of HOP.1 in EE1 description as anchor for HOP tests.
7. To train unified design for LOP jointly based on design and procedure described in [JVET-AE0281](https://jvet-experts.org/doc_end_user/current_document.php?id=13244): filter architecture design - combination of [JVET-AE0067](https://jvet-experts.org/doc_end_user/current_document.php?id=13015) (split aspect) and [JVET-AE0165](https://jvet-experts.org/doc_end_user/current_document.php?id=13128) (simplified Back Bone Block); training - [JVET-AE0165](https://jvet-experts.org/doc_end_user/current_document.php?id=13128) (aligned with HOP except number of epochs); filter usage – [JVET-AE0238](https://jvet-experts.org/doc_end_user/current_document.php?id=13201) (aligned with HOP).
8. Establish following tests in EE1:
   * EE1-0.x: Unified LOP ([JVET-AE0281](https://jvet-experts.org/doc_end_user/current_document.php?id=13244));
   * EE1-1.1.x category: architectural changes for HOP ([JVET-AE0160](https://jvet-experts.org/doc_end_user/current_document.php?id=13123) , [JVET-AE0164](https://jvet-experts.org/doc_end_user/current_document.php?id=13127)), more tests from old EE1 to be added.
   * EE1-1.2.x category: filter usage aspects ([JVET-AE0072](https://jvet-experts.org/doc_end_user/current_document.php?id=13020), [JVET-AE0093](https://jvet-experts.org/doc_end_user/current_document.php?id=13041), [JVET-AE0161](https://jvet-experts.org/doc_end_user/current_document.php?id=13124), [JVET-AE0238](https://jvet-experts.org/doc_end_user/current_document.php?id=13201)), more tests from old EE1 to be added.
   * EE1-1.3.x category: additional model for HOP (place holder for “temporal filter”???).
   * EE1-2.x NN-Inter ([JVET-AE0112](https://jvet-experts.org/doc_end_user/current_document.php?id=13060) investigating complexity reduction possibilities).
   * EE1-3.x NN-Intra ([JVET-AE0144](https://jvet-experts.org/doc_end_user/current_document.php?id=13107) with training in cross-check).
9. Adopt to NNVC SW MS-SSIM bug fix from [JVET-AE0162](https://jvet-experts.org/doc_end_user/current_document.php?id=13125). Bug affects only Super-Resolution of RPR tests. Suggest this bug fix to VTM, ECM.

BoG report v1 was presented in JVET plenary on Friday 14 July 1430-1530

All BoG recommendations above were confirmed by JVET.

It was agreed that JVET-AE0144 could be included in NNVC6, provided that the training crosscheck would be successfully completed by the date of the telco meeting (Aug. 2) (conditional adoption)

Decision (SW/BF): Adopt JVET-AE0162 bug fix for MS-SSIM in case of SR/RPR. Should be implemented in NNVC, ECM, and VTM.

## Liaison communications (1)

…

A reply was drafted by JVET as …. The draft reply was presented in the AG 3 meeting Thursday XX:XX (XXXX) and was reviewed in JVET on XXday XX July XXXX (, kept for future use).

# Project planning

## Software timeline (update)

ECM 10 software (including all adoptions) was planned to be available 3 weeks after the meeting.

The NNVC 6.0 codebase software (integrating “low” operation point loop filter) was planned to be available 2 weeks after the meeting. An update 5.1 (also including verified “high” operation point) was planned to be available after training verification.

VTM21.0 software was planned to be available on 2023-05-30. (Note that further updates may be released later)

Updates on top of HM17.0 software were not planned, but might be released after merging pending requests, as appropriate.

## Core experiment and exploration experiment planning

An EE on neural network-based video coding was established, as recorded in output document JVET-AE2023.

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-AE2024.

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:

* No review of normative contributions without draft specification text
* VTM algorithm description text is strongly encouraged for non-normative contributions
* Early upload deadline to enable substantial study prior to the meeting
* Using a clock timer to ensure efficient proposal presentations (5 min) and discussions

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.
* 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 is requested by contacting the CE coordinator.
* All software, results, and documents produced in the CE should be announced and made available to JVET in a timely manner.
* A JVET CE reflector will be established and announced on the main JVET reflector. Discussion of logistics arrangements, exchange of data, minor refinement of the test plans, and preparation of documents shall be conducted on the JVET CE reflector, with subject lines prefixed by “[CEx: ]”, where “x” is the number of the CE. All substantial communications about a CE other than such details shall take place on main JVET reflector. In the case that large amounts of data are to be distributed, it is recommended to send a link to the data rather than the data itself, or upload the data as an input contribution to the next meeting.

General timeline for CEs

T1= 3 weeks after the JVET meeting: To revise the CE description and refine questions to be answered. Questions should be discussed and agreed on JVET reflector. Any changes of planned tests after this time need to be announced and discussed on the JVET reflector. Initially assigned description numbers shall not be changed later. If a test is skipped, it is to be marked as “withdrawn”.

T2 = Test model software release + 2 weeks: Integration of all tools into a separate CE branch of the VTM is completed and announced to JVET reflector.

* Initial study by cross-checkers can begin.
* Proponents may continue to modify the software in this branch until T3.
* 3rd parties are encouraged to study and make contributions to the next meeting with proposed changes

T3: 3 weeks before the next JVET meeting or T2 + 1 week, whichever is later: Any changes to the CE test branches of the software must be frozen, so the cross-checkers can know exactly what they are cross-checking. A software version tag should be created at this time. The name of the cross-checkers and list of specific tests for each tool under study in the CE plan description shall be documented in an updated CE description by this time.

T4: Regular document deadline minus 1 week: CE contribution documents including specification text and complete test results shall be uploaded to the JVET document repository (particularly for proposals targeting to be promoted to the draft standard at the next meeting).

The CE summary reports shall be available by the regular contribution deadline. This shall include documentation about crosscheck of software, matching of CE description and confirmation of the appropriateness of the text change, as well as sufficient crosscheck results to create evidence about correctness (crosscheckers must send this information to the CE coordinator at least 3 days ahead of the document deadline). Furthermore, any deviations from the timelines above shall be documented. The numbers used in the summary report shall not be changed relative to the description document.

CE reports may contain additional information about tests of straightforward combinations of the identified technologies. Such supplemental testing needs to be clearly identified in the report if it was not part of the CE plan.

New branches may be created which combine two or more tools included in the CE document or the VTM (as applicable).

It is not necessary to formally name cross-checkers in the initial version of the CE description document. To adopt a proposed feature at the next meeting, JVET would like to see comprehensive cross-checking done, with analysis of whether the description matches the software, and a recommendation of the value of the tool and given tradeoffs.

The establishment of a CE does not indicate that a proposed technology is mature for adoption or that the testing conducted in the CE is fully adequate for assessing the merits of the technology, and a favourable outcome of CE does not indicate a need for adoption of the technology into a standard or test model.

Availability of specification text is important to have a detailed understanding of the technology and also to judge what its impact on the complexity of the specification will be. There must also be sufficient time to study this in detail. CE contributions without sufficiently mature draft specification text in the CE input document should not be considered for adoption.

Lists of participants in CE documents should be pruned to include only the active participants. Read access to software will be available to all members.

# Establishment of ad hoc groups

The ad hoc groups established to progress work on particular subject areas until the next meeting are described in the table below. The discussion list for all of these ad hoc groups was agreed to be the main JVET reflector ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de)).

Review of AHG plans was conducted during the plenary on XXday XX July 2023 at XXXX.

|  |  |  |
| --- | --- | --- |
| **Title and Email Reflector** | **Chairs** | **Mtg** |
| **Project Management (AHG1)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Coordinate overall JVET interim efforts. * Supervise AHG and experiment studies. * Report on project status to JVET reflector. * Provide a report to the next meeting on project coordination status. * Supervise processing and delivery of output documents | J.-R. Ohm (chair), G. J. Sullivan (vice‑chair) | N |
| **Draft text and test model algorithm description editing (AHG2)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Produce and finalize draft text outputs of the meeting (JVET-AD1003, JVET-AD1008, JVET-AD2002, JVET-AD2005, JVET-AD2006, and JVET-AD2027. * Collect reports of errata for the VVC, VSEI, HEVC, AVC, CICP, and the published related technical reports and produce the JVET-AD1004 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. * Discuss and make recommendations on the software development process. * Perform comparative tests of test model behaviour using common test conditions, including HDR, high bit depth and high bit rate. * Suggest configuration files for additional testing of tools. * Investigate how to minimize the number of separate codebases maintained for group reference software. * Coordinate with AHG on Draft text and test model algorithm description editing (AHG2) to identify any mismatches between software and text, and make further updates and cleanups to the software as appropriate. * Prepare drafts of merged and updated CTC documents for HM and VTM, as applicable. | F. Bossen, X. Li, K. Sühring (co-chairs), E. François, Y. He, K. Sharman, V. Seregin, A. Tourapis (vice‑chairs) | N |
| **Test material and visual assessment (AHG4)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Consider plans for additional verification testing of VVC capability, particularly target conducting a first test for VVC multi-layer features by the next meeting, and update the test plan according to subsequent tests. * 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. * Prepare availability of viewing equipment and facilities arrangements for future meetings. | V. Baroncini, T. Suzuki, M. Wien (co-chairs), S. Iwamura, S. Liu, S. Puri, A. Segall, P. Topiwala, S. Wenger, J. Xu, Y. Ye (vice-chairs) | Y (tel., 2 weeks notice) |
| **Conformance testing (AHG5)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Produce and finalize the draft of additional conformance bitstreams for VVC multilayer configurations JVET-AD2028, and investigate the need for future improvements of conformance testing specifications. * 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. | D. Rusanovskyy, I. Moccagatta (co-chairs), F. Bossen, K. Kawamura, T. Ikai, S. Iwamura, H.-J. Jhu, 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-9.0 software version and the reference configuration encodings according to the ECM common test conditions. * Investigate encoder speedup and other encoder software optimization. * 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, F. Le Léannec, K. Zhang (vice-chairs) | Y (tel., 2 weeks notice) |
| **ECM tool assessment (AHG7)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Investigate methodology of tool assessment. * Coordinate with AHG6 on resolving tool-off test related software issues (missing tool controls and software bugs). * Prepare configuration files and generate bitstreams and results of tool-on/tool-off testing. * Prepare reporting of tool assessment results. * Collect simulation results on non-CTC sequences, and report any issues identified with non-CTC sequences * Develop methodology of more reliable runtime measurement | X. Li (chair), L.-F. Chen, Z. Deng, J. Gan, E. François, H.-J. Jhu, X. Li, H. Wang (vice‑chairs) | 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. * Generate anchors according to the common test conditions JVET-AD2031. * Discuss improvements on the evaluation framework, including evaluation procedures and methodologies. * Coordinate software development, and investigate the possibility of migrating the software basis to newest VTM version. * Coordinate experiments on optimization of encoders and receiving systems for machine analysis of coded video content. * Maintain the software implementation example algorithms in the repository, including sufficient documentation in terms of operation and performance. * Evaluate proposed technologies and their suitability for machine analysis applications. * Propose improvements to the draft technical report JVET-AD2030 on optimization of encoders and receiving systems for machine analysis of coded video content. * Coordinate with WG 4 VCM AHG on common interests and activities such as common test conditions, test and training materials, and on studying characteristics and requirements of targeted machine analysis tasks, etc. | C. Hollmann, S. Liu, S. Wang, M. Zhou (AHG chairs) | Y (tel., 2 weeks notice) |
| **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. * 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. | S. McCarthy, Y.-K. Wang (co-chairs), T. Chujoh, S. Deshpande, C. Fogg, Hendry, P. de Lagrange, G. J. Sullivan, A. Tourapis, S. Wenger (vice-chairs) | N |
| **Encoding algorithm optimization (AHG10)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the impact of using techniques such as tool adaptation and configuration, and perceptually optimized adaptive quantization for encoder optimization. * Study the impact of non-normative techniques of preprocessing for the benefit of encoder optimization. * Study encoding techniques of optimization for objective quality metrics and their relationship to subjective quality. * Study optimized encoding for reference picture resampling and scalability modes in VTM. * Study optimized encoding and tool combinations for low latency and low complexity. * Consider neural network-based encoding optimization technologies for video coding standards. * Investigate other methods of improving objective and/or subjective quality, including adaptive coding structures and multi-pass encoding. * Study methods of rate control and rate-distortion optimization and their impact on performance, subjective and objective quality. * Study the potential of defining default or alternate software configuration settings and test conditions optimized for either subjective quality, or higher objective quality, and coordinate such efforts with AHG3 and AHG6. * Study the effect of varying configuration parameters depending on temporal layer, such as those related to deblocking, partitioning, chroma QP. | P. de Lagrange, A. Duenas, R. Sjöberg, A. Tourapis (AHG chairs) | N |
| **Neural network-based video coding (AHG11)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Evaluate and quantify the performance improvement potential of NN-based video coding technologies compared to existing video coding standards such as VVC, including both individual coding tools and novel architectures. * Discuss potential refinements of the test conditions for NN-based video coding in JVET-AD2016. 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, perform complexity analysis, and develop complexity reductions of candidate technology. * Finalize and discuss the EE on neural network-based video coding. * Coordinate with other relevant groups, including SC29/AG5 on the evaluation and assessment of visual quality, and AHG12 on the interaction with ECM coding tools. If possible, prepare encodings with combinations of tools included in the NNVC software for visual quality assessment at the next meeting. * Coordinate with AHG14 on items related to NNVC software development. | E. Alshina, S. Liu, A. Segall (co‑chairs), F. Galpin, J. Li, R.-L. Liao, D. Rusanovskyy, T. Shao, M. Wien, P. Wu (vice‑chairs) | Y (tel., 2 weeks notice), first on Aug.2 |
| **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 ECM9 algorithm description JVET-AD2025. * 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-AD2017. * Analyse the results of exploration experiments described in JVET-AD2024 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, H. Yang (vice-chairs) | Y (tel., 2 weeks notice) |
| **Film grain technologies (AHG13)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the benefits and characteristics of film grain technologies, including autoregressive and frequency-filtering technologies. * Study alternative film grain models and their associated documentation. * In consultation with AHG4, study and define content characteristics and test conditions that are desirable for the study and testing of film grain technologies. * 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 finalizing the draft plan for subjective quality testing of the FGC SEI message JVET-AD2022, and conduct preparations for such testing. * Coordinate with AHG3 for software support of the FGC SEI message. | W. Husak, M. Radosavljević (co-chairs), A. Duenas, D. Grois, Y. He, P. de Lagrange, X. Meng, A. Segall, A. Tourapis, W. Zhang (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-5.0 software version and the reference configuration encodings according to the NNVC common test conditions as described in JVET-AD2016. * Investigate combinations of tools included in the NNVC software, prepare and release anchor data for all configurations of the software, including anchor for High Operation Point configuration. * 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-AD2019) editors to identify any mismatches between software and description document, suggest further updates to the description document as appropriate. * Develop the software basis for High Operation Point in NNVC-5.1 as described in JVET-AD0380. * Coordinate with AHG11 on items related to NNVC activities. | S. Eadie, F. Galpin, Y. Li, J. Shingala, L. Wang, Z. Xie (AHG chairs) | Y (tel., 2 weeks notice), first on Aug. 2 |

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 XXX) 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 212, as noted in section 9.

[JVET-AE1000](https://jvet-experts.org/doc_end_user/current_document.php?id=12968) Meeting Report of the 31st JVET Meeting [J.-R. Ohm] [WG 5 N XXX] (2023-08-XX)

Initial versions of the meeting notes (d0 … d9) 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)]

[JVET-AD1003](https://jvet-experts.org/doc_end_user/current_document.php?id=12970) Coding-independent code points for video signal type identification (Draft 2 of 3rd edition) [WG 5 preliminary FDIS N 206] [G. J. Sullivan, A. Tourapis] (2023-06-30)

Text for IPT-C2 from JVET-AC1008 is carried over, with swapping of the number assignment of IPT-C2 and YCgCo-Rx as suggested in the related ballot comment. A Draft DoCR N 205 was reviewed Thursday 27 April 1510. It was reported that SMPTE ST 2128 might be published before the July meeting, when it is planned to issue the FDIS.

[JVET-AD1004](https://jvet-experts.org/doc_end_user/current_document.php?id=12971) Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP [Y.-K. Wang, B. Bross, I. Moccagatta, C. Rosewarne, G. J. Sullivan] (2023-06-30, near next meeting)

New aspects included the issues pointed out in JVET-AD0078 (for AVC) Editors were asked to check if any items for VVC, VSEI, HEVC and Video CICP could be removed that have already been resolved in new editions. A new edition of H.264 might be produced at next meeting for ITU-T consent, and a new edition of ISO/IEC 14496-10 might also be requested and started with the DIS stage. If additional SEI messages would be carried over from initial development for other standards, it might be better to start with a CD consultation stage.

Remains valid – not updated: [JVET-Z1005](https://jvet-experts.org/doc_end_user/current_document.php?id=11707) New levels for HEVC (Draft 3) [T. Suzuki, A. Tourapis, Y.-K. Wang]

The content of this document will be included in a new edition of HEVC. For additional corrigenda items included, see documents JVET-AB1004, as well as notes under JVET-AC0311 and JVET-AC0346. A DoCR on ISO/IEC DIS23008-2 was submitted as WG 5 N178 (reviewed during session 24), and the FDIS was submitted as WG 5 N 179.

(JVET-Z1005 can be removed after publication of the new edition of ISO/IEC 23008-2.)

No output: JVET-Axx1006

Remains valid – not updated: [JCTVC-V1007](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=10312) SHVC Test Model 11 (SHM 11) Introduction and Encoder Description [G. Barroux, J. Boyce, J. Chen, M. M. Hannuksela, Y. Ye] [WG 11 N 15778]

[JVET-AD1008](https://jvet-experts.org/doc_end_user/current_document.php?id=12972) Additional colour type identifiers for AVC and HEVC (Draft 4) [G. J. Sullivan, W. Husak, A. Tourapis] [WG 5 Preliminary WD N 200] (2023-06-30)

Remove items related to CICP, and swap number assignment IPT-C2 and YCgCo-Rx.

Remains valid – not updated: [JCTVC-AC1009](https://jvet-experts.org/doc_end_user/current_document.php?id=12569) Common test conditions for SHVC [K. Sühring]

This requires an update, as the previous version referred to an outdated location of test sequences.

Remains valid – not updated [JCTVC-O1010](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=8511) Guidelines for Conformance Testing Bitstream Preparation [T. Suzuki, W. Wan]

Remains valid – not updated: [JVET-AA1011](https://jvet-experts.org/doc_end_user/current_document.php?id=11944) HEVC multiview profiles supporting extended bit depth (draft 1) [A. Tourapis, W. Husak] [WG 5 preliminary WD N 143]

Proponents of the new profiles reported that they were currently developing software and conformance bitstreams.

JVET-[AD1012](https://jvet-experts.org/doc_end_user/current_document.php?id=12973) Overview of IT systems used in JVET [J.-R. Ohm, I. Moccagatta, K. Sühring, M. Wien] (2023-05-19)

Adding an annex from JVET-AD0004.

Remains valid – not updated: [JCT3V-G1003](http://phenix.int-evry.fr/jct3v/doc_end_user/current_document.php?id=1884) 3D-AVC Test Model 9 [ D. Rusanovskyy, F. C. Chen, L. Zhang, T. Suzuki ] [WG 11 N 14239]

Remains valid – not updated: [JCT3V-K1003](http://phenix.int-evry.fr/jct3v/doc_end_user/current_document.php?id=2499) Test Model 11 of 3D-HEVC and MV-HEVC [Y. Chen, G. Tech, K. Wegner, S. Yea] [WG 11 N 15141]

Reserved number – [JVET-AE1013](https://jvet-experts.org/doc_end_user/current_document.php?id=12570) Common test conditions of 3DV experiments [K. Sühring]

New licensing available from JVET-AE0179.

Remains valid – not updated [JCTVC-V1014](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=10316) Screen Content Coding Test Model 7 Encoder Description (SCM 7) [R. Joshi, J. Xu, R. Cohen, S. Liu, Y. Ye] [WG 11 N 16049]

Remains valid – not updated: [JVET-AC1015](https://jvet-experts.org/doc_end_user/current_document.php?id=12571) Common test conditions for SCM-based screen content coding [K. Sühring]

This requires an update, as the previous version referred to an outdated location of test sequences.

No output: JVET-Axx1016 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**

[JVET-AD2002](https://jvet-experts.org/doc_end_user/current_document.php?id=12974) Algorithm description for Versatile Video Coding and Test Model 20 (VTM 20) [A. Browne, Y. Ye, S. Kim] [WG 5 N 204] (2023-07-07, near next meeting)

Updates from JVET-AD0169 (description of RPR downsampling), and JVET-AD0045.

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-AC2003](https://jvet-experts.org/doc_end_user/current_document.php?id=12573) Guidelines for VTM-based software development [F. Bossen, X. Li, K. Sühring]

Remains valid – not updated: [JVET-T2004](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10542) Algorithm descriptions of projection format conversion and video quality metrics in 360Lib (Version 12) [Y. Ye, J. Boyce]

[JVET-AD2005](https://jvet-experts.org/doc_end_user/current_document.php?id=12975) New level and systems-related supplemental enhancement information for VVC (Draft 5) [E. François, B. Bross, M. M. Hannuksela, A. Tourapis, Y.-K. Wang] (2023-06-30)

New version, changes to NNPFA according to JVET-AD0362, and editorial change from JVET-AD0077, and potentially additional tickets at editors’ discretion.

The content of this delta (amendment style) document was included in a new edition of VVC. A draft DoCR on ISO/IEC 23090-3/DAM1 was issued as WG 5 N 201 (reviewed Thursday 27 April 1530), and the preliminary FDIS text was issued as WG 5 N 202.

[JVET-AD2006](https://jvet-experts.org/doc_end_user/current_document.php?id=12976) Additional SEI messages for VSEI (Draft 4) [S. McCarthy, T. Chujoh, M. M. Hannuksela, G. J. Sullivan, Y.-K. Wang] (2023-06-30)

The content of this delta (amendment style) document was included in a new edition of VSEI. A draft DoCR on ISO/IEC 23007-7/DAM1 was issued as WG 5 N 197 (reviewed Thursday 27 April 1535), and the preliminary FDIS text was issued as WG 5 N 198.

See notes under JVET-AD0362 for elements included.

[JVET-AD2007](https://jvet-experts.org/doc_end_user/current_document.php?id=12977) Guidelines for NNVC software development [F. Galpin, S. Eadie, L. Wang, Z. Xie, Y. Li] (2023-05-26)

Developed from JVET-AD0111

Remains valid – not updated: [JVET-X2008](https://jvet-experts.org/doc_end_user/current_document.php?id=11228) Conformance testing for versatile video coding (Draft 7) [J. Boyce, F. Bossen, K. Kawamura, I. Moccagatta, W. Wan]

Remains valid – not updated: [JVET-Y2009](https://jvet-experts.org/doc_end_user/current_document.php?id=11470) Reference software for versatile video coding (Draft 3) [F. Bossen, K. Sühring, X. Li]

Remains valid – not updated [JVET-AB2010](https://jvet-experts.org/doc_end_user/current_document.php?id=12216) VTM and HM common test conditions and software reference configurations for SDR 4:2:0 10 bit video [F. Bossen, X. Li, V. Seregin, K. Sharman, K. Sühring]

Remains valid – not updated: [JVET-AC2011](https://jvet-experts.org/doc_end_user/current_document.php?id=12575) VTM and HM common test conditions and evaluation procedures for HDR/WCG video [A. Segall, E. François, W. Husak, S. Iwamura, D. Rusanovskyy]

Remains valid – not updated: [JVET-U2012](https://jvet-experts.org/doc_end_user/current_document.php?id=10681) JVET common test conditions and evaluation procedures for 360° video [Y. He, J. Boyce, K. Choi, J.-L. Lin]

Remains valid – not updated: [JVET-T2013](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10546) VTM common test conditions and software reference configurations for non-4:2:0 colour formats [Y.-H. Chao, Y.-C. Sun, J. Xu, X. Xu]

Remains valid – not updated: [JVET-Q2014](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9683) JVET common test conditions and software reference configurations for lossless, near lossless, and mixed lossy/lossless coding [T.-C. Ma, A. Nalci, T. Nguyen]

Remains valid – not updated: [JVET-Q2015](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9684) JVET functionality confirmation test conditions for reference picture resampling [J. Luo, V. Seregin]

[JVET-AD2016](https://jvet-experts.org/doc_end_user/current_document.php?id=12978) Common test conditions and evaluation procedures for neural network-based video coding technology [E. Alshina, R.-L. Liao, S. Liu, A. Segall] (2023-05-12)

This included the new definition of low-complexity operation point as anchor.

[JVET-AD2017](https://jvet-experts.org/doc_end_user/current_document.php?id=12979) Common test conditions and evaluation procedures for enhanced compression tool testing [M. Karczewicz and Y. Ye] (2023-05-13)

This included editorial improvements as suggested in JVET-AD0242, and an increase of runtime reporting precision as suggested in JVET-AD0401.

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-AD2019](https://jvet-experts.org/doc_end_user/current_document.php?id=12980) Description of algorithms and software in neural network-based video coding (NNVC) version 3 [F. Galpin, Y. Li, D. Rusanovskyy, J. Ström, L. Wang] [WG 5 N 208] (2023-06-16)

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-AC2020](https://jvet-experts.org/doc_end_user/current_document.php?id=12577) Film grain synthesis technology for video applications (Draft 4) [D. Grois, Y. He, W. Husak, P. de Lagrange, A. Norkin, M. Radosavljević, A. Tourapis, W. Wan] [WG 5 CDTR N 176] (2023-03-31)

[JVET-AD2021](https://jvet-experts.org/doc_end_user/current_document.php?id=12981) Verification test plan for VVC multilayer coding [S. Iwamura, P. de Lagrange, M. Wien] (2023-05-26)

See notes under section 4.5 for updates. See notes under JVET-AD0102 and JVET-AD0399.

[JVET-AD2022](https://jvet-experts.org/doc_end_user/current_document.php?id=12982) Draft plan for subjective quality testing of FGC SEI message [P. de Lagrange, W. Husak, M. Radosavljević, M. Wien] (2023-06-16)

See notes under sections 4.4 and 4.5 for updates. Update for phase 1 based on conclusions from expert viewing (AD0382). More test material needed. Update for phase 2 concepts based on considerations discussed under JVET-AD0268 and JVET-AD0369. For judging preservation of artistic intent, this might need involving (non-technical) expert viewers. To be discussed in interim AHG meeting.

[JVET-AD2023](https://jvet-experts.org/doc_end_user/current_document.php?id=12967) Exploration experiment on neural network-based video coding (EE1) [E. Alshina, F. Galpin, Y. Li, D. Rusanovskyy, M. Santamaria, J. Ström, L. Wang, Z. Xie] [WG 5 N 207] (2023-05-12)

An initial draft of this document was reviewed and approved at 0920-0945 on Friday 28 April.

Categories are:

* …

[JVET-AD2024](https://jvet-experts.org/doc_end_user/current_document.php?id=12966) Exploration experiment on enhanced compression beyond VVC capability (EE2) [V. Seregin, J. Chen, R. Chernyak, K. Naser, J. Ström, F. Wang, M. Winken, X. Xiu, K. Zhang] [WG 5 N 209] (2023-05-26)

An initial draft of this document was reviewed and approved at 0945-1000 on Friday 28 April.

Categories are partitioning, intra prediction, inter prediction, transforms, and in-loop filters.

[JVET-AD2025](https://jvet-experts.org/doc_end_user/current_document.php?id=12983) Algorithm description of Enhanced Compression Model 9 (ECM 9) [M. Coban, R.-L. Liao, K. Naser, J. Ström, L. Zhang] [WG 5 N 210] (2023-06-30)

New elements from notes elsewhere in this report:

* …

It is noted that the list above may not be complete; if some adoption is missing that is recorded somewhere else in the meeting notes it shall also be considered included.

Remains valid – not updated: [JVET-AC2026](https://jvet-experts.org/doc_end_user/current_document.php?id=12581) Conformance testing for VVC operation range extensions (Draft 4) [D. Rusanovskyy, T. Hashimoto, H.-J. Jhu, I. Moccagatta, Y. Yu] (2023-04-14)

[JVET-AD2027](https://jvet-experts.org/doc_end_user/current_document.php?id=12984) SEI processing order SEI message in VVC (draft 4) [S. McCarthy, M. M. Hannuksela, Y.-K. Wang] [WG 5 preliminary WD 3 N 203] (2023-05-19)

Updated from JVET-AD0386 and JVET-AD0362.

[JVET-AD2028](https://jvet-experts.org/doc_end_user/current_document.php?id=12985) Additional conformance bitstreams for VVC multilayer configurations [S. Iwamura, I. Moccagatta] (2023-05-26)

Include new streams from JVET-AD0101.

Remains valid – not updated: [JVET-AB2029](https://jvet-experts.org/doc_end_user/current_document.php?id=12225) Visual quality comparison of ECM/VTM encoding [V. Baroncini, J.-R. Ohm, M. Wien] [AG 5 N75]

[JVET-AD2030](https://jvet-experts.org/doc_end_user/current_document.php?id=12986) Optimization of encoders and receiving systems for machine analysis of coded video content (draft 2) [J. Chen, C. Hollmann, S. Liu] [WG 5 N 199)] (2023-06-09)

From JVET-AD0042, plus annex about example algorithms from JVET-AD0047 and JVET-AD0135.

[JVET-AD2031](https://jvet-experts.org/doc_end_user/current_document.php?id=12987) Common test conditions for optimization of encoders and receiving systems for machine analysis of coded video content [S. Liu, C. Hollmann] (2023-05-12)

This includes a new Excel template from JVET-AD0043, potentially new anchor settings from JVET-AD0122, and the new sequences from JVET-AD0181.

# Future meeting plans, expressions of thanks, and closing of the meeting

Future meeting plans were established according to the following guidelines (assuming face-to-face meetings):

* Meeting under ITU-T SG16 auspices when it meets (ordinarily starting meetings on the Wednesday of the first week and closing it on the Wednesday of the second week of the SG16 meeting – a total of 8 meeting days), and
* Otherwise meeting under ISO/IEC JTC 1/‌SC 29 auspices when its MPEG WGs meet (ordinarily starting meetings on the Friday prior to the main week of such meetings and closing it on the same day as other MPEG WGs – a total of 8 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 Fri. 13 – Fri. 20 October 2023, 32nd meeting under ISO/IEC JTC 1/‌SC 29 auspices in Hannover, DE,
* During Wed. 17 – Fri. 19 and Mon. 22 – Fri. 26 January 2024, 33rd meeting under ISO/IEC JTC 1/‌SC 29 auspices, to be held as teleconference meeting,
* During April 2024, 34th meeting under ITU-T SG16 auspices, date and location t.b.d.,
* During Fri. 12 – Fri. 19 July 2024, 35th meeting under ISO/IEC JTC 1/‌SC 29 auspices in Sapporo, JP,
* During October 2024, 36th meeting under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.,
* During January 2025, 37th meeting under ITU-T SG16 auspices, date and location t.b.d.,
* During April 2025, 38th meeting under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.,
* During Thu. 26 June – Fri. 4 July 2025, 39th meeting under ISO/IEC JTC 1/‌SC 29 auspices in Daejeon, KR.

The agreed document deadline for the 32nd JVET meeting was planned to be Friday 6 October 2023.

It was suggested that it would be interesting to perform subjective quality investigation with the elements adopted to NNVC software in a future meeting (provided that close enough rate/quality matching with VTM anchor is achieved).

Marius Preda was thanked for his support in maintaining the document site jvet-experts.org, as well as the document sites of JCT-VC and JCT-3V. Institut Mines-Télécom is thanked for hosting the sites.

(update) Ali Begen was thanked for the dedication put into organizing, and for the hospitality experienced during the 31st JVET meeting. Mustafa Bay, Alev Yavuz, Basak Erel and Merve Dağlı of Dekon Congress & Tourism, and the staff of Mirage Park Resort Hotel were thanked for the excellent support during the meeting, including the capability for remote access. Further thanks were expressed to the sponsors Adobe, Ateme, Comcast, Dolby, Ericsson, Nokia, özyegin University, ofinno, Perculus, Pico TV, Pixery, Turkish Airlines, Unified Streaming, and V-Nova.

The 31st JVET meeting was closed at approximately XXXX hours CEST (UTC+2) on Wednesday 19 July 2023.

# Annex A to JVET report: List of documents

(Dates and times in the table below are in Paris/Geneva time.)

# 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-first meeting of the JVET, according to an attendance sheet circulated during the JVET meeting sessions (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 thirtieth 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:

# Annex C to JVET report:

# Recommendations of the 12th meeting of ISO/IEC JTC 1/SC 29/WG 5 MPEG Joint Video Coding Team(s) with ITU-T SG 16

**ISO/IEC JTC 1/SC 29/WG 5 N XXX**