|  |  |
| --- | --- |
| **Joint Video Experts Team (JVET)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29**  32nd Meeting, Hannover, DE, 13–20 October 2023 | Document: JVET-AF\_notes\_d9 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Meeting Report of the 32nd Meeting of the Joint Video Experts Team (JVET), Hannover, DE, 13–20 October 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-second meeting during 13–20 October 2023 at the Hannover Congress Centrum (HCC) in Hannover, Germany. 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 thirteenth 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 0900 CEST (UTC+2) on Friday 13 October 2023. Meeting sessions were held on all days, including the weekend days of Saturday and Sunday 14 and 15 October 2023, until the meeting was closed at approximately XXXX hours CEST on Friday 20 October 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 document(s) were discussed. in coordination with a meeting of various SC29 Working Groups and Advisory Groups – where WG 5 is representing the Joint Video Coding Team(s) and their activities from the perspective of the SC 29 parent body, under whose auspices this JVET meeting was held. The subject matter of the JVET meeting activities consisted of work on further development and maintenance of the twin-text video coding technology standards *Advanced Video Coding* (AVC), *High Efficiency Video Coding* (HEVC), *Versatile Video Coding* (VVC)*, Coding-independent Code Points (Video)* (CICP), and *Versatile Supplemental Enhancement Information Messages for Coded Video Bitstreams* (VSEI), as well as related technical reports, reference software and conformance testing packages. Further important goals were reviewing the results of the Exploration Experiment (EE) on Neural Network-based Video Coding, of the EE on Enhanced Compression beyond VVC capability, of other technical input on novel aspects of video coding technology, and to plan next steps for investigation of candidate technology towards further standard development.

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

a) JVET documents

* [JVET-AE1004](https://jvet-experts.org/doc_end_user/current_document.php?id=12567) Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP
* [JVET-AE1006](https://jvet-experts.org/doc_end_user/current_document.php?id=12568) New profiles, colour decriptors, and SEI messages for HEVC (draft 1), also issued as WG 5 CDAM N 226
* [JVET-AE1011](https://jvet-experts.org/doc_end_user/current_document.php?id=12569) HEVC multiview profiles supporting extended bit depth (draft 2)
* [JVET-AE1013](https://jvet-experts.org/doc_end_user/current_document.php?id=12569) Common test conditions of 3DV experiments
* [JVET-AE1016](https://jvet-experts.org/doc_end_user/current_document.php?id=12568) AVC with extensions and corrections (draft 1), also issued as WG 5 CD of AVC 11th edition N 218
* [JVET-AE2005](https://jvet-experts.org/doc_end_user/current_document.php?id=12574) New level and systems-related supplemental enhancement information for VVC (Draft 6)
* [JVET-AE2006](https://jvet-experts.org/doc_end_user/current_document.php?id=12574) Additional SEI messages for VSEI (Draft 5)
* [JVET-AE2016](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-AE2017](https://jvet-experts.org/doc_end_user/current_document.php?id=12576) Common test conditions and evaluation procedures for enhanced compression tool testing
* [JVET-AE2019](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 4
* [JVET-AD2020](https://jvet-experts.org/doc_end_user/current_document.php?id=12579) Film grain synthesis technology for video applications (Draft 5), also issued as WG 5 DTR N 222
* [JVET-AE2021](https://jvet-experts.org/doc_end_user/current_document.php?id=12578) Verification test plan for VVC multilayer coding (update 1)
* [JVET-AE2023](https://jvet-experts.org/doc_end_user/current_document.php?id=12560) Exploration experiment on neural network-based video coding (EE1)
* [JVET-AE2024](https://jvet-experts.org/doc_end_user/current_document.php?id=12562) Exploration experiment on enhanced compression beyond VVC capability (EE2)
* [JVET-AE2025](https://jvet-experts.org/doc_end_user/current_document.php?id=12580) Algorithm description of Enhanced Compression Model 10 (ECM 10)
* [JVET-AE2027](https://jvet-experts.org/doc_end_user/current_document.php?id=12582) SEI processing order SEI message in VVC (Draft 5), also issued as WG 5 preliminary WD N 229
* [JVET-AE2028](https://jvet-experts.org/doc_end_user/current_document.php?id=12583) Additional conformance bitstreams for VVC multilayer configurations
* [JVET-AE2030](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 3), also issued as WG 5 preliminary WD N 224
* [JVET-AE2031](https://jvet-experts.org/doc_end_user/current_document.php?id=12584) Common test conditions for optimization of encoders and receiving systems for machine analysis of coded video content
* [JVET-AE2032](https://jvet-experts.org/doc_end_user/current_document.php?id=12584) Technologies under consideration for future extensions of VSEI (draft 1)

b) documents produced as WG 5 documents only:

* WG 5 N 217 Request for ISO/IEC 14496-10 11th edition
* WG 5 N 219 Disposition of comments received on ISO/IEC 23002-7:2022 DAM 1
* WG 5 N 220 Text of ISO/IEC FDIS 23002-7:202x Versatile supplemental enhancement infor­mation messages for coded video bitstreams (3rd edition)
* WG 5 N 222 Disposition of comments received on ISO/IEC CDTR 23002-9
* WG 5 N 225 Request for ISO/IEC 23008-2:202x/Amd.1
* WG 5 N 227 Disposition of comments received on ISO/IEC 23090-3:2022 DAM 1
* WG 5 N 228 Text of ISO/IEC FDIS 23090-3:202x Versatile video coding (3rd edition)

c) draft revised ITU-T Recommendations forwarded by JVET and Q6/16 for ITU-T Consent:

* TD169/Plen ITU-T H.265 (V9) "High efficiency video coding" (Rev.)
* TD171/Plen ITU-T H.266 (V3) "Versatile video coding" (Rev.)
* TD159/Plen ITU-T H.266.1 (V2) "Conformance specification for ITU-T H.266 versatile video coding" (Rev.)
* TD172/Plen ITU-T H.273 (V3) "Coding-independent code points for video signal type identification" (Rev.)
* TD170/Plen (A.5: TD173/Plen) ITU-T H.274 (V3) "Versatile supplemental enhancement information messages for coded video bitstreams" (Rev.)

As main results, the JVET produced 15 output documents from the current meeting:

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

The following 2 documents were produced as WG 5 documents only:

* WG 5 N 240 Disposition of comments received on ISO/IEC CD 14496-10:202x
* WG 5 N 243 Disposition of comments received on ISO/IEC 23008-2:202x (5th ed.) CDAM 1

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

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

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

# Administrative topics

## Organization

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

The Joint Video Experts Team (JVET) of ITU-T WP3/16 and ISO/IEC JTC 1/‌SC 29 held its thirty-second meeting during 13–20 October 2023 at the Hannover Congress Centrum (HCC) in Hannover, Germany. 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 thirteenth meeting as WG 5. The JVET meeting was held under the chairmanship of Dr Jens-Rainer Ohm (RWTH Aachen/Germany).

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

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

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

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

This report contains three important annexes, as follows:

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

## Meeting logistics

Information regarding logistics arrangements for the meeting had been provided via the email reflector [jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de) and at <http://wftp3.itu.int/av-arch/jvet-site/2023_10_AF_Hannover/>.

## Primary goals

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

a) JVET documents

* [JVET-AE1004](https://jvet-experts.org/doc_end_user/current_document.php?id=12567) Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP
* [JVET-AE1006](https://jvet-experts.org/doc_end_user/current_document.php?id=12568) New profiles, colour decriptors, and SEI messages for HEVC (draft 1), also issued as WG 5 CDAM N 226
* [JVET-AE1011](https://jvet-experts.org/doc_end_user/current_document.php?id=12569) HEVC multiview profiles supporting extended bit depth (draft 2)
* [JVET-AE1013](https://jvet-experts.org/doc_end_user/current_document.php?id=12569) Common test conditions of 3DV experiments
* [JVET-AE1016](https://jvet-experts.org/doc_end_user/current_document.php?id=12568) AVC with extensions and corrections (draft 1), also issued as WG 5 DIS of AVC 11th edition N 218
* [JVET-AE2005](https://jvet-experts.org/doc_end_user/current_document.php?id=12574) New level and systems-related supplemental enhancement information for VVC (Draft 6)
* [JVET-AE2006](https://jvet-experts.org/doc_end_user/current_document.php?id=12574) Additional SEI messages for VSEI (Draft 5)
* [JVET-AE2016](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-AE2017](https://jvet-experts.org/doc_end_user/current_document.php?id=12576) Common test conditions and evaluation procedures for enhanced compression tool testing
* [JVET-AE2019](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 4
* [JVET-AD2020](https://jvet-experts.org/doc_end_user/current_document.php?id=12579) Film grain synthesis technology for video applications (Draft 5), also issued as WG 5 DTR N 222 (DTR to be delivered)
* [JVET-AE2021](https://jvet-experts.org/doc_end_user/current_document.php?id=12578) Verification test plan for VVC multilayer coding (update 1)
* [JVET-AE2023](https://jvet-experts.org/doc_end_user/current_document.php?id=12560) Exploration experiment on neural network-based video coding (EE1)
* [JVET-AE2024](https://jvet-experts.org/doc_end_user/current_document.php?id=12562) Exploration experiment on enhanced compression beyond VVC capability (EE2)
* [JVET-AE2025](https://jvet-experts.org/doc_end_user/current_document.php?id=12580) Algorithm description of Enhanced Compression Model 10 (ECM 10)
* [JVET-AE2027](https://jvet-experts.org/doc_end_user/current_document.php?id=12582) SEI processing order SEI message in VVC (Draft 5), also issued as WG 5 preliminary WD N 229
* [JVET-AE2028](https://jvet-experts.org/doc_end_user/current_document.php?id=12583) Additional conformance bitstreams for VVC multilayer configurations
* [JVET-AE2030](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 3), also issued as WG 5 preliminary WD N 224
* [JVET-AE2031](https://jvet-experts.org/doc_end_user/current_document.php?id=12584) Common test conditions for optimization of encoders and receiving systems for machine analysis of coded video content
* [JVET-AE2032](https://jvet-experts.org/doc_end_user/current_document.php?id=12584) Technologies under consideration for future extensions of VSEI (draft 1)

b) documents produced as WG 5 documents only:

* WG 5 N 217 Request for ISO/IEC 14496-10 11th edition
* WG 5 N 219 Disposition of comments received on ISO/IEC 23002-7:2022 DAM 1
* WG 5 N 220 Text of ISO/IEC FDIS 23002-7:202x Versatile supplemental enhancement infor­mation messages for coded video bitstreams (3rd edition) (not delivered yet)
* WG 5 N 222 Disposition of comments received on ISO/IEC CDTR 23002-9
* WG 5 N 225 Request for ISO/IEC 23008-2:202x/Amd.1
* WG 5 N 227 Disposition of comments received on ISO/IEC 23090-3:2022 DAM 1
* WG 5 N 228 Text of ISO/IEC FDIS 23090-3:202x Versatile video coding (3rd edition) (not delivered yet)

c) draft revised ITU-T Recommendations forwarded by JVET and Q6/16 for ITU-T Consent:

* TD169/Plen ITU-T H.265 (V9) "High efficiency video coding" (Rev.)
* TD171/Plen ITU-T H.266 (V3) "Versatile video coding" (Rev.)
* TD159/Plen ITU-T H.266.1 (V2) "Conformance specification for ITU-T H.266 versatile video coding" (Rev.)
* TD172/Plen ITU-T H.273 (V3) "Coding-independent code points for video signal type identification" (Rev.)
* TD170/Plen (A.5: TD173/Plen) ITU-T H.274 (V3) "Versatile supplemental enhancement information messages for coded video bitstreams" (Rev.)

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

## Documents and document handling considerations

### General

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

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

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

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

* Decisions made by the group that might affect the normative content of a future standard are identified in this report by prefixing the description of the decision with the string “Decision:”.
* Decisions that affect one of the various software packages but have no normative effect on text are marked by the string “Decision (SW):”.
* Decisions that fix a “bug” in one of the test model descriptions such as VTM, HM, etc. (an error, oversight, or messiness) or in the associated software package are marked by the string “Decision (BF):”.
* Decisions that are merely editorial without effect on the technical content of a draft standard are marked by the string "Decision (Ed.):". Such editorial decisions are merely suggestions to the editor, who has the discretion to determine the final action taken if their judgment differs.
* Other parenthetical comments may be used for describing the impact or motivation of a decision. Some decisions are recorded with the word “agreed” rather than “Decision:”, especially for 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 Friday, 6 October 2023. Any documents uploaded after 1159 hours Paris/Geneva time on Saturday 7 October 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-AF0207 were registered after the “officially late” deadline (and therefore were also uploaded late). However, some documents in the “late” range might include break-out activity reports that were generated during the meeting, or documents which were requested to be produced for the purpose of improving specification text, and are therefore better considered as report documents rather than as late contributions.

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

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

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

* JVET-AF0XXX (a proposal on …), uploaded 10-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-AF0XXX (a document presenting …), uploaded 10-XX,
* …

All cross-verification reports at this meeting (except for JVET-AF0074, JVET-AF0124, and JVET-AF0125) 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-AF0046, JVET-AF0054, JVET-AF0077, JVET-AF0078, JVET-AF0123, JVET-AF0127, JVET-AF0195, JVET-AF0196, JVET-AF0246, JVET-AF0248, JVET-AF0300, JVET-AF0312.

The cross-verification report JVET-AF0239 was still missing by the end of the meeting on Friday 20 Oct., but became available later. The following cross-verification reports were still missing three weeks after the end of the meeting: JVET-AF0XXX, and … . These were marked as withdrawn by the JVET chair, assuming their registrations had become obsolete.

“Placeholder” contribution documents that were basically empty of content, or lacking any results showing benefit for the proposed technology, and obviously uploaded with an intent to provide a more complete submission as a revision, had been agreed to be considered unacceptable and to be rejected in the document management system until a more complete version was available (which would then be counted as a late contribution if the update was after the document deadline). At the current meeting, this situation applied with documents JVET-AF0XXX and … , which were categorized as late in the list above, based on the time of the first reasonable document upload; this case did not happen at the current meeting, but the sentence is kept for future use.

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

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

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

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

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

### Outputs of the preceding meeting

All output documents of the previous meeting, particularly the meeting report JVET-AE1000, the Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP JVET-AE1004, the New profiles, colour descriptors, and SEI messages for HEVC (draft 1) JVET-AE1006, the HEVC multiview profiles supporting extended bit depth (draft 2) JVET-AE1011, the Common test conditions of 3DV experiments [JVET-AE1013](https://jvet-experts.org/doc_end_user/current_document.php?id=12569), the AVC specification with extensions and corrections (draft 1) JVET-AE1016, the New level and systems-related supplemental enhancement information for VVC (Draft 6) JVET-AE2005, the Additional SEI messages for VSEI (Draft 5) JVET-AE2006, the Common test conditions and evaluation procedures for neural network-based video coding technology JVET-AE2016, the Common test conditions and evaluation procedures for enhanced compression tool testing JVET-AE2017, the Description of algorithms and software in neural network-based video coding (NNVC) version 4 JVET-AE2019, the Film grain synthesis technology for video applications (Draft 5) JVET-AE2020, the Verification test plan for VVC multilayer coding (update 1) JVET-AE2021, the Description of the EE on Neural Network-based Video Coding JVET-AE2023, the Description of the EE on Enhanced Compression beyond VVC capability JVET-AE2024, the Algorithm description of Enhanced Compression Model 10 (ECM 10) JVET-AE2025, the SEI processing order SEI message in VVC (Draft 5) JVET-AE2027, the Additional conformance bitstreams for VVC multilayer configurations JVET-AE2028, the Optimization of encoders and receiving systems for machine analysis of coded video content (Draft 3) [JVET-AE2030](https://jvet-experts.org/doc_end_user/current_document.php?id=12564), the Common test conditions for optimization of encoders and receiving systems for machine analysis of coded video content [JVET-AE2031](https://jvet-experts.org/doc_end_user/current_document.php?id=12584), and the Technologies under consideration for future extensions of VSEI (draft 1) JVET-AE2032, had been completed and were approved. In a few cases, the corresponding WG 5 N-numbered documents had not yet been uploaded, and this was requested to be done as soon as possible. The software implementations of VTM version 22.0, ECM version 10.0, and NNVC (versions 6.0 and 6.1) were also approved. The latter include modifications of loop filters confirmed by post-meeting AHG recommendations which were also approved by JVET.

Only minor editorial issues were found in the meeting report JVET-AE1000; no need to produce an update was identified (see section 2.14 for details).

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

## Attendance

The list of participants in the JVET meeting can be found in Annexes B1 and B2 of this report.

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

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

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

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

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

## Agenda

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

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

The agenda was approved as suggested.

The 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 of June/July 2009. Both ITU and ISO/IEC will be identified in the <OWNER> and <ORGANIZATION> tags in the header. This software header is also used in the process of designing the VTM and ECM software, and for evaluating proposals for technology to be potentially included in these designs. This software or parts thereof might be published by ITU-T and ISO/IEC as an example implementation of a future video coding standard and for use as the basis of products to promote adoption of such technology.

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

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

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

## Communication practices

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

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

It was emphasized that reflector subscriptions and email sent to the reflector must use real names when subscribing and sending messages and subscribers must respond to inquiries regarding the nature of their interest in the work. The current number of subscribers on the JVET email list was XXXX (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
* **HOP**: High-complexity operating point for neural network-based filter.
* **HRD**: Hypothetical reference decoder
* **HTM**: HEVC-based multiview and 3D test model (developed by JCT-3V)
* **HyGT**: Hyper-cube Givens transform (a type of NSST)
* **IBC** (also **Intra BC**): Intra block copy, also known as CPR – a technique by which sample values are predicted from other samples in the same picture by means of a displacement vector called a block vector, in a manner conceptually similar to motion-compensated prediction
* **IBDI**: Internal bit-depth increase – a technique by which lower bit-depth (8 bits per sample) source video is encoded using higher bit-depth signal processing, ordinarily including higher bit-depth reference picture storage (ordinarily 12 bits per sample)
* **IBF**: Intra boundary filtering
* **ILP**: Inter-layer prediction (in scalable coding)
* **ILRP**: Inter-layer reference picture
* **IPCM**: Intra pulse-code modulation (similar in spirit to IPCM in AVC and HEVC)
* **IRAP**: Intra random access picture
* **ISP**: Intra subblock partitioning
* **JCCR**: Joint coding of chroma residuals
* **JCT-3V**: Joint collaborative team on 3D video (for AVC and HEVC)
* **JCT-VC**: Joint collaborative team on video coding (for HEVC)
* **JEM**: Joint exploration model – a software codebase previously used for video coding exploration
* **JM**: Joint model – the primary software codebase that has been developed for the AVC standard
* **JSVM**: Joint scalable video model – another software codebase that has been developed for the AVC standard, which includes support for scalable video coding extensions
* **JVET**: Joint video experts team (initially for VVC, later expanded)
* **JVT**: Joint video team (for AVC)
* **KLT**: Karhunen-Loève transform
* **LB** or **LDB**: Low-delay B – the variant of the LD conditions that uses B pictures
* **LD**: Low delay – one of two sets of coding conditions designed to enable interactive real-time communication, with less emphasis on ease of random access (contrast with RA). Typically refers to LB, although also applies to LP
* **LFNST**: Low-frequency non-separable transform
* **LIC**: Local illumination compensation
* **LM**: Linear model
* **LMCS**: Luma mapping with chroma scaling (formerly sometimes called “in-loop reshaping”)
* **LOP**: Low-complexity operating point for neural network-based filter.
* **LP** or **LDP**: Low-delay P – the variant of the LD conditions that uses P frames
* **LUT**: Look-up table
* **LTRP**: Long-term reference picture
* **MANE**: Media-aware network element
* **MC**: Motion compensation
* **MCP**: Motion compensated prediction
* **MCTF**: Motion compensated temporal pre-filtering
* **MDNSST**: Mode dependent non-separable secondary transform
* **MIP**: Matrix-based intra prediction
* **MMLM**: Multi-model (cross component) linear mode
* **MMVD**: Merge with MVD
* **MPEG**: Moving picture experts group (an alliance of working groups and advisory groups in ISO/IEC JTC 1/‌SC 29, one of the two parent bodies of the JVET)
* **MPM**: Most probable mode (in intra prediction)
* **MRL**: Multiple reference line intra prediction
* **MV**: Motion vector
* **MVD**: Motion vector difference
* **NAL**: Network abstraction layer
* **NNVC**: Neural network-based video coding (experimental software package)
* **NSQT**: Non-square quadtree
* **NSST**: Non-separable secondary transform
* **NUH**: NAL unit header
* **NUT**: NAL unit type (as in AVC and HEVC)
* **OBMC**: Overlapped block motion compensation (e.g., as in H.263 Annex F)
* **OETF**: Opto-electronic transfer function – a function that converts to input light (e.g., light input to a camera) to a representation value
* **OLS**: Output layer set.
* **OOTF**: Optical-to-optical transfer function – a function that converts input light (e.g. l,ight input to a camera) to output light (e.g., light emitted by a display).
* **ONNX**: Open Neural Network Exchange – a format used to convert code from common neural network software packages into SADL code.
* **operation point**: A temporal subset of an OLS.
* **PDPC**: Position-dependent (intra) prediction combination.
* **PERP**: Padded equirectangular projection (a 360° projection format).
* **PH**: Picture header.
* **PHEC**: Padded hybrid equiangular cubemap (a 360° projection format).
* **PMMVD**: Pattern-matched motion vector derivation.
* **POC**: Picture order count.
* **PoR**: Plan of record.
* **PROF**: Prediction refinement with optical flow
* **PPS**: Picture parameter set (as in AVC and HEVC).
* **PTL**: Profile/tier/level combination.
* **QM**: Quantization matrix (as in AVC and HEVC).
* **QP**: Quantization parameter (as in AVC and HEVC, sometimes confused with quantization step size).
* **QT**: Quadtree.
* **RA**: Random access – a set of coding conditions designed to enable relatively-frequent random access points in the coded video data, with less emphasis on minimization of delay (contrast with LD).
* **RADL**: Random-access decodable leading (type of picture).
* **RASL**: Random-access skipped leading (type of picture).
* **R-D**: Rate-distortion.
* **RDO**: Rate-distortion optimization.
* **RDOQ**: Rate-distortion optimized quantization.
* **RDPCM**: Residual DPCM
* **ROT**: Rotation operation for low-frequency transform coefficients.
* **RPL**: Reference picture list.
* **RPLM**: Reference picture list modification.
* **RPR**: Reference picture resampling (e.g., as in H.263 Annex P), a special case of which is also known as ARC or DRC.
* **RPS**: Reference picture set.
* **RQT**: Residual quadtree.
* **RRU**: Reduced-resolution update (e.g. as in H.263 Annex Q).
* **RVM**: Rate variation measure.
* **SADL**: Small adhoc deep learning library
* **SAO**: Sample-adaptive offset.
* **SBT**: Subblock transform.
* **SbTMVP**: Subblock based temporal motion vector prediction.
* **SCIPU**: Smallest chroma intra prediction unit.
* **SD**: Slice data; alternatively, standard-definition.
* **SDH**: Sign data hiding.
* **SDT**: Signal-dependent transform.
* **SE**: Syntax element.
* **SEI**: Supplemental enhancement information (as in AVC and HEVC).
* **SH**: Slice header.
* **SHM**: Scalable HM.
* **SHVC**: Scalable high efficiency video coding.
* **SIF**: Switchable (motion) interpolation filter.
* **SIMD**: Single instruction, multiple data.
* **SMVD**: Symmetric MVD.
* **SPS**: Sequence parameter set (as in AVC and HEVC).
* **STMVP**: Spatial-temporal motion vector prediction.
* **STRP**: Short-term reference picture.
* **STSA**: Step-wise temporal sublayer access.
* **TBA/TBD/TBP**: To be announced/determined/presented.
* **TGM**: Text and graphics with motion – a category of content that primarily contains rendered text and graphics with motion, mixed with a relatively small amount of camera-captured content.
* **TIMD**: Template-based intra mode derivation
* **TM**: Template matching.
* **TMVP**: Temporal motion vector prediction.
* **TS**: Transform skip.
* **TSRC**: Transform skip residual coding.
* **TT**: Ternary tree.
* **UCBDS**: Unrestricted center-biased diamond search.
* **UGC**: User-generated content.
* **UWP**: Unequal weight prediction.
* **VCEG**: Visual coding experts group (ITU-T Q.6/16, the relevant rapporteur group in ITU-T WP3/16, which is one of the two parent bodies of the JVET).
* **VCM**: Video coding for machines.
* **VPS**: Video parameter set – a parameter set that describes the overall characteristics of a coded video sequence – conceptually sitting above the SPS in the syntax hierarchy.
* **VQA**: Visual quality assessment.
* **VT**: Verification testing.
* **VTM**: VVC Test Model.
* **VUI**: Video usability information.
* **VVC**: Versatile Video Coding, the standardization project developed by JVET.
* **WAIP**: Wide-angle intra prediction
* **WCG**: Wide colour gamut.
* **WG**: Working group, a group of technical experts (usually used to refer to WG 11, a.k.a. MPEG).
* **WPP**: Wavefront parallel processing (usually synonymous with ECS).
* Block and unit names in HEVC:
  + **CTB**: Coding tree block (luma or chroma) – unless the format is monochrome, there are three CTBs per CTU.
  + **CTU**: Coding tree unit (containing both luma and chroma, synonymous with LCU), with a size of 16x16, 32x32, or 64x64 for the luma component.
  + **CB**: Coding block (luma or chroma), a luma or chroma block in a CU.
  + **CU**: Coding unit (containing both luma and chroma), the level at which the prediction mode, such as intra versus inter, is determined in HEVC, with a size of 2Nx2N for 2N equal to 8, 16, 32, or 64 for luma.
  + **PB**: Prediction block (luma or chroma), a luma or chroma block of a PU, the level at which the prediction information is conveyed or the level at which the prediction process is performed in HEVC.
  + **PU**: Prediction unit (containing both luma and chroma), the level of the prediction control syntax within a CU, with eight shape possibilities in HEVC:
    - **2Nx2N**: Having the full width and height of the CU.
    - **2NxN (or Nx2N)**: Having two areas that each have the full width and half the height of the CU (or having two areas that each have half the width and the full height of the CU).
    - **NxN**: Having four areas that each have half the width and half the height of the CU, with N equal to 4, 8, 16, or 32 for intra-predicted luma and N equal to 8, 16, or 32 for inter-predicted luma – a case only used when 2N×2N is the minimum CU size.
    - **N/2x2N** paired with **3N/2x2N** or **2NxN/2** paired with **2Nx3N/2**: Having two areas that are different in size – cases referred to as AMP, with 2N equal to 16 or 32 for the luma component.
  + **TB**: Transform block (luma or chroma), a luma or chroma block of a TU, with a size of 4x4, 8x8, 16x16, or 32x32.
  + **TU**: Transform unit (containing both luma and chroma), the level of the residual transform (or transform skip or palette coding) segmentation within a CU (which, when using inter prediction in HEVC, may sometimes span across multiple PU regions).
* Block and unit names in VVC:
  + **CTB**: Coding tree block (luma or chroma) – there are three CTBs per CTU in a P or B slice or in an I slice that uses a single tree, and one CTB per luma CTU and two CTBs per chroma CTU in an I slice that uses separate trees.
  + **CTU**: Coding tree unit (synonymous with LCU, containing both luma and chroma in a P or B slice or in an I slice that uses a single tree, containing only luma or only chroma in an I slice that uses separate trees), with a size of 16x16, 32x32, 64x64, or 128x128 for the luma component.
  + **CB**: Coding block, a luma or chroma block in a CU.
  + **CU**: Coding unit (containing both luma and chroma in P/B slice, containing only luma or chroma in I slice), a leaf node of a QTBT. It’s the level at which the prediction process and residual transform are performed in JEM. A CU can be square or rectangle shape.
  + **PB**: Prediction block, a luma or chroma block of a PU.
  + **PU**: Prediction unit, has the same size as a CU in the VVC context.
  + **TB**: Transform block, a luma or chroma block of a TU.
  + **TU**: Transform unit, has the same size as a CU in the VVC context.

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

* MPEG-2 | H.262 (coding specification is common text)
  + ITU-T H.262 V3 was approved in 2012-02; Amd.1 was approved in 2013-03 and was not published separately; it was instead incorporated directly into the V3 text and published 2013-09
  + ISO/IEC 13818-2:2013 V3 FDIS ballot closed 2012-05-08; FDAM 1 ballot closed 2013-04-12 and was not published separately; it was instead incorporated directly into the V3 text and published 2013-10
  + Conformance testing (not joint with ITU-T)
    - ISO/IEC 13818-4:2004 V2 FDIS closed 2004-08-22, published 2004-12-12; it specifies conformance testing for Part 1 (Systems), Part 2 (Video), Part 3 (Audio), and Part 7 (AAC)
    - ISO/IEC 13818-4:2004/Amd 3:2009 Level for 1080@50p/60p conformance testing
    - Cor 1:2007, Cor 2:2009, Cor 3:2012, Cor 4:2011 may also have video relevance
  + Reference software (not joint with ITU-T)
    - ISO/IEC TR 13818-5:2005 V2 FDIS closed 2005-07-24, published 2005-10; it specifies reference software for Part 1 (Systems), Part 2 (Video), Part 3 (Audio), Part 7 (AAC) and Part 11 (IPMP)
* AVC (twin text)
  + ITU-T H.264 V14 was Consented at 22nd meeting on 2021-04-30 (with annotated regions, shutter interval, and miscellaneous corrections), approved 2021-08-22, published 2021-10-13
  + ISO/IEC 14496-10:2020 (Ed. 9) FDIS ballot closed 2020-11-27, published 2020-12-15
  + ISO/IEC 14496-10:2022 (Ed. 10), had been forwarded from DIS directly for publication 2022-01-21 (with annotated regions, shutter interval, and miscellaneous corrections) with an editing period, submitted to ITTF in 2022-05 after consultation with ISO staff on format of graphics files, upgraded to “DIS approved for registration” in ISO Project system 2022-07-04, published 2022-11-07
  + Preliminary draft text for YCgCo-Re and YCgCo-Ro issued at 26th meeting, second draft including SMPTE ST 2128 issued at 28th meeting 2022-10, third draft issued at 29th meeting 2023-01, fourth draft issued at 30th meeting 2023-04, formal project requested and CD of 11th edition issued at 31st meeting 2023-07, ready for progression to DIS at the current meeting.
  + Conformance testing (twin text)
    - ITU-T H.264.1 V6 Approved 2016-02-13, published 2016-06-17
    - Various amendments of ISO/IEC 14496-4:2004, including:
      * ISO/IEC 14496-4:2004/AMD 6:2005 Advanced Video Coding conformance
      * ISO/IEC 14496-4:2004/AMD 9:2006 AVC fidelity range extensions conformance
      * ISO/IEC 14496-4:2004/AMD 30:2009 Conformance testing for new profiles for professional applications
      * ISO/IEC 14496-4:2004/AMD 31:2009 Conformance testing for SVC profiles
      * ISO/IEC 14496-4:2004/AMD 38:2010 Conformance testing for Multiview Video Coding
      * ISO/IEC 14496-4:2004/AMD 41:2014 Conformance testing of MVC plus depth extension of AVC
      * ISO/IEC 14496-4:2004/AMD 42:2014 Conformance testing of Multi-Resolution Frame Compatible Stereo Coding extension of AVC
      * ISO/IEC 14496-4:2004/AMD 43:2015 3D-AVC conformance testing
      * ISO/IEC 14496-4:2004/AMD 45:2016 Conformance Testing for the Multi-resolution Frame Compatible Stereo Coding with Depth Maps Extension of AVC
  + Reference software (twin text)
    - ITU-T H.264.2 V7 Approved 2016-02-13, published 2016-05-30
    - Various amendments of ISO/IEC 14496-5:2001 have been published, including:
      * ISO/IEC 14496-5:2001/AMD 6:2005 Advanced Video Coding (AVC) and High Efficiency Advanced Audio Coding (HE AAC) reference software
      * ISO/IEC 14496-5:2001/AMD 8:2006 AVC fidelity range extensions reference software
      * ISO/IEC 14496-5:2001/AMD 15:2010 Reference software for Multiview Video Coding
      * ISO/IEC 14496-5:2001/AMD 18:2008 Reference software for new profiles for professional applications
      * ISO/IEC 14496-5:2001/AMD 19:2009 Reference software for Scalable Video Coding
      * ISO/IEC 14496-5:2001/AMD 33:2015 Reference software for MVC plus depth extension of AVC
      * ISO/IEC 14496-5:2001/AMD 34:2014 Reference software of the multi-resolution frame compatible stereo coding of AVC
      * ISO/IEC 14496-5:2001/AMD 35:2015 3D-AVC Reference software
      * ISO/IEC 14496-5:2001/AMD 39:2016 Reference software for the Multi-resolution Frame Compatible Stereo Coding with Depth Maps of AVC
      * ISO/IEC 14496-5:2001/AMD 42:2017 Reference software for the alternative depth information SEI message extension of AVC
* HEVC (twin text)
  + ITU-T H.265 V7 approved 2019-11-29, published 2020-01-10
  + ISO/IEC 23008-2:2020 (Ed. 4) FDIS ballot closed 2020-07-16, published 2020-08-27
  + ITU-T H.265 V8 Consented at the 22nd meeting (shutter interval information SEI message and miscellaneous corrections), published 2020-10-13
  + ISO/IEC 23008-2:2020/AMD 1:2021 (shutter interval information SEI message) published 2021-07-12
  + ISO/IEC 23008-2:202x (Ed. 5) began as CDAM 2 High-range levels output of 25th meeting of January 2022, CDAM ballot closed 2022-04-15, conversion to 5th edition with miscellaneous corrections planned at 26th meeting of 2022-04, text submitted for DIS ballot 2022-07-10, DIS ballot closed 2023-01-10, FDIS issued 29th meeting of 2023-01, FDIS ballot opened 2023-08-06, closed 2023-10-02, pending publication
  + ITU-T H.265 V9 Consented at 31st meeting 2023-07, approved and pre-published 2023-09.
  + 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, formal work item requested and CDAM1 issued 31st meeting 2023-07.
  + Conformance testing (twin text)
    - ITU-T H.265.1 V3 approved 2018-10-14, published 2019-01-15
    - ISO/IEC 23008-8:2018 (Ed. 2) Conformance specification for HEVC, published 2018-08-06
    - ISO/IEC 23008-8:2018/AMD 1:2019 Conformance testing for HEVC screen content coding (SCC) extensions and non-intra high throughput profiles, published 2019-10-15
  + Reference software (twin text)
    - ITU-T H.265.2 V4 approved 2016-12-22, published 2017-04-10
    - ISO/IEC 23008-5:2017 (Ed. 2) Reference software for high efficiency video coding, published 2017-03-01
    - ISO/IEC 23008-5:2017/AMD 1:2017 Reference software for screen content coding extensions, published 2017-11-09
* VVC (twin text)
  + ITU-T H.266 V1 approved 2020-08-29, published 2020-11-10
  + ISO/IEC 23090-3:2021 (Ed. 1) published 2021-02-16
  + ITU-T H.266 V2 with operation range extensions, Consented 2022-01-28, Last Call began 2022-04-01, Approved 2022-04-29, pre-published 2022-06-06, published 2022-07-12
  + ISO/IEC 23090-3:2022 (Ed. 2) with operation range extensions, approval at WG level to proceed to FDIS 2022-01-21, published 2022-09-25
  + ISO/IEC 23090-3:202x (Ed. 2) / Amd.1 New level and systems-related supplemental enhancement information, CDAM 1 issued from 26th meeting, ballot closed 2022-07-14, DAM 1 issued from 27th meeting, ballot closed 2023-01-03, FDIS issued 2023-07
  + ITU-T H.266 V3 Consented 2023-07, approved and pre-published 2023-09
  + 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 3rd edition issued as an output of the 29th meeting in January 2023 (awaiting ballot at the time of this meeting)
    - ITU-T H.266.1 V4 Consented 2023-07, approved and pre-published 2023-09.
  + 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,[m62571](https://dms.mpeg.expert/doc_end_user/current_document.php?id=86620&id_meeting=194) FDIS 3rd edition issued 2023-07 (pending ballot)
  + ITU-T H.274 V3 Consent 2023-07, approved 2023-07, pending publication.
* CICP (twin text)
  + ISO/IEC 23091-2:2021 (Ed. 2) had been forwarded from DIS directly for publication in 2021-04 and published 2021-10-18
  + ITU-T H.273 V2 (with 4:2:0 sampling alignment and corrections for range of values for sample aspect ratio, ICTCP equations for HLG, and transfer characteristics function for sYCC of IEC 61966-2-1) Consented on 2021-04-30, Last Call closed during the 23rd meeting with approval on 2021-07-14, published 2021-09-24
  + ISO/IEC 23091-2:202x (Ed. 3) Request for new edition and CD for new edition (including YCgCo-Re and YCoCg-Ro) issued at 27th meeting, ballot closed 2022-10-10, DIS registered 2022-11-13, DIS ballot closed 2023-04-06, preliminary draft text for including SMPTE ST 2128 issued at 28th meeting, incorporated into preliminary FDIS at 30th meeting 2023-04, FDIS issued 2023-07
  + ITU-T H.273 Consent 2023-07, approved 2023-09, publication waiting for publication of SMPTE ST 2128.
* Conversion and coding practices for HDR/WCG Y′CbCr 4:2:0 video with PQ transfer characteristics (twin text)
  + H.Sup15 V1, approved 2017-01-27, published 2017-04-12
  + ISO/IEC TR 23008-14:2018 published 2018-08
* Signalling, backward compatibility and display adaptation for HDR/WCG video coding (twin text)
  + H.Sup18 V1, approved 2017-10-27, published 2018-01-18
  + ISO/IEC TR 23008-15:2018 published 2018-08
* Usage of video signal type code points (twin text)
  + H.Sup19 V3 approved 2021-04-30, published 2021-06-04
  + ISO/IEC TR 23091-4 (Ed. 3) published 2021-05-23
* Working practices using objective metrics for evaluation of video coding efficiency experiments (twin text)
  + HSTP-VID-WPOM V1: approved 2020-07-03, published 2020-11
  + ISO/IEC TR 23002-8 (Ed. 1) published 2021-05-20
* Film grain synthesis technologies for video applications (twin text)
  + ISO/IEC TR 23002-9 Request for subdivision and WD 1 issued at 25th meeting 2022-01-21, WD 2 issued at 27th meeting, WD 3 issued at 28th meeting, CDTR issued at 29th meeting 2023-07, 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 2023-10-09:
  + ISO/IEC 13818-4:2004 Conformance for MPEG-2
  + ISO/IEC 13818-4:2004/Amd 3:2009 Level for 1080@50p/60p conformance testing
  + ISO/IEC TR 13818-5:2005 Software simulation for MPEG-2
  + Various amendments of ISO/IEC 14496-4:2004 Conformance for AVC
  + Various amendments of ISO/IEC 14496-5:2001 Reference software for AVC
  + ISO/IEC 14496-10:2022 (Ed. 10) AVC
  + ISO/IEC 23002-7:2022 (Ed. 2) – VSEI
  + ISO/IEC 23008-2:2020 (Ed. 4) HEVC
  + ISO/IEC 23090-3:2022 (Ed. 2) VVC
  + ISO/IEC 23090-15:2022 (Ed. 1) Conformance for VVC
  + ISO/IEC 23090-16:2022 (Ed. 1) Reference software for VVC
  + ISO/IEC 23091-2:2021 (Ed. 2) Video CICP
* The following standards that have been intended by JVET to be publicly available were not available at <https://standards.iso.org/ittf/PubliclyAvailableStandards/index.html> as of 2023-05-26 and this was still the case as of 2023-10-06. (These should be checked for previously issued requests for free availability.)
  + ISO/IEC 23008-2:2020 (Ed. 4) Amd.1:2021: Shutter interval information SEI message, published 2021-07-12 (has not been requested)
  + ISO/IEC 23008-5:2017 (Ed. 2) Reference software for high efficiency video coding, published 2017-03-01
  + ISO/IEC 23008-5:2017/AMD 1:2017 Reference software for screen content coding extensions, published 2017-11-09
  + ISO/IEC 23008-8:2018 (Ed. 2) Conformance specification for HEVC, published 2018-08, published 2018-08-06
  + ISO/IEC 23008-8:2018/AMD 1:2019 Conformance testing for HEVC screen content coding (SCC) extensions and non-intra high throughput profiles, published 2019-10-15
* It appears necessary to check if all older software and conformance packages are publicly available – it might be that it was never requested, e.g. for those that were produced by JCT-3V. This topic was left TBD until the next meeting – perhaps it would be best to compile a list of all relevant software and conformance parts of AVC, HEVC, MPEG-2 aka H.262, CICP, and request these in bulk.

## Draft standards progression status

* AVC colour type indicators for YCgCo-Re, YCgCo-Ro, and SMPTE ST 2128 (IPT-PQ-C2) and XXX are in a CD of the 11th edition issued in July 2023 (WG 5 N 218).
* HEVC 23008-2:202x (5th ed.) CDAM1 New profiles, colour descriptors, and SEI messages, with colour type indicators for YCgCo-Re, YCgCo-Ro, and SMPTE ST 2128 (IPT-PQ-C2) and XXX was issued in July 2023 (WG 5 N 226)
* HEVC new levels (from JVET-Z1005) – ISO/IEC 23008-2 DIS of new edition of HEVC was issued from the April 2022 26th meeting, incorporating Amd.1 and corrigenda items (ballot closed 2023-01-10, ballot comments in the Summary of Voting document [m61834](https://dms.mpeg.expert/doc_end_user/current_document.php?id=85619&id_meeting=193)); note that Amd.1 = shutter interval SEI is already included in latest ITU-T edition of H.265. It is noted that there are potential additional items (corrigenda+tickets, YCgCo-Re and YCgCo-Ro draft, SMPTE ST 2128, multiview profiles draft) where only corrigenda items were included in the FDIS text based on ballot comments, ballot had not been started yet. ITU-T consent for a new edition is planned for July 2023. It was noted that the referencing of VSEI is also somewhat different in the ITU-T and ISO/IEC versions of HEVC and/or AVC, which might be aligned at the next convenient time (basically editorial – e.g., the ITU version of AVC specifies the annotated regions SEI message without referencing VSEI, whereas the ISO/IEC version references VSEI for the syntax and semantics of that SEI message). However, there is currently no other need for HEVC to reference the VSEI standard. An FDIS for HEVC was issued as an output of the 29th meeting in January 2023 (and it does not reference VSEI). Its ballot began 2023-08-06 and closed 2023-10-02, and it was pending publication. A new edition of H.265 (v9) was Consented in July 2023, and approved and pre-published in September 2023 (not referencing VSEI).
* VVC new level and systems-related supplemental enhancement information (from JVET-AA2005) – VVC DAM was issued from 27th meeting, ballot closed 2023-01-03, ballot comments in the Summary of Voting document [m61833](https://dms.mpeg.expert/doc_end_user/current_document.php?id=85618&id_meeting=193). This was converted into a preliminary FDIS of VVC 3rd edition ([WG 5 N 183](https://dms.mpeg.expert/doc_end_user/current_document.php?id=86365&id_meeting=193)) at the 29th meeting of January 2023, anticipating that some alignment would be necessary with the ongoing VSEI amendment. Another preliminary FDIS was issued (WG 5 N 202) from the April meeting. The FDIS was then issued (WG 5 N 228) from the 31st meeting in July 2023. A new edition of H.266 was Consented in July 2023, approved in September 2023 and currently pending publication.
* VVC Conformance testing for operation range extensions – (from JVET-Y2026) – the DAM ballot closed 2022-11-15 (ballot comments in the Summary of Voting document [m61832](https://dms.mpeg.expert/doc_end_user/current_document.php?id=85617&id_meeting=193)), and this was consolidated into an FDIS at the 29th meeting, but the ballot had not been started yet. ITU-T H.266.1 was Consented in July 2023 and approved and pre-published in September 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). The FDIS of a new edition of ISO/IEC 23002-7 was issued (WG 5 N 220) from the 31st meeting in July 2023 and also reached ITU-T Consent at that meeting. H.274 v3 was approved in September 2023 and currently pending publication.
* 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. A DTR ballot was issued from the 31st meeting in July 2023 and results are expected prior to the 33rd meeting in January 2024. (It was noted that a second DTR could become necessary in case of comments). The publication limit date was reportedly 2023-08-09, so action to extend that date may be needed. ITU-T approval would be anticipated in April 2024.
* Video CICP new edition draft for YCgCo-Re and YCgCo-Ro (from JVET-Z1003), an ISO/IEC 23091-2 preliminary FDIS was issued from the 30th meeting and the Summary of Voting document was available as [m62572](https://dms.mpeg.expert/doc_end_user/current_document.php?id=86621&id_meeting=194) and a draft DoC had been issued as WG 5 [N 205](https://sd.iso.org/documents/ui/#!/browse/iso/iso-iec-jtc-1/iso-iec-jtc-1-sc-29/iso-iec-jtc-1-sc-29-wg-5/library/2/Draft%20disposition%20of%20comments%20received%20on%20ISO-IEC%20DIS%2023091-2%3A202X). There was a delay in the submittal of the FDIS due to dependency on the status of SMPTE ST 2128, which was tentatively included in the preliminary FDIS, based on an NB comment. The video CICP colour type indicator for SMPTE ST 2128 had been drafted and incorporated into the preliminary FDIS issued at the 30th meeting of April 2023. It had been reported that the specification was expected to become finalized in the SMPTE meeting in March 2023, but this had not yet happened, so the production of the FDIS was delayed. ITU-T Consent for H.273 v3 proceeded at the 31st meeting of July 2023 (to prevent undue delay since SG16 does not meet very frequently) and the text was approved in September 2023, but the text was on hold pending the publication of SMPTE ST 2128.
* A request for free availability in ISO/IEC has to be made for each edition, amendment and corrigendum, and the request needs to be approved in the WG 5 Recommendations. A request form also needs to be filled out (but the form does not need to be issued as a WG 5 document). A freely available URL for the ITU publication should be provided for the following parts:
  + For the ongoing work items, when they become finalized
  + ISO/IEC 23008-2:2020/Amd.1:2021 – HEVC FDAM issued 20th meeting (October 2020), public availability not yet requested but may not be necessary as it becomes included in next edition

## Opening remarks

Remarks during the opening session of the meeting Friday 13 October at 0900 CEST were as follows.

* Timing and organization of the meeting and online access, calendar posting of session plans
  + The initial number of documents was increased relative to the previous meeting (160->180) – parallel sessions might be necessary.
  + Scheduling of NNVC discussions – might need alignment with NNPF discussions (but no overlap with JPEG meeting this time, JPEG meets from Oct. 30)
* Plans for subsequent F2F meeting in April (Rennes), July (Sapporo), and October/November 2024 (tbd).
* 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). Presentations in the room are also managed via zoom – experts who present need to connect to the zoom meeting for screen sharing. Recording of the meeting notes by the session chair will also permanently be shared via zoom.
  + Having text and software available is crucial (and not just arriving at the end of the meeting).
  + There were no objections voiced in the opening plenary to the consideration of late contributions.
* The results of the previous meeting and the meeting report JVET-AE1000 were reviewed. The following small issues in the meeting report were noted and were not considered sufficient to warrant issuing a revision. These are obviously left over from a previous report, and the correct information can be found in other places of the report:
  + In section 1 (summary), JVET-AE1016 is alternatively denoted as “WG 5 DIS of AVC 11th edition N 218”, which actually was issued as WG 5 CD, not DIS.
  + In section 2.1, annex B is mentioned, but information is missing that it consists of two parts, (B1) in-person attendees as recorded by ITU-T via registration and badge pickup, (B2) remote attendees as recorded by the teleconferencing tool used for the meeting.
  + In annex B2, the remote attendees are associated with the thirtieth JVET meeting, which should have been the thirty-first meeting.
* There was a somewhat decreased number of late non-cross-check documents, compared to previous meetings. In particular, all non-cross-check documents that had been registered before the deadline were also available in time.
* 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 were asked not to pick a specific JVET number for regular documents – this function is reserved for AHG reports, summary reports, and output docs
* Experts were asked to always register JVET documents via the “jvet-experts.org” site, not via the MPEG dms site.
* Experts were asked to inform the chair when the title of a document is changed, or if authors are added. Otherwise, that might not be correct in the meeting notes.
* The primary goals of the meeting were:
  + 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.
  + New edition video CICP FDIS (DIS ballot response in [m62572](https://dms.mpeg.expert/doc_end_user/current_document.php?id=86621&id_meeting=194) draft DoCR in [MDS22710](https://dms.mpeg.expert/doc_end_user/current_document.php?id=87856&id_meeting=194), and preliminary FDIS text in [MDS22711](https://dms.mpeg.expert/doc_end_user/current_document.php?id=87857&id_meeting=194)) – inclusion of ST 2128 descriptor needed clarification of the status, therefore it was not yet submitted for FDIS yet; was consented and approved in ITU-T, but cannot be published as long as ST2128 is not available. It was verbally reported that availability can be expected around 6 weeks after the meeting – therefore it was concluded to issue the FDIS in January.
  + TR on film grain synthesis technology for video applications – DTR not yet submitted for ballot in ISO, ballot results cannot be expected before the January meeting. ITU consent is targeted for April 2024, possibly second DTR by April as well.
  + Optimization of encoders and receiving systems for machine analysis of coded video content – new WD of TR to be issued
  + Preparation of subjective tests for film grain, also new content proposed – expert viewing to be planned
  + Also plan expert viewing for multi-layer verification test
  + Any action items on reference software JM/HM? Status of MV-HEVC software and test conditions (refer to resolution of a previous meeting for the latter)?
  + New edition of AVC
    - 11th ed. 14496-10 CD ballot response in [m65567](https://dms.mpeg.expert/doc_end_user/current_document.php?id=90188&id_meeting=196), DIS to be issued (no long editing period if ballot results shall return before April 2024 meeting)
    - ITU consent for new H.264 edition was targeted for April 2024
  + HEVC updates
    - CDAM includes elements that are not in the ISO 5th edition – ballot response in [m65568](https://dms.mpeg.expert/doc_end_user/current_document.php?id=90189&id_meeting=196), DAM to be issued, check/confirm registered title, possibly recommendation (no long editing period if ballot results shall return before April 2024 meeting).
    - 5th edition not published yet, ballot results in [m65579](https://dms.mpeg.expert/doc_end_user/current_document.php?id=90200&id_meeting=196). Would be important to get the IS published still in 2023, to allow for a next edition including the ongoing Amd.1 in 2024.
    - An updated H.265 edition was sent for ITU consent in the last meeting; consent for next H.265 edition was targeted for April 2024 with new elements from CDAM
  + Exploration Experiments
    - Neural network-based video coding
    - Enhanced compression beyond VVC
* Liaison communication:
  + Incoming liaison statements: JPEG [m65594](https://dms.mpeg.expert/doc_end_user/current_document.php?id=90215&id_meeting=196), 3GPP [m65645](https://dms.mpeg.expert/doc_end_user/current_document.php?id=90266&id_meeting=196).
* Joint meetings were expected with AG 5 (on Monday 16 Oct.) and possibly with other groups.
* Principles of standards development were discussed.
* Lunch: Oct. 13-15 1200-1400 in the conference center. Lunch break 1230-1330 on the first three days. Oct. 16-20 in the conference center, JVET shall use the time slot 1245-1330 to avoid congestion with other WGs.

## 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 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 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:

Coffee breaks 1030, 1530

* Fri. 13 October, 1st day
  + Morning session:
    - 0900–1000 Opening remarks, review of practices, agenda, IPR policy reminder
    - 1000–1240 Reports of AHGs 1–11, 14
  + Afternoon session:
    - 1340–1400 Reports of AHGs 12, 13
    - 1400–1900 EE2 summary report
* Sat. 14 October, 2nd day
  + Morning session:
    - 0830–1230 Review of EE1 summary report and related
  + Afternoon session:
    - 1330–1900 Review of EE2.4 summary (RPR) EE2-related and non-EE2 (Leibniz rm., track chaired by Y. Ye)
    - 1330-1900 Review of HLS 6.1, 6.2 (rm. 27, track chaired by J. Ohm)
* Sun. 15 October, 3rd day
  + Morning session:
    - 0830–1230 Review of EE2-related and non-EE2 (Leibniz rm., track chaired by Y. Ye)
    - 0830–1230 Review of HLS/SEI 6.4 (rm. 27, track chaired by JRO)
  + Afternoon session:
    - 1330–1900 Review of EE2-related and non-EE2 (Leibniz rm., track chaired by Y. Ye)
    - 1330–1800 Review of HLS/SEI 6.3, 6.6ff. (rm. 27, track chaired by JRO)
* Mon. 16 October, 4th day
  + 0900–1230 MPEG information sharing session
  + 1300–1600 BoG on NNVC HOP LF (E. Alshina, F. Galpin)
  + 1610-1755 JVET plenary: Coordination, doc review ECM related (5.2.4)
  + 1800-2200 JVET session: HLS 6.7, 6.6, 6.8, 6.9 (Future Meeting Space B)
* Tue 17 October, 5th day
  + Morning session:
    - 0830-1100 BoG on NNVC HOP LF (E. Alshina, F. Galpin) (Leibniz)
    - 0830–1100 HLS/SEI 6.8/6.9 (Neuer Saal)
    - 1230-1345 remaining EE1 (Leibniz)
  + 1100-1215 Joint meeting with WG2/WG7/VCEG: Generative face video (Room 27)
  + Afternoon session:
    - 1400–1800 JVET session: 4.10, HLS/SEI 6.9 (track chaired by JRO)
    - 1400–1800 JVET session: ECM/CTC 5.2.5, encoder opt. 4.11 (track chaired by Y. Ye)
* Wed. 18 October, 6th day
  + 0900–1030 MPEG information sharing session
  + 1030-1100 Joint meeting with WG4/VCEG: MPI/MIV relationship (Future Meeting Space B)
  + 1100-1200 Joint meeting with WG4/AG5/VCEG: Studies of multi-layer coding (4.5) (Future Meeting Space B)
  + Afternoon session:
    - 1300–1830 Coordination, NNVC (BoG report+1 doc), Review 4.x (4.12 film grain incl. SEI @1600), remaining 6.9 (Leibniz room)
  + Social event starts 1900
* Thu. 19 October, 7th day
  + Morning session (Leibniz room):
    - 0830–0900 Outgoing liaison review
    - 0900–1315 Remaining doc review, revisits
  + Afternoon session (Leibniz room):
    - 1400–1800 Review 4.1, 4.13, 4.14, review draft DoCRs, planning of output docs and AHGs, a.o.b.
  + 1800-2200 Evening sessions possible
* Fri. 20 Oct., 8th day
  + 0900–1225 JVET wrap-up plenary (Leibniz room):
    - Approval of output docs
    - Establishment of AHGs
    - Review of meeting recommendations
    - Future planning, a.o.b.
  + 1400–1600 MPEG information sharing session
  + 1716–1724 WG 5 approval of meeting recommendations, closing of meeting

## Contribution topic overview

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

* AHG reports (14) (section 3)
* Project development (section 4)
  + AHG1: Deployment and advertisement of standards (3) TBP
  + AHG2: Text development and errata reporting (1)
  + AHG3: Test conditions (0)
  + AHG3: Software development (0)
  + AHG4: Subjective quality testing and verification testing (4)
  + AHG4: Test Material (2)
  + AHG4: Codec performance with alternative test material (1)
  + AHG5: Conformance test development (0)
  + AHG7: ECM tool assessment (0)
  + AHG8: Optimization of encoders and receiving systems for machine analysis of coded video content (3)
  + AHG10: Encoding algorithm optimization (3)
  + AHG13: Film grain synthesis (5) 1 TBP
  + Implementation studies (1) TBP
  + Profile/tier/level specification (1)
  + General aspects of standards development and applications of standards (1)
* Low-level tool technology proposals (section 5) with subtopics (number counts excluding BoG and summary reports)
  + AHG11/AHG14 and EE1: Neural network-based video coding (28) (section 5.1)
  + AHG6/AHG12 and EE2: Enhanced compression beyond VVC capability (78) (section 5.2)
* AHG9: High-level syntax (HLS) proposals (section 6) with subtopics
  + SEI messages on neural-network post filter (13) (sections 6.1, 6.2, 6.3) 1 TBP
  + SEI messages on topics other than NNPF (31) (sections 6.4, 6.5, 6.6, 6.7, 6.8, 6.9)
  + Non-SEI HLS aspects (0) (section 6.10)
* Joint meetings, plenary discussions, BoG reports (X), liaison (1), 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 XXXX–XXXX on Friday 13 Oct. 2023 (chaired by JRO).

[JVET-AF0001](https://jvet-experts.org/doc_end_user/current_document.php?id=13372) JVET AHG report: Project Management (AHG1) [J.-R. Ohm (chair), G. J. Sullivan (vice chair)]

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

The number of subscribers (by the beginning of the current meeting) was 1233. Furthermore, the previous lists of joint teams (which were still kept open as archives) had the following number of subscribers:

* JCT-VC – 1159 subscribers
* JCT-3V – 682 subscribers
* JVT-experts – 2068 subscribers

1. **Goals and activity**

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

Output documents from the preceding meeting had been made initially available at the JVET web site (<https://jvet-experts.org/>) or the ITU-based JVET site (<http://wftp3.itu.int/av-arch/jvet-site/2023_07_AE_Geneva/>). 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:

a) JVET documents

* The meeting report (JVET-AE1000) [Posted 2023-08-16, also submitted as WG 5 N 215]
* [JVET-AE1004](https://jvet-experts.org/doc_end_user/current_document.php?id=12567) Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP [Posted 2023-07-26]
* [JVET-AE1006](https://jvet-experts.org/doc_end_user/current_document.php?id=12568) New profiles, colour decriptors, and SEI messages for HEVC (draft 1) [Posted 2023-08-21, last update 2023-09-04, also issued as WG 5 CDAM N 226]
* [JVET-AE1011](https://jvet-experts.org/doc_end_user/current_document.php?id=12569) HEVC multiview profiles supporting extended bit depth (draft 2) [Posted 2023-09-23]
* [JVET-AE1013](https://jvet-experts.org/doc_end_user/current_document.php?id=12569) Common test conditions of 3DV experiments [Posted 2023-10-12]
* [JVET-AE1016](https://jvet-experts.org/doc_end_user/current_document.php?id=12568) AVC with extensions and corrections (draft 1) [Posted 2023-08-16, also issued as WG 5 CD of AVC 11th edition N 218]
* [JVET-AE2005](https://jvet-experts.org/doc_end_user/current_document.php?id=12574) New level and systems-related supplemental enhancement information for VVC (Draft 6) [Posted 2023-08-11, last update 2023-10-03]
* [JVET-AE2006](https://jvet-experts.org/doc_end_user/current_document.php?id=12574) Additional SEI messages for VSEI (Draft 5) [Posted 2023-07-21, last update 2023-08-23]
* [JVET-AE2016](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 2023-08-01]
* [JVET-AE2017](https://jvet-experts.org/doc_end_user/current_document.php?id=12576) Common test conditions and evaluation procedures for enhanced compression tool testing [Posted 2023-08-01]
* [JVET-AE2019](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 4 [Posted 2023-09-01, last update 2023-09-04]
* [JVET-AD2020](https://jvet-experts.org/doc_end_user/current_document.php?id=12579) Film grain synthesis technology for video applications (Draft 5) [Posted 2023-09-222, last update 2023-09-28, also to be issued as WG 5 DTR N 222]
* [JVET-AE2021](https://jvet-experts.org/doc_end_user/current_document.php?id=12578) Verification test plan for VVC multilayer coding (update 1) [Posted 2023-10-06]
* [JVET-AE2023](https://jvet-experts.org/doc_end_user/current_document.php?id=12560) Exploration experiment on neural network-based video coding (EE1) [Posted 2023-07-17, last update 2023-08-10]
* [JVET-AE2024](https://jvet-experts.org/doc_end_user/current_document.php?id=12562) Exploration experiment on enhanced compression beyond VVC capability (EE2) [Posted 2023-07-18, last update 2023-08-22]
* [JVET-AE2025](https://jvet-experts.org/doc_end_user/current_document.php?id=12580) Algorithm description of Enhanced Compression Model 10 (ECM 10) [Posted 2023-10-02]
* [JVET-AE2027](https://jvet-experts.org/doc_end_user/current_document.php?id=12582) SEI processing order SEI message in VVC (Draft 5) [Posted 2023-09-02, also issued as WG 5 preliminary WD N 229]
* [JVET-AE2028](https://jvet-experts.org/doc_end_user/current_document.php?id=12583) Additional conformance bitstreams for VVC multilayer configurations [Posted 2023-09-01]
* [JVET-AE2030](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 3) [Posted 2023-09-14, also issued as WG 5 preliminary WD N 224]
* [JVET-AE2031](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 2023-08-11]
* [JVET-AE2032](https://jvet-experts.org/doc_end_user/current_document.php?id=12584) Technologies under consideration for future extensions of VSEI (draft 1) [Posted 2023-09-15]

b) documents produced as WG 5 documents only:

* WG 5 N 217 Request for ISO/IEC 14496-10 11th edition
* WG 5 N 219 Disposition of comments received on ISO/IEC 23002-7:2022 DAM 1
* WG 5 N 220 Text of ISO/IEC FDIS 23002-7:202x Versatile supplemental enhancement infor­mation messages for coded video bitstreams (3rd edition) (to be delivered)
* WG 5 N 222 Disposition of comments received on ISO/IEC CDTR 23002-9
* WG 5 N 225 Request for ISO/IEC 23008-2:202x/Amd.1
* WG 5 N 227 Disposition of comments received on ISO/IEC 23090-3:2022 DAM 1
* WG 5 N 228 Text of ISO/IEC FDIS 23090-3:202x Versatile video coding (3rd edition) (to be delivered)
* Liaison statement to SC 29/WG 1 (JPEG) on JPEG AI, neural network-based video coding, and SEI messages for neural network-based post filtering (WG 5 N 234)
* List of AHGs established at the 12th WG 5 meeting (WG 5 N 235)

c) draft revised ITU-T Recommendations forwarded by JVET and Q6/16 for ITU-T Consent:

TD169/Plen ITU-T H.265 (V9) "High efficiency video coding" (Rev.)

TD171/Plen ITU-T H.266 (V3) "Versatile video coding" (Rev.)

TD159/Plen ITU-T H.266.1 (V2) "Conformance specification for ITU-T H.266 versatile video coding" (Rev.)

TD172/Plen ITU-T H.273 (V3) "Coding-independent code points for video signal type identification" (Rev.)

TD170/Plen (A.5: TD173/Plen) ITU-T H.274 (V3) "Versatile supplemental enhancement information messages for coded video bitstreams" (Rev.)

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 finally been delivered 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 32nd meeting had been announced in the JVET reflector, in the JVET logistics document (<https://www.itu.int/wftp3/av-arch/jvet-site/2023_10_AF_Hannover/JVET-AF_Logistics.docx>), and in the WG 5 calling notice (N 236) and agenda (N 237) for the 13th 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 180 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 32nd meeting had been made publicly available on the ITU-hosted ftp site as <http://wftp3.itu.int/av-arch/jvet-site/2023_10_AF_Hannover/JVET-AF_notes_d0.docx>.

1. **Recommendations**

The AHG recommends its continuation.

[JVET-AF0002](https://jvet-experts.org/doc_end_user/current_document.php?id=13373) JVET AHG report: Draft text and test model algorithm description editing (AHG2) [B. Bross, C. Rosewarne (co-chairs), F. Bossen, A. Browne, S. Kim, S. Liu, J.-R. Ohm, G. J. Sullivan, A. Tourapis, Y.-K. Wang, Y. Ye (vice chairs)]

1. **Ad hoc group activity**
   1. **Output documents produced**
      1. **JVET-AE1004 - 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-AE meeting:

* Updated the general descriptions for VVC, VSEI, HEVC, and AVC.
* Removed obsolete/resolved errata items.
* Added 16 items for the ISO text only for HEVC, identified when working on the ITU-T consent text for H.265 version 9 (edition 5).
* Added 9 items on the semantics of post-filter hint SEI message from JVET-AE0155 for both HEVC and AVC.
* Added 4 items for AVC, on semantics of the omnidirectional video SEI messages. These items were among the items for HEVC and have been resolved for HEVC.
* Added an item for AVC on use of the "should" wording for the use of the film grain characteristics SEI messages when present.
* Removed the section for codepoint usage TR as the TR part has been dropped from the title of the document a while ago.
* Added an item for HEVC regarding replacing of "Max slice segments per picture MaxSliceSegmentsPerPicture" in Table A.8 with "Max # of slice segments per picture MaxSliceSegmentsPerPicture" (reported by [Cliff Reader](mailto:cliff@reader.com) – thanks!).
  + 1. **JVET-AE1006 - New profiles, colour descriptors, and SEI messages for HEVC (draft 1)**

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

Draft 1 incorporated items:

* Changes for NNPF and phase indication SEI messages (based on JVET-AE0101 as well as the latest changes to the interface text for the NNPF SEI messages agreed at the JVET-AE meeting
* JVET-AD1008 Additional colour type identifiers
* Fixes and changes for the film grain characteristics SEI message
* New multiview profiles from JVET-AE0296, plus an additional Multiview Monochrome 10 profile
* Some errata items to subclauses A.4.1, A.4.2, G.11.2.3, and H.11.2.2 reported by Cliff Reader (thanks!)
* Changes to D.1 related to referring to VSEI in HEVC
  + 1. **JVET-AE1016 – AVC with extensions and corrections (Draft 1)**

This document contains the draft text for changes of the Advanced Video Coding (AVC) standard (Rec. ITU-T H.264 | ISO/IEC 14496-10). Text modifications are provided for 1) the support of additional colour type identifiers; 2) some technical corrections and editorial improvements to the 10th edition text of AVC; and 3) the AVC-specific supports for some supplemental enhancement information (SEI) messages that may be included in AVC bitstreams but are not to be specified in the AVC specification. These SEI messages include three other SEI messages, namely the neural network post-filter characteristics SEI message, the neural-network post-processing filter activation SEI message, and the phase indication 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).

The affected sections are provided relative to the 2022 edition of ISO/IEC 14496-10, i.e. the 10th edition.

Draft 1 incorporated items:

* JVET-AD1008 Additional colour type identifiers for AVC
* JVET-AE0101 Use of neural network post filter and phase indication SEI messages for AVC
* JVET-AE1004 Errata items on AVC
* Fix for film grain synthesis equation in FGC SEI message (see notes under BoG report JVET-AE0294)
  + 1. **JVET-AE2005 - New level and systems-related supplemental enhancement information for VVC (Draft 6)**

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 four 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, and the phase indication 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 6 incorporated items:

* Fixed bug tickets [#1594](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1594), [#1602](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1602) and [#1606](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1606).
* Allow multiple NNPFC and NNPFA SEI messages with different content in the same picture unit by excluding NNPFC and NNPFA SEI messages from the existing constraint in JVET-AD2005, subclause D.3.1. The NNPFA SEI message is constrained in a manner that the payload shall be identical for the same nnpfa\_target\_id value. Editors are requested to make sure the blanket constraint in VVC on constraining SEI payload content to be the same within a PU does not impose a problem for SEI messages that contain ID values. No such blanket constraint appears to exist in HEVC, and this was intentional for some SEI messages such as tone mapping information and colour remapping information SEI messages. (JVET-AE0049 item 1).
* Constrain the repeated inference to generate interpolated pictures up to the end of the bitstream only to NNPFs that perform only picture rate upsampling. (JVET-AE0049 item 2)
* Change the index range in input picture assignment for repeated inference at the end of a bitstream: it is proposed to remove "− j" from "When numInputPics is greater than 1, the following applies for each value of i in the range of j + 1 to numInputPics − j − 1, inclusive, in increasing order of i:" (JVET-AE0049 item 3, JVET-AE0142 item 1)
* In the following sentence in the NNPFC interface text, replace "inputPic" with "resampledPic": If sourceWidth is equal to CroppedWidth and sourceHeight is equal to CroppedHeight, inputPic is set to be the same as sourcePic. (JVET-AE0142 item 2)
* In the NNPFC interface text, condition the following constraint on "when pictureRateUpsamplingFlag is equal to 1", and num\_interpolated\_pics[ i − 1 ] shoud be nnpf\_interpolated\_pics[ i − 1 ]: It is a requirement of bitstream conformance that num\_interpolated\_pics[ i − 1 ] shall not be greater than 0. (JVET-AE0142 item 3)
* VVC changes corresponding to the addition of nnpfa\_no\_prev\_clvs\_flag. (JVET-AE0050 item 1)
* VVC changes corresponding to the addition of nnpfa\_no\_foll\_clvs\_flag. (JVET-AE0050 item 2)
* At the end of a bitstream, add repeated inference of NNPF for creating pictures corresponding to input pictures (regardless of the purpose indicated by nnpfc\_purpose except for picture rate upsampling) when there are multiple input pictures to which the NNPF applies, with an update to use the nnpfa\_output\_flag rather than the nnpfc\_input\_pic\_output\_flag. (JVET-AE0134 item 2)
* Add a constraint based on nnpfa\_output\_flag[ i ], as part of the process for derivation of the NNPF input pictures in the interface text, to ensure that input pictures that are not present in the bitstream do not have a corresponding NNPF output picture. (JVET-AE0070 item 2)
* Editorial changes (JVET-AE0127 attachment)
* Do not include the post-filter hint SEI message in VVC v3 due to its apparent lack of clarity. (JVET-AE0155)
  + 1. **JVET-AE2006 - Additional SEI messages for VSEI (Draft 5)**

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 for VSEI v3 and JVET-AE2006:**

* Add missing value ranges for resampling syntax elements nnpfc\_pic\_width\_num\_minus1, nnpfc\_pic\_width\_denom\_minus1, nnpfc\_pic\_height\_num\_minus1, and nnpfc\_pic\_height\_denom\_minus1 to be in the range of 0 to 65535, inclusive. (JVET-AE0048 item 1)
* Add missing value ranges for nnpfc\_inp\_format\_idc and nnpfc\_out\_format\_idc to be in the range of 0 to 255, inclusive. (JVET-AE0048 item 2).
* Arrange the definitions of DeriveInputTensors(), the syntax elements related to overlap and padding type, and the definition of the InpSampleVal function so that forward references are avoided by moving the derivations of strengthControlScaledVal (Formula 87), DeriveInputTensor (Formula 88), and StoreOutputTensor (91) to a later position that immediately follows the derivation of the InpSampleVal function. (JVET-AE0048 item 3)
* Clarify that indicated or inferred values of vui\_colour\_primaries, vui\_transfer\_characteristics, vui\_matrix\_coeffs, and vui\_full\_range\_flag apply when nnpfc\_separate\_colour\_description\_present\_flag is equal to 0. Here, the possibility of inferring these syntax elements is made explicit when compared to JVET-AD2006. (JVET-AE0048 item 4)
* Signal nnpfc\_chroma\_loc\_info\_present\_flag only when nnpfc\_out\_order\_idc is not equal to 0. Infer nnpfc\_chroma\_loc\_info\_present\_flag to be equal to 0 when not present. (JVET-AE0060 item 1)
* Clarify using the approach in the following suggestion: “nnpfc\_mode\_idc equal to 0 indicates that the stuff is in the SEI message. nnpfc\_mode\_idc equal to 1 indicates the stuff is in a location identified by a URI. When nnpfc\_mode\_idc is equal to 0, the stuff is in the format specified by ISO/IEC 15938-17.” (JVET-AE0175)
* Change nnpfc\_reserved\_zero\_bit(\_a/\_b) to nnpfc\_alignment\_zero\_bit(\_a/\_b). (JVET-AE0126 item 2)
* Change the condition for the constraint on either nnpfcOutputPicWidth is not equal to CroppedWidth or nnpfcOutputPicHeight is not equal to CroppedHeight from "when nnpfc\_pic\_width\_num\_minus1, nnpfc\_pic\_width\_denom\_minus1, nnpfc\_pic\_height\_num\_minus1, and nnpfc\_pic\_height\_denom\_minus1 are present" to "when resolutionResamplingFlag is equal to 1". (JVET-AE0126 item 5)
* Move nnpfa\_target\_base\_flag to be immediately after nnpfa\_persistence\_flag, such that the persistence flag, when present, immediately follows the cancel flag. (JVET-AE0126 item 7)
* State that “All NNPFC SEI messages in a CLVS that have a particular nnpfc\_id value and nnpfc\_base\_flag equal to 1 shall have identical SEI payload content.” (or editorial equivalent) (JVET-AE0128 item 2)
* Remove ChromaFormatIdc from the constraint for when colourizationFlag shall be equal to 0. (JVET-AE0141)
* Cancel the persistence of an NNPFA SEI message with any subsequent NNPFA SEI message with the same nnpfa\_target\_id. (JVET-AE0048 item 5)
* Update the constraint related to nnpfcTargetPictures and nnpfaTargetPictures by also considering the value of nnpfa\_target\_base\_flag. (JVET-AE0189 item 3)
* Add the following constraint: When nnpfa\_target\_base\_flag in an NNPFA SEI message is equal to 0, there shall be at least one NNPFC SEI message with nnpfc\_id equal to nnpfa\_target\_id and nnpfc\_base\_flag equal to 0 that precedes the NNPFA SEI message in decoding order. (JVET-AE0189 item 2)
* Add nnpfa\_no\_prev\_clvs\_flag to the NNPFA SEI message. (JVET-AE0050 item 1)
* Add nnpfa\_no\_foll\_clvs\_flag to the NNPFA SEI message, gated by nnpfa\_persistence\_flag being equal to 1. (JVET-AE0050 item 2)
* Disallow generating NNPF output pictures between any particular pair of consecutive input pictures more than once. (JVET-AE0135 item 2)
* Add a constraint to ensure that there is at least one NNPF output picture when an NNPF is activated and applied. (JVET-AE0069)
* Specify that all NNPFC SEI messages with a particular value of nnpfc\_id within a CLVS shall have the same value of nnpfc\_purpose. (JVET-AE0142 item 4)
* Do not include the post-filter hint SEI message in VSEI v3 due to its apparent lack of clarity. (JVET-AE0155)
* Editorial changes (JVET-AE0048 item 6, JVET-AE0126 items 3, 4, 6, JVET-AE0126 attachment, and ticket #1605, JVET-AE0054)
* Use “should” for film grain synthesis and retain prior drafted language for other SEI messages. (JVET-AE0155)
* Editors are requested to make sure it is clear that activating an “update” can (and basically always would) require using information from a “base” filter in order to interpret the update. (JVET-AE0189 item 1)
* Clarify, as necessary, that an NNPFA message can activate a base with a particular ID when an update is active, which will switch the filter from the updated filter to the base filter. (JVET-AE0173)
* Add a NOTE to clarify an implied, hidden constraint on picture-rate-upsampling NNPFs with persistent activation, with expansion to consider all possible cases that have a similar implication: (JVET-AE0135 item 1)
* Editors are requested to check/clarify the existing intent that if the current picture contains an SEI message (any SEI message that uses the typical persistence scope scheme) with a cancel flag equal to 1, the current picture is not within the scope of the previous SEI message (regardless of whether the SEI message is a suffix or prefix SEI message). (JVET-AE0048 item 5)
* Text related to the remaining revisit from the BoG report JVET-AE0272 was presented on Tuesday 18 July at 1035. This is related to the issue that for frame interpolation between frames A and B only one activation shall be performed for all frames that are to be interpolated altogether. It was commented that the purpose may be difficult to understand for implementers. It was suggested to add a NOTE explaining the purpose of that constraint.
  + 1. **JVET-AE2027 - SEI processing order SEI message in VVC (draft 5)**

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:**

none

**Changes that have been integrated:**

Updates from JVET-AE0156. Leave out the change in CVS (for further study), and add a note about the usage of wrapping.

1. **Related input contributions**

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

* [JVET-AF0045](https://jvet-experts.org/doc_end_user/current_document.php?id=13289) AHG2/AHG9: On the use of NNPF SEI messages in AVC
* [JVET-AF0063](https://jvet-experts.org/doc_end_user/current_document.php?id=13311) On MV-HEVC profiles
* [JVET-AF0064](https://jvet-experts.org/doc_end_user/current_document.php?id=13312) AHG2: Some editorial issues in AVC, HEVC, VVC, and VSEI

1. **Remaining bug tickets**

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

1. **Recommendations**

The AHG recommends to:

* Approve JVET-AE1004, JVET-AE1006, JVET-AE1016, JVET-AE2005, JVET-AE2006, and JVET-AE2027 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-AF0003](https://jvet-experts.org/doc_end_user/current_document.php?id=13374) 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 22.0](https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/-/releases/VTM-22.0) (Sept. 2023)
* [HM-18.0](https://vcgit.hhi.fraunhofer.de/jvet/HM/-/releases/HM-18.0) (Apr. 2023)
* [HM-16.21+SCM-8.8](https://vcgit.hhi.fraunhofer.de/jvet/HM/-/tags/HM-16.21+SCM-8.8) (Mar. 2020)
* [SHM 12.4](https://vcgit.hhi.fraunhofer.de/jvet/SHM/-/tags/SHM-12.4) (Jan. 2018)
* [HTM 16.3](https://vcgit.hhi.fraunhofer.de/jvet/HTM/-/tags/HTM-16.3) (Jul. 2018)
* [JM 19.1](https://vcgit.hhi.fraunhofer.de/jvet/JM/-/releases/JM-19.1) (Apr. 2023)
* [JSVM 9.19.15](https://vcgit.hhi.fraunhofer.de/jvet/jsvm/-/tags/JSVM_9_19_15)
* [JMVC 8.5](https://vcgit.hhi.fraunhofer.de/jvet/jmvc/-/tags/JMVC_8_5)
* [3DV ATM 15.0](https://vcgit.hhi.fraunhofer.de/jvet/3dv-atm/-/tags/3DV-ATM_v15.0) (no version history)
* [HDRTools 0.24](https://gitlab.com/standards/HDRTools/-/tags/v0.24) (March 2023)

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

1. **Software development**

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

The server is located at:

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

The registration and development workflow are documented at:

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

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

1. **VTM related activities**

The VTM software can be found at

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

The software development continued on the GitLab server. VTM versions 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 21.1 was tagged on July 21, 2023. Changes include:

* JVET-AD0057 item 1: Allow inclusion of NNPFC and NNPFA SEI messages in suffix SEI NAL units
* JVET-AD0054 item 3: NNPF interpolated pictures constraint
* Add encoder parameter WriteVuiHrdFromY4m
* fix RPL with inter-layer references

VTM 21.2 was tagged on July 22, 2023. Changes include:

* Remove macros from previous cycle

VTM 22.0 was tagged Sept. 24, 2023. Changes include:

* remove macro JVET\_Z0120\_SII\_SEI\_PROCESSING
* Clean up data types in ALF code
* Only one VPS can be active at any time
* JVET-AE0060: Signal nnpfc\_chroma\_loc\_info\_present\_flag only when nnpfc\_out\_order\_idc is not equal to 0
* JVET-AE0180: Add missing condition on PIC\_RECON\_WRAP buffer allocation
* JVET-AE0057: MTT split modes early termination
* JVET-AE0162: fix MS-SSIM calculation for RPR/SR
* Use PLTRunMode instead of bool for palette data
* Fix access to unitialized value in SPS
* Clean up code from JVET-AD0045 (encoder DMVR avoidance)
* Use std::vector to avoid memory deallocation/reallocation
* Port JVET-Z0150 on memory usage log
* Add TGM config files from ECM
* Remove unnecessary memory allocation for CCALF
* JVET-AE0141: Fix a bug in NNPFC SEI message for colourization
* Fix overflow condition in static\_vector
* JVET-AE0135 item2: On NNPF picture rate upsampling constraint
* JVET-AE0126: NNPF cleanup and editorial changes for VSEI include item 2, item 3, item 5, and item 7
* JVET-AE0142 item 3 and item 4: Fix two bugs including an added condition and a nnpfc\_purpose related constraint.
* JVET-AE0128 item 2: Update the constraint on when an NNPFC SEI message shall be a repetition of the first NNPFC SEI message in the CLVS
* JVET-AE0122: Fix for ppsid when both GOP based RPR and scalable coding is used
* JVET-AE0048 item 1: Add missing value ranges for nnpfc\_pic\_width\_num\_minus1,...
* JVET-AE0048 item 5: Cancel the persistence of an NNPFA SEI message with any...
* JVET\_AE0048 item 2: Add missing value ranges for nnpfc\_inp\_format\_idc and nnpfc\_out\_format\_idc
* Clean up CommonDefX86.h
* JVET-AE0049 item 2: Constrain the repeated inference to NNPFs that perform...
* JVET-AE0181: Scaling window support
* JVET-AE0189 Item 2: Check when activating NNPFA with base flag equal to 0
* JVET-AE0050 item 1: Add nnpfa\_no\_prev\_clvs\_flag to the NNPFA SEI message
* Clean up AffineGradientSearch SIMD code
* JVET-AE0049 item 3: Change the index range in input picture assignment for...
* JVET-AE0049 item 1: Update NNPF SEI identical content constraints
* JVET-AE0050 item 2: Add nnpfa\_no\_foll\_clvs\_flag to the NNPFA SEI message
* JVET-AE0134 item 2: add repeated inference of NNPF for creating pictures corresponding to input pictures at the end of a bitstream or a CLVS under some conditions
* Reduce memory for ALF covariance data

VTM 22.1 is expected to be tagged during the 32nd JVET meeting. Changes are expected to include the remaining implementations of the 31st meeting, issue fixes and code optimizations.

* 1. ***CTC Performance***

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

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

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

There is no change in coding performance or run time between **VTM 22.0** and **VTM 21.1** using SDR CTC.

For the high bit depth CTCs, there is no change in coding performance or run time between VTM 22.0 and VTM 21.0 for the low QP range; however, a change in coding performance and run-time had been observed for the standard QP range (22-37), as shown in the summary tables below. The coding performance changes were attributed to an optimisation of ALF causing small floating point calculation discrepancies (ALF is turned off for the low QP range experiments).

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access** | | | | | | | | | |
|  | **VTM 22.0 vs VTM21.0 for High Bit Depth, Standard QP range (JVET-AA2018)** | | | | | | | | | |
|  |  |  | **wPSNR** |  |  | **PSNR** |  |  |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | 0.17% | 0.02% | 0.00% | 0.12% | 0.23% | 0.00% | 0.08% | 0.21% | 99% | 105% |
| Class H2 |  |  |  |  |  | 0.00% | 0.00% | 0.00% | 99% | 104% |
| **Overall** | 0.17% | 0.02% | 0.00% | 0.12% | 0.23% | 0.00% | 0.04% | 0.11% | 99% | 105% |
|  |  |  |  |  |  |  |  |  |  |  |
|  | **All Intra** | | | | | | | | | |
|  | **VTM 22.0 vs VTM21.0 for High Bit Depth, Standard QP range (JVET-AA2018)** | | | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | 0.06% | 0.00% | 0.00% | 0.09% | 0.08% | 0.00% | 0.05% | 0.07% | 98% | 102% |
| Class H2 |  |  |  |  |  | 0.00% | 0.00% | 0.00% | 100% | 102% |
| **Overall** | 0.06% | 0.00% | 0.00% | 0.09% | 0.08% | 0.00% | 0.03% | 0.03% | 99% | 102% |

The coding performance differences had been almost eliminated in a pending merge request (!2682), as shown in the table below:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access** | | | | | | | | | |
|  | **(VTM 22.0 + !2682) vs VTM21.0 for High Bit Depth, Standard QP range (JVET-AA2018)** | | | | | | | | | |
|  |  |  | **wPSNR** |  |  | **PSNR** |  |  |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | -0.01% | 0.01% | 102% | 107% |
| Class H2 |  |  |  |  |  | 0.00% | 0.00% | 0.00% | 102% | 107% |
| **Overall** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | -0.01% | 0.00% | 102% | 107% |
|  |  |  |  |  |  |  |  |  |  |  |
|  | **All Intra** | | | | | | | | | |
|  | **(VTM 22.0 + !2682) vs VTM21.0 for High Bit Depth, Standard QP range (JVET-AA2018)** | | | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 101% | 104% |
| Class H2 |  |  |  |  |  | 0.00% | 0.00% | 0.00% | 102% | 103% |
| **Overall** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 101% | 103% |

Results using HDR CTC show slight differences between **VTM 22.0** and **VTM 21.0**, as depicted in table below.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access** | | | | | | | | | |
|  | **Over VTM21.0** | | | | | | | | | |
|  |  |  | **wPSNR** |  |  | **PSNR** |  |  |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | 0.08% | 0.02% | 0.02% | 0.25% | 0.19% | 0.01% | 0.22% | 0.20% | 101% | 101% |
| Class H2 |  |  |  |  |  | 0.00% | 0.00% | 0.00% | 101% | 101% |
| **Overall** | 0.08% | 0.02% | 0.02% | 0.25% | 0.19% | 0.00% | 0.14% | 0.13% | 101% | 101% |
|  |  |  |  |  |  |  |  |  |  |  |
|  | **All Intra** | | | | | | | | | |
|  | **Over VTM21.0** | | | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | 0.08% | 0.00% | 0.00% | 0.09% | 0.06% | 0.00% | 0.06% | 0.04% | 99% | 100% |
| Class H2 |  |  |  |  |  | 0.00% | 0.00% | 0.00% | 98% | 99% |
| **Overall** | 0.08% | 0.00% | 0.00% | 0.09% | 0.06% | 0.00% | 0.04% | 0.02% | 98% | 100% |

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

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access** | | | | | | | | | |
|  | **Over HM18.0** | | | | | | | | | |
|  |  |  | **wPSNR** |  |  | **PSNR** |  |  |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | -39.51% | -37.55% | -36.66% | -54.89% | -48.36% | -33.68% | -49.39% | -40.75% | 289% | 102% |
| Class H2 |  |  |  |  |  | -31.97% | -57.69% | -63.36% | 257% | 90% |
| **Overall** | -39.51% | -37.55% | -36.66% | -54.89% | -48.36% | -33.06% | -52.41% | -48.97% | 277% | 97% |
|  |  |  |  |  |  |  |  |  |  |  |
|  | **All Intra** | | | | | | | | | |
|  | **Over HM18.0** | | | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | -41.41% | -27.31% | -26.79% | -57.69% | -52.79% | -24.00% | -52.89% | -45.12% | 1506% | 117% |
| Class H2 |  |  |  |  |  | -21.76% | -47.21% | -50.57% | 1280% | 109% |
| **Overall** | -41.41% | -27.31% | -26.79% | -57.69% | -52.79% | -23.18% | -50.83% | -47.10% | 1420% | 114% |

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

The following issues in VTM master branch may affect conformance:

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

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

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

1. **HM related activities**

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

The following MRs are pending [with status indicated]:

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

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

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

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

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

Help to address these tickets would be appreciated.

1. **360Lib related activities**

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

The following table is for the projection formats comparison using VTM-22.0 according to 360-degree video CTC (JVET-U2012) compared to that using VTM-21.0 (VTM-21.0 as anchor).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **End-to-end WS-PSNR** | | | **End-to-end S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| Class S2 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| **Overall** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |

The following table compares generalized cubemap (GCMP) coding and padded equi-rectangular projection (PERP) coding using VTM-22.0 (PERP as anchor).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **End-to-end WS-PSNR** | | | **End-to-end S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | -11.50% | -5.44% | -6.08% | -11.49% | -5.37% | -6.02% |
| Class S2 | -3.69% | 0.81% | 1.28% | -3.67% | 0.90% | 1.37% |
| **Overall** | -8.38% | -2.94% | -3.13% | -8.36% | -2.86% | -3.07% |

The following tables are for PERP and GCMP coding comparison between VTM-22.0 and HM-16.22 (HM as anchor), respectively.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **End-to-end WS-PSNR** | | | **End-to-end S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | -30.91% | -38.82% | -41.28% | -30.91% | -38.88% | -41.28% |
| Class S2 | -36.89% | -37.33% | -39.65% | -36.88% | -37.36% | -39.71% |
| **Overall** | -33.30% | -38.23% | -40.63% | -33.29% | -38.27% | -40.65% |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **End-to-end WS-PSNR** | | | **End-to-end S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | -35.60% | -40.82% | -42.80% | -35.56% | -40.75% | -42.75% |
| Class S2 | -39.84% | -39.52% | -41.44% | -39.85% | -39.52% | -41.48% |
| **Overall** | -37.29% | -40.30% | -42.25% | -37.27% | -40.26% | -42.24% |

1. **SCM related activities**

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

1. **SHM related activities**

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

1. **HTM related activities**

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

One merge request is pending:

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

The next release will include the following changes:

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

1. **HDRTools related activities**

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

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

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

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

1. **Bug tracking**

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

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

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

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

Bug tracking for HDRTools is located at:

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

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

1. **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.

1. **CTC alignment and merging**

There are currently 8 JVET CTC documents:

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

JVET-Z2011 VTM/HM HDR test conditions

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

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

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

JVET-U2012 VTM 360 video test conditions

JVET-AC1009 SHVC test conditions

JVET-AC1015 SCM test conditions

JVET-AE1013 3DV test conditions

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

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

1. **Guidelines for reference software development**

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

1. **Recommendations**

The AHG recommends to:

* Continue to develop reference software.
* Improve documentation, especially the software manual.
* Encourage people to test VTM and other reference software more extensively outside of common test conditions.
* Encourage people to report all (potential) bugs that they are finding.
* Encourage people to submit bit-streams/test cases that trigger bugs in VTM and other reference software.
* Encourage people to submit non-normative changes that either reduce encoder run time without significantly sacrificing compression performance or improve compression performance without significantly increasing encoder run time.
* Design and add configuration files to the VTM software for testing of HLS features.
* Review VTM-related contributions and determine whether features should be added (or removed) from the software.
* Continue to investigate the merging of branches.
* Continue to investigate merging of CTC documents.
* Verify correctness of CTC documents and issue updates as appropriate
* Keep common test conditions aligned for the different standards.
* Consider documents (including late documents) related to AHG3 activities.

The software coordinators expressed thanks to experts who diligently delivered implementations of HLS related software in a timely fashion.

[JVET-AF0004](https://jvet-experts.org/doc_end_user/current_document.php?id=13375) JVET AHG report: Test material and visual assessment (AHG4) [V. Baroncini, T. Suzuki, M. Wien (co-chairs), W. Husak, S. Iwamura, P. de Lagrange, S. Liu, S. Puri, A. Segall, S. Wenger (vice-chairs)]

1. **Activities**
   1. ***Verification test preparation for VVC multilayer coding***

A teleconference was organized by AG 5 and JVET AhG4 on 2023-09-20 13:00h UTC. The topic was the Verification Test preparation for VVC Multilayer Coding.

The AhG reviewed the status and progress with respect to the compression performance assessment of dual-layer coding category of the verification test. Potential issues with the current configuration and implementation of the VTM software were identified. Furthermore, the suitability of the current test sequences under consideration and the intended rate range were discussed. The provision of alternative test material, potentially including other resolutions, and further study was encouraged. The conduction of corresponding expert viewing tests at the meeting site in Hannover was solicited.

The report of the AhG meeting is available in JVET-AF0044.

* 1. ***Test sequences***

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

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).

1. **Related contributions**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| [JVET number](https://jvet-experts.org/doc_end_user/current_meeting.php?id_meeting=195&type_order=&sql_type=document_number) | MPEG number | [Created](https://jvet-experts.org/doc_end_user/current_meeting.php?id_meeting=195&type_order=&sql_type=document_date_time) | First upload | [Last upload](https://jvet-experts.org/doc_end_user/current_meeting.php?id_meeting=195&type_order=&sql_type=upload_document_date_time) | [Title](https://jvet-experts.org/doc_end_user/current_meeting.php?id_meeting=195&type_order=&sql_type=title) | [Source](https://jvet-experts.org/doc_end_user/current_meeting.php?id_meeting=195&type_order=&sql_type=authors) |
| [JVET-AF0004](https://jvet-experts.org/doc_end_user/current_document.php?id=13375) | m64938 | 2023-10-06 14:09:48 |  |  | JVET AHG report: Test material and visual assessment (AHG4) | V. Baroncini, T. Suzuki, M. Wien (co-chairs), W. Husak, S. Iwamura, P. de Lagrange, S. Liu, S. Puri, A. Segall, S. Wenger |
| [JVET-AF0044](https://jvet-experts.org/doc_end_user/current_document.php?id=13288) | m64729 | 2023-09-20 14:59:44 | 2023-09-20 16:26:38 | 2023-09-20 16:26:38 | AHG4: Report on AhG meeting on verification testing for VVC multilayer coding | [M. Wien (RWTH)](mailto:wien@lfb.rwth-aachen.de) |
| [JVET-AF0047](https://jvet-experts.org/doc_end_user/current_document.php?id=13294) | m64758 | 2023-09-28 12:30:50 | 2023-10-02 08:11:08 | 2023-10-02 08:11:08 | [AHG4] ECM10.0 evaluation on V3C test content | [S. Schwarz](mailto:sebastian.schwarz@nokia.com), M. M. Hannuksela (Nokia) |
| [JVET-AF0075](https://jvet-experts.org/doc_end_user/current_document.php?id=13323) | m64859 | 2023-10-06 04:56:24 | 2023-10-06 09:00:29 | 2023-10-11 06:44:02 | Evaluation results of Low Complexity Enhancement Video Codec (LCEVC) with HM and VTM on 4K content | O. Chubach, Y.-L. Hsiao, C.-Y. Chen, C.-W. Hsu, T.-D. Chuang, Y.-W. Chen, Y.-W. Huang, S.-M. Lei (MediaTek) |
| [JVET-AF0186](https://jvet-experts.org/doc_end_user/current_document.php?id=13444) | m65031 | 2023-10-06 23:27:22 | 2023-10-06 23:56:26 | 2023-10-06 23:56:26 | AHG4: experiments related to VVC spatial scalability visual testing | [P. de Lagrange](mailto:philippe.delagrange@interdigital.com), [F. Urban (InterDigital)](mailto:fabrice.urban@interdigital.com) |
| [JVET-AF0187](https://jvet-experts.org/doc_end_user/current_document.php?id=13445) | m65032 | 2023-10-06 23:28:48 | 2023-10-06 23:33:14 | 2023-10-06 23:33:14 | Compression of Gaming Contents | Z. Lin, [K. Cai](mailto:caikangying@huawei.com), J. Sauer, Y. Zhao, E. Alshina (Huawei) |
| [JVET-AF0253](https://jvet-experts.org/doc_end_user/current_document.php?id=13516) | m65538 | 2023-10-11 10:26:29 | 2023-10-11 17:38:45 | 2023-10-11 17:38:45 | [AHG-4] On the proposed gaming sequences from InterDigital and class gaming | [S. Puri](mailto:saurabh.puri@interdigital.com), T. Poirier, C. Bonnineau, I. Marzuki, R. Utida, E. Francois (InterDigital) |
| [JVET-AF0262](https://jvet-experts.org/doc_end_user/current_document.php?id=13526) | m65563 | 2023-10-12 01:10:15 | 2023-10-12 01:20:34 | 2023-10-12 01:20:34 | New Film Grain Material based on a Ground Truth approach | D. Ugur, D. Podborski, [A. M. Tourapis (Apple Inc)](mailto:atourapis@apple.com) |
| [JVET-AF0263](https://jvet-experts.org/doc_end_user/current_document.php?id=13527) | m65564 | 2023-10-12 01:12:29 | 2023-10-12 01:20:56 | 2023-10-12 01:20:56 | Suggested process for measuring Grain Fidelity using the Ground Truth test set | [A. M. Tourapis](mailto:atourapis@apple.com), J. Kim, S. Paluri, D. Podborski (Apple Inc) |

1. **Recommendations**

The AHG recommends:

* To review and consider JVET-AF0186 in the finalization of the multi-layer VVC verification test plan.
* To consider the structure of a verification test report for multi-layer VVC taking into account the results achieved at the Geneva meeting in the content layering category of the verification test plan.
* To review and consider JVET-AE0262 and JVET-AE0263 in the development of the draft plan for subjective quality testing of FGC SEI message.
* To review document JVET-AF0075 in a joint meeting with WG 4 and AG 5.
* 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.

A joint meeting with WG 4 and AG 5 was agreed to be to be on JVET-AF0075.

[JVET-AF0005](https://jvet-experts.org/doc_end_user/current_document.php?id=13376) JVET AHG report: Conformance testing (AHG5) [I. Moccagatta (chair), F. Bossen, K. Kawamura, P. de Lagrange, T. Ikai, S. Iwamura, H.-J. Jhu, S. Paluri, K. Sühring, Y. Yu (vice chairs)]

1. **Activities**

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 31st and 32nd meetings.

1. **Timeline**

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

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

1. **Status on bitstream submission**

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

* conformance bitstreams for VVC:
  + 104 bitstream categories have been identified
  + At least one bitstream has been submitted in each identified category
  + 283 total bitstreams have been provided, checked, and made available
  + No changes in bitstream between 31st and 32nd meeting.
* conformance bitstreams for VVC operation range extensions:
  + 57 bitstream categories have been identified
  + 1 bitstream of 1 identified category has been re-generated
  + 128 bitstreams of 57 identified categories have been cross-checked and uploaded.
  + No changes between 31st and 32nd meeting.

1. **Activities and Discussion**

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

VVC activities:

In the 2nd edition company names and emails have been removed from the packages, 2nd edition packages are available at <https://www.itu.int/wftp3/av-arch/jvet-site/bitstream_exchange/VVC2ndEd/>

VVC operation range extensions activities:

In the 2nd edition company names and emails have been removed from the packages, 2nd edition packages are available at <https://www.itu.int/wftp3/av-arch/jvet-site/bitstream_exchange/VVC2ndEd/>

VVC Multilayer activities:

Six of the 7 additional conformance bitstreams for VVC multilayer configurations in JVET-AE2028 decoded correctly using VTM-20.2 are available in <https://www.itu.int/wftp3/av-arch/jvet-site/bitstream_exchange/VVCMultilayer/under_test/VTM-20.2/>.

Volunteers to cross check these 6 conformance bitstreams for VVC multilayer configurations have been identified and cross-check is ongoing.

According to the proponents, the VVC multilayer configurations bitstream in JVET-AE0111 can be decoded by VTM-20.2 and a volunteer to cross-check the bitstream has been identified. This bitstream has been added to JVET-AE2028 and will be uploaded to <https://www.itu.int/wftp3/av-arch/jvet-site/bitstream_exchange/VVCMultilayer/under_test/VTM-20.2/> shortly.

In total currently 7 additional conformance bitstreams for VVC multilayer configurations have been generated.

HEVC Multilayer supporting extended bit depth activities:

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

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

Cross-checking of the 4 HEVC Multilayer supporting extended bit depth bitstreams in AD0232 is contingent to the merging of the above MR. Review of the MR is on-going.

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

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

1. **Contributions**

None.

1. **Ftp site information**

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

* VVC:

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

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

* VVC operation range extensions:

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

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

The ftp site for uploading bitstream file is as follows.

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

(user id: avguest, passwd: Avguest201007)

If using FileZilla, the following configuration is suggested:

Graphical user interface, text, application, email

Description automatically generated

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

1. **Recommendations**

The AHG recommends the following:

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

It was requested to clarify whether additional work can be expected for defining more bitstreams in the category of VVC multi-layer, or if at some point the verified bitstreams could be submitted as an extension of VVC conformance.

[JVET-AF0006](https://jvet-experts.org/doc_end_user/current_document.php?id=13295) JVET AHG report: ECM software development (AHG6) [V. Seregin (chair), J. Chen, R. Chernyak, F. Le Léannec, K. Zhang (vice-chairs)]

(zoom connection broke down Fri 13 1100)

1. **Software development**

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

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

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

The following adopted aspects were integrated into ECM-10.0:

JVET-AE0091: High-Accuracy template matching (Test 3.3e) (MR 449)

JVET-AE0174: Add non-inter TM SPS flag for tool control (MR 452)

JVET-AE0139(Test 5.2c): Improved ALF fixed filters (MR 450)

JVET-AE0059(Test3.1b): Convolutional cross-component model for inter prediction (MR 453)

JVET-AE0169 (Test 2.2c/2.3b): Bi-predictive IBC (MR 448)

JVET-AE0100 (Test 2.1b): Block vector guided CCCM (MR 455)

JVET-AE0125 (Test 4.1) Shifting Quantization Center (MR 461)

JVET-AE0150: RPR filters and thresholds for scale factors 1.1x to 1.35x (MR 468)

JVET-AE0077(Test-2.9): Extended search areas for IntraTMP mode (MR 458)

JVET-AE0094(Test 2.6c): IBC with non-adjacent spatial candidates (MR 457)

JVET-AE0086(Test-4.2): Large NSPT (MR 466)

JVET-AE0091: Iterative BDOF Part (Test 3.5) (MR 460)

JVET-AE0084(Test-2.11a): Harmonization of IBC HMVP and IBC-LIC (MR 459)

[JVET-AE0046(Test 3.2): Bi-predictive GPM](https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/merge_requests/465#top) (MR 465)

[JVET-AE0159(Test 2.5c): Filtered Intra Block Copy](https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/merge_requests/462) (MR 462)

JVET-AE0169: IBC MBVD list construction (Test 2.4a) (MR 469)

JVET-AE0097: Remove offset update process in CCP merge (MR 463)

JVET-AE0043 (Test 2.7): Cross-component merge mode with temporal candidates (MR 467)

JVET-AE0102: LFNST/NSPT coefficients context (Test 4.3) (MR 478)

JVET-AE0151 (Test5.1b): CCSAO with history offsets and extended edge classifiers (MR 472)

JVET-AE0278: ALF, CCALF memory reduction (lossless change) (MR 480)

JVET-AE0078: IBC-LIC extension (Test 2.10a) (MR 482)

JVET-AE0136: Fix on TM-based reordering for affine MMVD (MR 456)

JVET-AE0162: fix MS-SSIM calculation for RPR/SR (MR 475)

JVET-AE0055: Add NumSignPredOverrideByQP to LD configurations (MR 454)

Bug fixes:

Fix: memory release for JVET\_AB0155\_SGPM (MR 452)

Fixes for DebugStream tool (MRs 488, 489, 490, 502)

Fix: buffer copy for AC0119 LM chroma fusion (MR 484)

Fix condition for checking buffer overflow (MR 494)

Fix: add initialization for BIF overhead parsing and copy (MR 504)

Fix: Add checks for scaled collocated picture, inter check for GPM (MR 508)

Fix: Valgrind error in spanGeoMMVDMotionInfo (MR 514)

Fix: use picture level reconstructed buffer for SBT sub-TUs in BIF (MR 512)

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

JVET-AE0057: MTT split modes early termination (MR 474)

JVET-AE0162: fix MS-SSIM calculation for RPR/SR (MR 476)

ECM-10.0 and VTM-11ecm10.0 were tagged on August 29, 2023.

* 1. ***CTC Performance***

In this section, ECM test results following ECM CTC configuration descried in JVET-AE2017 are summarized.

Next tables show ECM-10.0 performance over ECM-9.1 anchor.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | | |
|  | **Over ECM-9.1** | | | | | |
|  | Y | U | V | EncT | DecT | VmPeak |
| Class A1 | -0.93% | -0.63% | -1.17% | 108.3% | 102.2% | 74.2% |
| Class A2 | -1.02% | -1.59% | -1.37% | 106.5% | 100.6% | 74.2% |
| Class B | -1.06% | -1.10% | -1.31% | 106.0% | 102.1% | 70.5% |
| Class C | -0.89% | -0.66% | -0.80% | 110.1% | 103.5% | 87.7% |
| Class E | -1.06% | -1.05% | -0.98% | 105.8% | 102.1% | 79.9% |
| **Overall** | -1.00% | -1.00% | -1.13% | 107.3% | 102.2% | 76.9% |
| Class D | -0.62% | -0.79% | -0.72% | 108.0% | 99.2% | 95.1% |
| Class F | -2.56% | -2.81% | -3.05% | 106.7% | 103.7% | 80.1% |
| Class TGM | -3.21% | -3.53% | -3.50% | 104.9% | 108.7% | 71.4% |
|  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | |
|  | **Over ECM-9.1** | | | | | |
|  | Y | U | V | EncT | DecT | VmPeak |
| Class A1 | -1.74% | -3.59% | -6.37% | 108.4% | 97.7% | 79.4% |
| Class A2 | -2.07% | -4.14% | -5.39% | 106.1% | 96.0% | 79.6% |
| Class B | -1.24% | -3.84% | -3.50% | 103.1% | 95.7% | 72.3% |
| Class C | -1.25% | -2.08% | -1.93% | 105.3% | 99.5% | 85.5% |
| Class E |  |  |  |  |  |  |
| **Overall** | -1.51% | -3.38% | -4.04% | 105.3% | 97.2% | 78.5% |
| Class D | -1.27% | -2.12% | -2.15% | 105.8% | 100.2% | 92.8% |
| Class F | -2.13% | -2.98% | -3.03% | 107.9% | 97.6% | 80.2% |
| Class TGM | -2.23% | -2.82% | -2.95% | 102.1% | 90.8% | 76.0% |
|  |  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | | |
|  | **Over ECM-9.1** | | | | | |
|  | Y | U | V | EncT | DecT | VmPeak |
| Class A1 |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |
| Class B | -0.69% | -8.86% | -7.04% | 94.0% | 98.5% | 69.8% |
| Class C | -0.97% | -3.14% | -3.09% | 106.7% | 101.6% | 86.2% |
| Class E | -0.98% | -3.47% | -3.09% | 105.8% | 98.2% | 78.5% |
| **Overall** | -0.86% | -5.61% | -4.74% | 101.0% | 99.4% | 77.1% |
| Class D | -0.87% | -4.16% | -3.75% | 105.0% | 99.7% | 94.0% |
| Class F | -1.59% | -6.03% | -7.65% | 106.9% | 95.9% | 79.7% |
| Class TGM | -1.98% | -4.41% | -5.46% | 103.8% | 94.5% | 74.6% |
|  |  |  |  |  |  |  |
|  | **Low delay P Main 10** | | | | | |
|  | **Over ECM-9.1** | | | | | |
|  | Y | U | V | EncT | DecT | VmPeak |
| Class A1 |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |
| Class B | -0.42% | -8.57% | -7.36% | 93.1% | 95.5% | 69.6% |
| Class C | -0.84% | -3.48% | -3.68% | 105.2% | 94.6% | 86.3% |
| Class E | -0.59% | -2.68% | -2.66% | 107.0% | 98.0% | 78.5% |
| **Overall** | -0.60% | -5.40% | -4.96% | 100.4% | 95.8% | 77.1% |
| Class D | -1.01% | -2.98% | -4.80% | 105.1% | 95.4% | 94.2% |
| Class F | -2.00% | -6.37% | -6.84% | 109.5% | 96.1% | 79.5% |
| Class TGM | -2.71% | -5.41% | -6.28% | 105.5% | 98.0% | 73.5% |

The below tables show ECM-10.0 performance comparing to VTM-11.0ecm10.0 anchor.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | | |
|  | **Over VTM-11.0ecm10** | | | | | |
|  | Y | U | V | EncT | DecT | VmPeak |
| Class A1 | -10.55% | -21.81% | -29.24% | 915.8% | 439.3% | #DIV/0! |
| Class A2 | -16.68% | -28.60% | -31.54% | 916.2% | 458.0% | #DIV/0! |
| Class B | -11.02% | -26.58% | -24.82% | 826.4% | 455.0% | #DIV/0! |
| Class C | -11.07% | -16.52% | -17.06% | 898.0% | 452.7% | #DIV/0! |
| Class E | -14.59% | -24.40% | -22.89% | 804.1% | 467.8% | #DIV/0! |
| **Overall** | -12.49% | -23.52% | -24.63% | 867.2% | 454.4% | #DIV/0! |
| Class D | -9.17% | -14.38% | -14.35% | 851.0% | 445.1% | #DIV/0! |
| Class F | -26.21% | -35.26% | -35.17% | 541.1% | 452.7% | #DIV/0! |
| Class TGM | -39.28% | -47.12% | -46.45% | 457.3% | 496.2% | #DIV/0! |
|  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | |
|  | **Over VTM-11.0ecm10** | | | | | |
|  | Y | U | V | EncT | DecT | VmPeak |
| Class A1 | -22.25% | -27.01% | -36.07% | 882.3% | 784.9% | #DIV/0! |
| Class A2 | -25.50% | -35.13% | -38.75% | 791.3% | 889.0% | #DIV/0! |
| Class B | -20.50% | -34.42% | -32.13% | 673.3% | 731.6% | #DIV/0! |
| Class C | -21.91% | -26.29% | -26.07% | 708.2% | 777.7% | #DIV/0! |
| Class E |  |  |  |  |  |  |
| **Overall** | -22.23% | -30.91% | -32.63% | 744.0% | 784.1% | #DIV/0! |
| Class D | -22.87% | -27.80% | -28.10% | 738.6% | 851.2% | #DIV/0! |
| Class F | -28.30% | -36.71% | -37.10% | 602.2% | 482.5% | #DIV/0! |
| Class TGM | -36.63% | -43.02% | -43.08% | 570.5% | 415.0% | #DIV/0! |
|  |  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | | |
|  | **Over VTM-11.0ecm10** | | | | | |
|  | Y | U | V | EncT | DecT | VmPeak |
| Class A1 |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |
| Class B | -17.55% | -38.06% | -35.34% | 589.9% | 668.2% | #DIV/0! |
| Class C | -19.29% | -28.96% | -29.31% | 648.1% | 710.0% | #DIV/0! |
| Class E | -16.12% | -28.03% | -27.57% | 642.1% | 459.1% | #DIV/0! |
| **Overall** | -17.77% | -32.52% | -31.39% | 621.7% | 620.8% | #DIV/0! |
| Class D | -21.17% | -30.48% | -30.27% | 656.7% | 758.7% | #DIV/0! |
| Class F | -24.60% | -37.20% | -36.64% | 588.5% | 498.3% | #DIV/0! |
| Class TGM | -35.27% | -43.12% | -43.29% | 560.3% | 418.2% | #DIV/0! |
|  |  |  |  |  |  |  |
|  | **Low delay P Main 10** | | | | | |
|  | **Over VTM-11.0ecm10** | | | | | |
|  | Y | U | V | EncT | DecT | VmPeak |
| Class A1 |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |
| Class B | -16.23% | -37.00% | -34.66% | 529.3% | 651.6% | #DIV/0! |
| Class C | -18.30% | -28.47% | -28.23% | 561.3% | 660.2% | #DIV/0! |
| Class E | -15.42% | -26.79% | -27.20% | 626.8% | 468.8% | #DIV/0! |
| **Overall** | -16.72% | -31.60% | -30.65% | 563.0% | 602.8% | #DIV/0! |
| Class D | -21.10% | -29.72% | -30.38% | 592.5% | 707.9% | #DIV/0! |
| Class F | -23.63% | -37.06% | -36.33% | 602.7% | 467.7% | #DIV/0! |
| Class TGM | -32.74% | -41.78% | -41.94% | 624.2% | 417.1% | #DIV/0! |

1. **ECM memory consumption**

ECM memory consumption (VmPeak, GiB) is provided in ECM encoder log files and is summarized in the table below.

|  |  |  |  |
| --- | --- | --- | --- |
|  | AI | RA | LB |
| class A (A1& A2) | 9 | 20 |  |
| class B | 7 | 10 | 8 |
| class C | 5 | 6 | 5 |
| class D | 5 | 5 | 5 |
| class E | 6 |  | 6 |
| class F | 8 | 10 | 8 |
| class TGM | 8 | 12 | 11 |

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

Memory consumption was further reduced after ECM-10.0 release, for the current master branch it is reduced by ~10% to ~60% depending on configuration and resolution. It is planned to tag it as ECM-10.1.

1. **Contributions**

[JVET-AF0090](https://jvet-experts.org/doc_end_user/current_document.php?id=13337), “AHG6: Encoder memory profiling on ECM software“, Y. Yasugi, T. Ikai (Sharp), R. Chernyak, S. Liu (Tencent)

[JVET-AF0156](https://jvet-experts.org/doc_end_user/current_document.php?id=13414), “AHG6: ECM software optimizations”, F. Urban, T. Poirier, F. Le Léannec (InterDigital)

[JVET-AF0177](https://jvet-experts.org/doc_end_user/current_document.php?id=13435), “AHG6: Memory reduction for ECM encoder”, N. Hu, V. Seregin, M. Karczewicz (Qualcomm), Y. Yasugi, T. Ikai (Sharp)

[JVET-AF0201](https://jvet-experts.org/doc_end_user/current_document.php?id=13462), “AHG6: On ECM SW memory consumption”, R. Chernyak, M. Xu, S. Liu (Tencent), Y. Yasugi, T. Ikai (Sharp)

1. **Recommendations**

The AHG recommends to:

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

One expert commented that as a general practice in ECM development, the person issuing a merge request should not also do the merge. In case a SW coordinator does a merge request, another SW coordinator should do the merge.

[JVET-AF0007](https://jvet-experts.org/doc_end_user/current_document.php?id=13377) 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)]

(zoom connection still down Fri 13 1110)

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

The same four groups were used in this meeting cycle.

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

Five group-off tests were performed and crosschecked on top of ECM-10 and later versions with bug fixes. The two anchors are ECM 10.0 and VTM-11ECM10.0. 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. Note that in all the crosschecks, the special setting for LD class B QP22 (with encoder configuration parameter MaxMTTHierarchyDepthByTidOverrideByQP) was not set so that the LDB class B results are slightly different from the results in tests. As the issue was spotted too late and LDB class B QP22 are the slowest points, reruns could not be finished before the JVET-AF meeting.

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

The software versions used in the tests are summarized below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Tests | Tester | | Crosschecker | |
| Anchor | Test | Anchor | Test |
| Group 1 off | ECM-10 | ECM-10 | ECM-10 | ECM-10 |
| Group 2 off | ECM-10 (41ef8a15) | ECM-10 (41ef8a15) | ECM-10 | ECM-10 (41ef8a15) |
| Group 3 off | ECM-10 | ECM-10 | ECM-10 | ECM-10 (41ef8a15) |
| Group 4 off | ECM-10 | ECM-10 | ECM-10 | ECM-10 (41ef8a15) |
| Group 1-4 off | ECM-10 | ECM-10 (b0b4f1e9b) | ECM-10 | ECM-10 (41ef8a15) |

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

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

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | | | | |
|  | **Over ECM-10.0** | | | | | | **Over VTM-11.0ecm10.0** | | | | |
|  | Y | U | V | EncT | DecT | VmPeak | Y | U | V | EncT | DecT |
| Class A1 | 0.00% | 0.00% | 0.00% | 97.8% | 98.9% | 100.0% | -10.55% | -21.81% | -29.24% | 799.4% | 482.6% |
| Class A2 | 0.00% | 0.00% | 0.00% | 95.9% | 94.5% | 99.9% | -16.68% | -28.60% | -31.54% | 785.6% | 473.1% |
| Class B | 0.00% | 0.00% | 0.00% | 98.2% | 97.0% | 100.0% | -11.03% | -26.58% | -24.82% | 736.7% | 471.3% |
| Class C | -0.01% | -0.01% | -0.01% | 103.6% | 109.4% | 100.0% | -11.08% | -16.53% | -17.07% | 746.9% | 424.0% |
| Class E | -0.01% | -0.01% | -0.01% | 105.3% | 110.6% | 100.0% | -14.60% | -24.41% | -22.90% | 736.8% | 509.9% |
| **Overall** | -0.01% | -0.01% | -0.01% | 100.1% | 101.7% | 100.0% | -12.50% | -23.53% | -24.63% | 757.2% | 468.6% |
| Class D | -0.03% | -0.02% | -0.02% | 99.2% | 102.0% | 100.0% | -9.20% | -14.40% | -14.37% | 736.8% | 412.4% |
| Class F | -0.01% | -0.01% | -0.01% | 98.0% | 102.8% | 100.0% | -26.22% | -35.27% | -35.17% | 501.7% | 496.4% |
| Class TGM | 0.00% | 0.00% | 0.00% | 96.8% | 100.0% | 100.0% | -39.29% | -47.12% | -46.45% | 419.3% | 538.1% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | | | | | | |
|  | **Over ECM-10.0** | | | | | | **Over VTM-11.0ecm10.0** | | | | |
|  | Y | U | V | EncT | DecT | VmPeak | Y | U | V | EncT | DecT |
| Class A1 | 3.92% | 3.93% | 3.99% | 83.6% | 80.9% | 100.1% | -19.20% | -24.26% | -33.44% | 666.7% | 758.6% |
| Class A2 | 4.36% | 4.59% | 4.61% | 83.8% | 77.6% | 100.0% | -22.26% | -32.18% | -35.93% | 598.3% | 836.7% |
| Class B | 3.71% | 3.94% | 3.69% | 81.3% | 74.5% | 99.9% | -17.52% | -31.80% | -29.60% | 507.3% | 678.0% |
| Class C | 4.06% | 4.24% | 4.64% | 81.0% | 78.8% | 100.0% | -18.71% | -23.13% | -22.58% | 508.5% | 602.3% |
| Class E |  |  |  |  |  |  |  |  |  |  |  |
| **Overall (Ref)** | 3.98% | 4.15% | 4.18% | 82.2% | 77.5% | 100.0% | -19.12% | -28.05% | -29.76% | 554.1% | 700.7% |
| Class D | 3.31% | 3.71% | 3.47% | 78.2% | 71.4% | 100.0% | -20.26% | -25.06% | -25.54% | 540.0% | 617.7% |
| Class F | 2.94% | 2.92% | 3.11% | 89.3% | 84.3% | 99.9% | -26.11% | -34.80% | -35.07% | 506.3% | 421.7% |
| Class TGM | 3.79% | 3.79% | 3.74% | 88.3% | 86.1% | 100.0% | -34.24% | -40.83% | -40.93% | 504.3% | 467.3% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | | | | | | | |
|  | **Over ECM-10.0** | | | | | | **Over VTM-11.0ecm10.0** | | | | |
|  | Y | U | V | EncT | DecT | VmPeak | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |  |  |  |  |
| Class B | 4.59% | 5.35% | 4.80% | 79.3% | 70.6% | 100.0% | -13.75% | -34.87% | -32.35% | 405.0% | 535.0% |
| Class C | 4.70% | 4.76% | 5.20% | 72.1% | 58.4% | 100.0% | -15.47% | -25.61% | -25.62% | 407.9% | 439.6% |
| Class E | 4.50% | 5.53% | 5.32% | 79.7% | 75.8% | 100.0% | -12.34% | -24.14% | -23.82% | 467.4% | 398.5% |
| **Overall (Ref)** | 4.60% | 5.20% | 5.06% | 76.9% | 67.5% | 100.0% | -13.97% | -29.10% | -27.98% | 420.8% | 465.5% |
| Class D | 4.13% | 4.64% | 4.25% | 72.8% | 60.7% | 100.0% | -17.89% | -27.18% | -27.32% | 468.6% | 513.5% |
| Class F | 4.24% | 3.47% | 4.11% | 82.1% | 72.8% | 100.0% | -21.33% | -34.99% | -34.06% | 470.0% | 412.4% |
| Class TGM | 5.78% | 5.60% | 5.50% | 91.3% | 85.5% | 100.6% | -31.51% | -39.83% | -40.05% | 469.8% | 374.6% |

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

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

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | | | | |
|  | **Over ECM-10 (41ef8a15)** | | | | | | **Over VTM-11ecm10** | | | | |
|  | Y | U | V | EncT | DecT | VmPeak | Y | U | V | EncT | DecT |
| Class A1 | 0.39% | 0.80% | 0.64% | 91.5% | 96.5% | 99.8% | -10.19% | -21.21% | -28.74% | 788.0% | 445.3% |
| Class A2 | 0.55% | 0.75% | 0.77% | 89.0% | 95.8% | 99.9% | -16.22% | -28.07% | -31.02% | 764.8% | 443.9% |
| Class B | 0.48% | 0.72% | 0.70% | 87.7% | 94.8% | 99.9% | -10.59% | -26.02% | -24.27% | 692.3% | 442.3% |
| Class C | 0.34% | 0.51% | 0.68% | 85.3% | 91.8% | 99.9% | -10.76% | -16.11% | -16.49% | 700.8% | 397.2% |
| Class E | 0.72% | 0.98% | 0.93% | 87.9% | 94.5% | 99.9% | -13.96% | -23.65% | -22.15% | 660.2% | 456.0% |
| **Overall** | 0.49% | 0.74% | 0.73% | 88.0% | 94.5% | 99.9% | -12.06% | -22.96% | -24.06% | 715.5% | 434.8% |
| Class D | 0.31% | 0.42% | 0.49% | 85.4% | 89.2% | 99.9% | -8.89% | -14.01% | -13.92% | 713.2% | 407.2% |
| Class F | 3.58% | 3.71% | 3.97% | 89.5% | 95.7% | 99.9% | -23.78% | -33.16% | -32.92% | 486.0% | 464.9% |
| Class TGM | 0.87% | 0.91% | 0.94% | 89.1% | 100.3% | 99.9% | -38.74% | -46.64% | -45.95% | 416.6% | 535.0% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | | | | | | |
|  | **Over ECM-10 (41ef8a15)** | | | | | | **Over VTM-11ecm10** | | | | |
|  | Y | U | V | EncT | DecT | VmPeak | Y | U | V | EncT | DecT |
| Class A1 | 2.37% | 0.74% | 1.65% | 91.2% | 98.4% | 99.9% | -20.39% | -26.36% | -34.98% | 776.8% | 861.1% |
| Class A2 | 1.68% | 0.94% | 1.39% | 90.1% | 99.7% | 100.2% | -24.23% | -34.48% | -37.87% | 669.6% | 996.2% |
| Class B | 1.65% | 1.11% | 1.29% | 89.6% | 96.9% | 99.7% | -19.18% | -33.64% | -31.27% | 607.6% | 832.9% |
| Class C | 1.25% | 0.97% | 0.53% | 87.7% | 95.2% | 99.9% | -20.95% | -25.63% | -25.72% | 627.8% | 872.8% |
| Class E |  |  |  |  |  |  |  |  |  |  |  |
| **Overall (Ref)** | 1.70% | 0.96% | 1.18% | 89.5% | 97.3% | 99.9% | -20.90% | -30.21% | -31.85% | 656.4% | 879.9% |
| Class D | 0.68% | -0.03% | 0.11% | 86.3% | 95.2% | 99.9% | -22.32% | -27.76% | -28.00% | 666.9% | 961.1% |
| Class F | 3.73% | 3.60% | 3.45% | 88.9% | 95.8% | 100.0% | -25.71% | -34.60% | -35.11% | 566.3% | 524.4% |
| Class TGM | 1.92% | 1.93% | 2.02% | 90.0% | 100.3% | 100.0% | -35.38% | -41.89% | -41.89% | 549.7% | 508.2% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | | | | | | | |
|  | **Over ECM-10 (41ef8a15)** | | | | | | **Over VTM-11ecm10** | | | | |
|  | Y | U | V | EncT | DecT | VmPeak | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |  |  |  |  |
| Class B | 1.90% | 0.53% | 0.59% | 84.8% | 97.2% | 99.9% | -15.97% | -37.26% | -34.59% | 501.8% | 783.5% |
| Class C | 1.58% | 0.81% | 1.11% | 84.6% | 93.3% | 99.9% | -18.09% | -28.43% | -28.57% | 539.4% | 750.0% |
| Class E | 1.16% | 0.97% | 0.14% | 87.4% | 96.7% | 99.9% | -15.13% | -27.23% | -27.12% | 571.3% | 537.9% |
| **Overall (Ref)** | 1.61% | 0.73% | 0.65% | 85.4% | 95.8% | 99.9% | -16.46% | -31.81% | -30.72% | 531.0% | 702.9% |
| Class D | 0.77% | 0.41% | 0.60% | 84.0% | 93.9% | 99.9% | -20.52% | -30.29% | -30.09% | 580.6% | 933.4% |
| Class F | 3.38% | 1.85% | 1.52% | 88.6% | 96.3% | 99.8% | -22.07% | -35.99% | -35.72% | 544.4% | 557.4% |
| Class TGM | 2.01% | 2.04% | 2.34% | 90.5% | 102.1% | 100.2% | -33.94% | -42.10% | -42.06% | 513.2% | 496.1% |

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

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

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | | | | |
|  | **Over ECM-10** | | | | | | **Over VTM-11ecm10** | | | | |
|  | Y | U | V | EncT | DecT | VmPeak | Y | U | V | EncT | DecT |
| Class A1 | 1.00% | 2.85% | 4.29% | 80.0% | 88.8% | 99.9% | -9.65% | -19.75% | -26.36% | 715.7% | 360.7% |
| Class A2 | 1.84% | 4.41% | 4.43% | 80.8% | 79.3% | 99.9% | -15.12% | -25.67% | -28.71% | 712.3% | 329.9% |
| Class B | 1.86% | 3.90% | 3.87% | 78.5% | 74.8% | 100.0% | -9.39% | -23.94% | -22.06% | 647.7% | 317.1% |
| Class C | 1.66% | 2.10% | 2.25% | 78.5% | 74.1% | 100.0% | -9.62% | -14.84% | -15.26% | 662.7% | 290.4% |
| Class E | 2.94% | 3.78% | 3.52% | 78.6% | 75.1% | 100.0% | -12.08% | -21.63% | -20.20% | 625.2% | 320.3% |
| **Overall** | 1.85% | 3.39% | 3.62% | 79.1% | 77.6% | 100.0% | -10.89% | -21.12% | -22.06% | 668.6% | 320.3% |
| Class D | 1.22% | 1.81% | 1.79% | 80.2% | 73.4% | 100.0% | -8.06% | -12.83% | -12.83% | 677.0% | 298.7% |
| Class F | 4.36% | 5.57% | 5.66% | 78.7% | 69.8% | 100.0% | -23.11% | -31.86% | -31.72% | 464.5% | 302.0% |
| Class TGM | 8.63% | 10.60% | 10.93% | 82.1% | 59.9% | 99.9% | -34.14% | -41.63% | -40.72% | 414.4% | 287.3% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | | | | | | |
|  | **Over ECM-10** | | | | | | **Over VTM-11ecm10** | | | | |
|  | Y | U | V | EncT | DecT | VmPeak | Y | U | V | EncT | DecT |
| Class A1 | 0.96% | 2.00% | 5.80% | 95.1% | 99.0% | 100.2% | -21.50% | -25.59% | -32.98% | 813.1% | 702.3% |
| Class A2 | 1.10% | 3.01% | 3.27% | 95.7% | 100.1% | 99.6% | -24.68% | -33.27% | -36.81% | 739.3% | 828.6% |
| Class B | 1.00% | 3.61% | 3.47% | 94.5% | 98.4% | 99.9% | -19.70% | -32.07% | -29.80% | 645.0% | 684.2% |
| Class C | 0.80% | 1.16% | 1.19% | 94.4% | 99.0% | 100.0% | -21.28% | -25.40% | -25.17% | 663.9% | 707.8% |
| Class E |  |  |  |  |  |  |  |  |  |  |  |
| **Overall (Ref)** | 0.96% | 2.51% | 3.29% | 94.8% | 99.1% | 99.9% | -21.48% | -29.24% | -30.60% | 699.6% | 721.1% |
| Class D | 0.68% | 1.56% | 1.85% | 95.2% | 98.4% | 100.0% | -22.34% | -26.62% | -26.73% | 711.7% | 728.1% |
| Class F | 2.80% | 3.63% | 3.81% | 92.3% | 99.5% | 100.2% | -26.43% | -34.60% | -34.82% | 589.0% | 412.3% |
| Class TGM | 5.50% | 7.34% | 7.70% | 93.9% | 98.0% | 100.0% | -33.15% | -38.85% | -38.74% | 585.9% | 381.1% |

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

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

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

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | | | | |
|  | **Over ECM-10** | | | | | | **Over VTM-11ecm10** | | | | |
|  | Y | U | V | EncT | DecT | VmPeak | Y | U | V | EncT | DecT |
| Class A1 | 1.60% | 5.59% | 7.45% | 86.9% | 94.4% | 99.9% | -9.12% | -17.75% | -24.20% | 752.4% | 416.8% |
| Class A2 | 2.48% | 6.17% | 7.18% | 83.7% | 96.3% | 99.9% | -14.70% | -24.44% | -26.94% | 702.7% | 429.4% |
| Class B | 1.01% | 6.50% | 5.30% | 87.5% | 96.9% | 100.0% | -10.14% | -22.27% | -21.13% | 691.2% | 442.6% |
| Class C | 0.89% | 2.91% | 2.83% | 84.5% | 95.8% | 100.0% | -10.29% | -14.24% | -14.84% | 689.1% | 416.4% |
| Class E | 1.16% | 6.61% | 3.92% | 87.6% | 98.0% | 100.0% | -13.60% | -19.81% | -20.08% | 666.8% | 463.0% |
| **Overall** | 1.35% | 5.51% | 5.19% | 86.1% | 96.3% | 100.0% | -11.34% | -19.68% | -21.04% | 698.3% | 433.4% |
| Class D | 0.87% | 2.53% | 2.34% | 85.6% | 96.9% | 100.0% | -8.38% | -12.27% | -12.38% | 696.3% | 433.3% |
| Class F | 1.21% | 4.10% | 4.28% | 92.3% | 97.1% | 100.0% | -25.33% | -32.83% | -32.59% | 477.0% | 458.9% |
| Class TGM | 2.27% | 3.84% | 3.76% | 91.4% | 96.2% | 100.0% | -37.97% | -45.22% | -44.56% | 415.0% | 500.0% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | | | | | | |
|  | **Over ECM-10** | | | | | | **Over VTM-11ecm10** | | | | |
|  | Y | U | V | EncT | DecT | VmPeak | Y | U | V | EncT | DecT |
| Class A1 | 1.18% | 2.92% | 5.22% | 91.1% | 98.9% | 100.5% | -21.33% | -24.98% | -33.17% | 779.5% | 786.8% |
| Class A2 | 1.08% | 3.53% | 4.33% | 94.2% | 99.8% | 99.8% | -24.69% | -32.95% | -36.23% | 709.2% | 913.4% |
| Class B | 0.52% | 6.08% | 4.97% | 93.5% | 100.0% | 100.1% | -20.08% | -30.65% | -28.98% | 623.8% | 788.6% |
| Class C | 0.41% | 1.88% | 1.76% | 94.0% | 99.8% | 100.0% | -21.60% | -24.89% | -24.75% | 644.2% | 828.5% |
| Class E |  |  |  |  |  |  |  |  |  |  |  |
| **Overall (Ref)** | 0.73% | 3.82% | 4.04% | 93.3% | 99.7% | 100.1% | -21.66% | -28.44% | -30.14% | 674.9% | 822.5% |
| Class D | 0.37% | 1.81% | 2.11% | 95.0% | 100.2% | 100.0% | -22.58% | -26.52% | -26.61% | 691.3% | 889.2% |
| Class F | 0.94% | 3.23% | 3.45% | 95.8% | 98.4% | 99.9% | -27.69% | -34.90% | -35.15% | 594.3% | 493.7% |
| Class TGM | 0.68% | 1.56% | 1.34% | 94.6% | 99.3% | 100.0% | -36.23% | -42.17% | -42.36% | 576.2% | 455.0% |

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

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

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | | | | |
|  | **Over ECM-10** | | | | | | **Over VTM-11ecm10** | | | | |
|  | Y | U | V | EncT | DecT | VmPeak | Y | U | V | EncT | DecT |
| Class A1 | 3.24% | 9.59% | 13.34% | 61.0% | 76.1% | 99.7% | -7.63% | -14.77% | -20.19% | 546.8% | 316.3% |
| Class A2 | 5.22% | 12.08% | 12.57% | 57.6% | 68.7% | 99.7% | -12.36% | -20.41% | -23.37% | 507.5% | 297.0% |
| Class B | 3.67% | 11.44% | 10.20% | 56.4% | 63.4% | 99.9% | -7.78% | -18.82% | -17.57% | 455.0% | 281.3% |
| Class C | 3.23% | 5.74% | 5.93% | 54.4% | 56.8% | 99.9% | -8.23% | -11.93% | -12.29% | 453.5% | 250.0% |
| Class E | 5.56% | 11.56% | 8.76% | 57.6% | 64.5% | 99.9% | -9.84% | -16.07% | -16.33% | 461.4% | 294.5% |
| **Overall** | 4.07% | 9.99% | 9.93% | 57.1% | 64.8% | 99.8% | -8.96% | -16.42% | -17.59% | 478.5% | 284.1% |
| Class D | 2.76% | 4.96% | 4.94% | 55.1% | 54.7% | 99.9% | -6.66% | -10.17% | -10.14% | 465.3% | 242.1% |
| Class F | 10.53% | 15.32% | 15.80% | 64.0% | 61.5% | 99.9% | -18.81% | -26.29% | -25.77% | 367.0% | 277.9% |
| Class TGM | 14.28% | 18.07% | 18.81% | 64.7% | 56.0% | 99.9% | -30.83% | -37.82% | -36.67% | 315.8% | 267.9% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | | | | | | | |
|  | **Over ECM-10** | | | | | | **Over VTM-11ecm10** | | | | |
|  | Y | U | V | EncT | DecT | VmPeak | Y | U | V | EncT | DecT |
| Class A1 | 9.01% | 10.53% | 18.25% | 65.8% | 80.5% | 99.8% | -15.22% | -19.48% | -25.41% | 564.9% | 579.9% |
| Class A2 | 8.71% | 12.98% | 14.22% | 65.7% | 77.3% | 99.6% | -18.99% | -26.94% | -30.19% | 499.4% | 641.7% |
| Class B | 7.23% | 14.79% | 13.93% | 63.7% | 73.7% | 99.8% | -14.70% | -24.88% | -22.86% | 434.1% | 532.6% |
| Class C | 6.73% | 8.54% | 8.96% | 60.7% | 65.9% | 99.9% | -16.60% | -19.91% | -19.32% | 433.5% | 499.6% |
| Class E |  |  |  |  |  |  |  |  |  |  |  |
| **Overall (Ref)** | 7.75% | 11.91% | 13.53% | 63.7% | 73.5% | 99.8% | -16.17% | -22.89% | -23.89% | 470.4% | 552.8% |
| Class D | 5.26% | 7.62% | 8.28% | 61.7% | 64.5% | 99.9% | -18.72% | -22.17% | -21.99% | 462.6% | 530.5% |
| Class F | 11.45% | 14.38% | 14.99% | 69.5% | 76.9% | 99.8% | -20.36% | -28.13% | -28.20% | 443.8% | 344.8% |
| Class TGM | 13.48% | 16.18% | 16.41% | 72.2% | 77.4% | 99.9% | -28.09% | -33.82% | -33.78% | 453.6% | 318.9% |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | | | | | | | |
|  | **Over ECM-10** | | | | | | **Over VTM-11ecm10** | | | | |
|  | Y | U | V | EncT | DecT | VmPeak | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |  |  |  |  |
| Class B | 7.07% | 9.46% | 9.24% | 56.8% | 64.8% | 100.0% | -11.69% | -32.34% | -29.54% | 337.7% | 424.7% |
| Class C | 6.74% | 7.49% | 8.02% | 54.4% | 53.9% | 99.9% | -13.83% | -23.70% | -23.63% | 346.6% | 369.5% |
| Class E | 6.48% | 14.60% | 11.45% | 66.3% | 75.6% | 99.9% | -10.66% | -17.94% | -19.36% | 435.9% | 334.0% |
| **Overall (Ref)** | 6.81% | 10.09% | 9.39% | 58.2% | 63.4% | 99.9% | -12.14% | -25.86% | -25.03% | 363.1% | 381.8% |
| Class D | 5.19% | 6.89% | 7.08% | 57.3% | 54.4% | 99.9% | -17.01% | -25.62% | -25.22% | 390.5% | 429.3% |
| Class F | 10.40% | 11.81% | 11.19% | 66.3% | 71.0% | 99.9% | -16.84% | -29.98% | -29.78% | 412.1% | 322.5% |
| Class TGM | 13.46% | 14.95% | 15.60% | 72.1% | 72.4% | 99.4% | -26.57% | -34.63% | -34.56% | 425.6% | 276.2% |

* 1. ***Summary***

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

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

* Software issues #51, #54 with tool off tests was resolved
  1. ***Open issues***
* #53, decoding mismatch was observed when AMVR is off
* The variation of runtime over VTM11ECM9 is relatively large, for which the version of compiler matters
* Different software versions were used in AHG7 tests, which leads to slightly different results between testing and crosschecking.
* It was commented that it is not consistent to set LDB class B QP22 setting for ECM-10 anchor without also setting it for VTM11-ECM10 anchor. JVET-AF0165 is related to this issue.

1. **Input contributions**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| [JVET-AF0165](https://jvet-experts.org/doc_end_user/current_document.php?id=13423) | m64994 | 2023-10-06 18:23:08 | 2023-10-06 19:45:18 | 2023-10-06 19:45:18 | Ahg7/Ahg12: On VTM-11ecm anchor low-delay test conditions | [F. Le Léannec](mailto:fabrice.leleannec@interdigital.com), S. Puri, E. François, K. Naser (InterDigital) |

1. **Recommendations**

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

[JVET-AF0008](https://jvet-experts.org/doc_end_user/current_document.php?id=13379) JVET AHG report: Optimization of encoders and receiving systems for machine analysis of coded video content (AHG8) [C. Hollmann, S. Liu, S. Wang, M. Zhou (AHG chairs)]

1. **Activities**

The AHG used the main JVET reflector, jvet@lists.rwth-aachen.de, for email discussion. There was a kick-off email sent to the reflector with announcement of CTC document upload. Offline discussions were conducted among co-chairs, editors and interested experts for technical report preparation, common test conditions and anchor generation, etc. There are seven input contriubtions related to AHG 8 mandates submitted to this meeting. They are listed in Section 3.

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

Common test conditions (CTC) for optimization of encoders and receiving systems for machine analysis of coded video content, are summarized in output document JVET-AE2031. This document includes detailed descriptions of test datasets, anchor software and configurations, anchor generation processes, machine task networks used, test and training conditions, evaluation methodologies and metrics.

Following the decision made in the last meeting, the CTC includes two anchors, generated by VTM12 and VTM20 software, respectively. Hence, two reporting templates in Excel format are enclosed in the document package for reporting VTM12 and VTM20 based restuls, respectively. This output document package JVET-AE2031 was uploaded on 2023-08-11 to JVET document management system and is also available at <https://vcgit.hhi.fraunhofer.de/jvet-ahg-ofm>/ofm-ctc.

* 1. ***Technical Report***

The draft 3 of the technical report (TR) has been prepared and uploaded as JVET-AE2030 based on discussions and decisions in the last (31st) JVET meeting on 2023-09-14. Descriptions about

* Denoising filter (JVET-AE0081)
* NNPF and post-filter hint SEI message (JVET-AE0099)

were included in the TR, together with some text improvements.

* 1. ***Git Management***

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

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

1. **Input contributions**

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

|  |  |  |
| --- | --- | --- |
| **Report** | | |
| JVET-AF0008 | 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-AF0060 | AHG8: An exemplar software implementation for temporal resampling for VCM | [S. Wang](mailto:shurun.wsr@alibaba-inc.com), [J. Chen](mailto:jiechen.cj@alibaba-inc.com), [Y. Ye (Alibaba)](mailto:yan.ye@alibaba-inc.com), [S. Wang (CityU)](mailto:shiqwang@cityu.edu.hk) |
| JVET-AF0157 | AHG8: A temporal resampling algorithm | D. Ding, X. Zhao, Z. Liu, S. Liu (Tencent) |
| JVET-AF0068 | AHG8/AHG9: Signalling of encoder preprocessing information | [W. Jia](mailto:wei.jia@bytedance.com), [Y. Li](mailto:yue.li@bytedance.com), [Y.-K. Wang](mailto:yekui.wang@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com) |
| JVET-AF0138 | AHG8/AHG9: Truncated bit depth support SEI messages | [D. Ding](mailto:ddding@global.tencent.com), [X. Zhao](mailto:xinzzhao@global.tencent.com), [S. Liu](mailto:shanl@global.tencent.com), [G. Teniou](mailto:teniou@global.tencent.com), [S. Wenger (Tencent)](mailto:swenger@global.tencent.com) |
| JVET-AF0147 | AHG8/AHG9: On Picture Modality Type | [J. Gao](mailto:jingying.gao@sg.panasonic.com), [H.-B. Teo](mailto:hanboon.teo@sg.panasonic.com), C.-S. Lim, [K. Abe (Panasonic)](mailto:abe.kiyo@jp.panasonic.com) |
| **Crosscheck** | | |
| JVET-AE0241 | Cross-check of JVET-AF0157 (AHG8: A temporal resampling algorithm) | C. Hollmann (Ericsson) |

1. **Recommendations**

The AHG recommends to:

* Review all input contributions.
* Discuss and refine test conditions, evalution and reporting procedures.
* 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.
* Discuss TR development and finalization schedule.

Potential timelines for the TR were to be discussed. It was commented that the topic has such a broad range of applications that it might be impossible to cover all of them. A first version could very well be a subset of those. Newly proposed SEI messages could not be included.

[JVET-AF0009](https://jvet-experts.org/doc_end_user/current_document.php?id=13380) 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)]

1. **Related contributions**

A total of 42 contributions are identified relating to the mandates of AHG9. Some contributions also relate to the work of AHG8.

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

* 10 contributions relate to the mandate to study the SEI messages in VSEI, VVC, HEVC, and AVC.
  + 8 contributions relate to SEI processing order SEI message aspects other than post-processing filter (PPF) grouping
  + 2 contributions relate to SEI messages related to film grain synthesis
* 24 contributions relate to the mandate to discuss the document for the TuC for future extensions of VSEI.
  + 3 contributions relate to the general direction of grouping of PPFs, including NNPFs
  + 9 contributions relate to detailed aspects of NNPF grouping
  + 6 contributions relate to the source picture timing information SEI message
  + 4 contributions relate to the encoder optimization information SEI message
  + 2 contributions relate to the object mask information SEI message
* 2 contributions relate to the mandate to collect software and showcase information for SEI messages;
* 13 contributions relate to the mandate to identify potential needs for additional SEI messages.
  + 4 contributions relate to SEI messages related to generative video
  + 9 contributions relate to other SEI messages
* 1 contribution relates to the mandate to study the alignments of the same SEI messages in different standards.

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

* 1. ***Study*** ***the SEI messages in VSEI, VVC, HEVC and AVC***
     1. **AHG9: SEI processing order SEI message aspects other than PPF grouping (8)**

[JVET-AF0049](https://jvet-experts.org/doc_end_user/current_document.php?id=13297) AHG9: On the SEI processing order SEI message [M. M. Hannuksela (Nokia)]

[JVET-AF0061](https://jvet-experts.org/doc_end_user/current_document.php?id=13309) AHG9: On grouping of post-processing filters [Y.-K. Wang, W. Jia, J. Xu, L. Zhang (Bytedance)]

JVET-AF0061 also relates to the mandate to discuss the TuC for future extensions of VSEI.

[JVET-AF0062](https://jvet-experts.org/doc_end_user/current_document.php?id=13310) AHG9: Some syntax and semantics changes for the SEI processing order SEI message [Y.-K. Wang (Bytedance)]

[JVET-AF0065](https://jvet-experts.org/doc_end_user/current_document.php?id=13313) AHG9: On the SEI processing order SEI message (Part 1) [L. Chen, O. Chubach, Y.-W. Huang, S. Lei (MediaTek)]

[JVET-AF0067](https://jvet-experts.org/doc_end_user/current_document.php?id=13315) AHG9: On the SEI processing order SEI message (Part 2) [L. Chen, O. Chubach, Y.-W. Huang, S. Lei (MediaTek)]

[JVET-AF0070](https://jvet-experts.org/doc_end_user/current_document.php?id=13318) AHG9: On the SEI processing order SEI message [T. Chujoh, Y. Yasugi, T. Ikai (Sharp)]

[JVET-AF0174](https://jvet-experts.org/doc_end_user/current_document.php?id=13432) AHG9: On the SEI processing order SEI message and NNPF groups [Y. Sanchez, C. Hellge, T. Schierl (Fraunhofer HHI)]

JVET-AF0061 also relates to the mandate to discuss the TuC for future extensions of VSEI.

[JVET-AF0189](https://jvet-experts.org/doc_end_user/current_document.php?id=13449) AHG9: Proposed modifications of the draft SEI processing order SEI message in VVC [G. J. Sullivan, S. McCarthy, P. Yin (Dolby Labs)]

* + 1. **AHG9: SEI messages related to film grain synthesis (2)**

[JVET-AF0142](https://jvet-experts.org/doc_end_user/current_document.php?id=13400) [AHG9] Local Film Grain Synthesis using Annotated Regions SEI [G. Teniou, S. Wenger (Tencent)]

JVET-AF0142 also relates to the mandate to identify potential needs for additional SEI messages.

[JVET-AF0144](https://jvet-experts.org/doc_end_user/current_document.php?id=13402) [AHG9] Adaptive Film grain synthesis using Alpha Channel Information SEI message [G. Teniou, S. Wenger (Tencent)]

JVET-AF0144 also relates to the mandate to identify potential needs for additional SEI messages.

* 1. ***Discuss the document for the TuC for future extensions of VSEI***
     1. **AHG9: General direction of grouping of post-processing filters (PPFs), including NNPFs (3)**

[JVET-AF0050](https://jvet-experts.org/doc_end_user/current_document.php?id=13298) AHG9: On NNPF groups [M. M. Hannuksela, F. Cricri (Nokia)]

[JVET-AF0061](https://jvet-experts.org/doc_end_user/current_document.php?id=13309) AHG9: On grouping of post-processing filters [Y.-K. Wang, W. Jia, J. Xu, L. Zhang (Bytedance)]

JVET-AF0061 also relates to the mandate to study SEI messages in VSEI, VVC, HEVC and AVC.

[JVET-AF0174](https://jvet-experts.org/doc_end_user/current_document.php?id=13432) AHG9: On the SEI processing order SEI message and NNPF groups [Y. Sanchez, C. Hellge, T. Schierl (Fraunhofer HHI)]

JVET-AF0174 also relates to the mandate to study SEI messages in VSEI, VVC, HEVC and AVC.

* + 1. **AHG9: Detailed aspects of NNPF grouping (9)**

[JVET-AF0050](https://jvet-experts.org/doc_end_user/current_document.php?id=13298) AHG9: On NNPF groups [M. M. Hannuksela, F. Cricri (Nokia)]

[JVET-AF0051](https://jvet-experts.org/doc_end_user/current_document.php?id=13299) AHG9: Signalling the gain provided by NNPFs and NNPF groups [M. M. Hannuksela, F. Cricri, A. Hallapuro (Nokia)]

[JVET-AF0067](https://jvet-experts.org/doc_end_user/current_document.php?id=13315) AHG9: On the SEI processing order SEI message (Part 2) [L. Chen, O. Chubach, Y.-W. Huang, S. Lei (MediaTek)]

JVET-AF0067 also relates to the mandate to study SEI messages in VSEI, VVC, HEVC and AVC.

[JVET-AF0091](https://jvet-experts.org/doc_end_user/current_document.php?id=13338) [AHG9] On design for signalling purpose information for NNPFGC [[Hendry](mailto:dr.hendry@lge.com), J. Nam, S. Kim, J. Lim (LGE)]

[JVET-AF0092](https://jvet-experts.org/doc_end_user/current_document.php?id=13339) [AHG9] On the output pictures from NNPFGC with parallel grouping type [[Hendry](mailto:dr.hendry@lge.com), J. Nam, S. Kim, J. Lim (LGE)]

[JVET-AF0093](https://jvet-experts.org/doc_end_user/current_document.php?id=13340) [AHG9] On intermediary output picture(s) from activation of an NNPFGC [[Hendry](mailto:dr.hendry@lge.com), J. Nam, S. Kim, J. Lim (LGE)]

[JVET-AF0094](https://jvet-experts.org/doc_end_user/current_document.php?id=13341) [AHG9] On activation of an NNPFGC that contains another NNPFGC [[Hendry](mailto:dr.hendry@lge.com), J. Nam, S. Kim, J. Lim (LGE)]

[JVET-AF0095](https://jvet-experts.org/doc_end_user/current_document.php?id=13342) [AHG9] On order of output pictures when skipped candidate input pictures are present in NNPFGA [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

[JVET-AF0096](https://jvet-experts.org/doc_end_user/current_document.php?id=13343) [AHG9] On miscellaneous aspects of NNPFGC and NNPFGA [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

* + 1. **AHG9:** **Source picture timing information SEI message aspects (6)**

[JVET-AF0055](https://jvet-experts.org/doc_end_user/current_document.php?id=13303) AHG9: On Source Picture Timing Information (SPTI) SEI message [[J. Samuelsson-Allendes](mailto:samuelssonj@sharplabs.com), S. Deshpande (Sharp)]

[JVET-AF0069](https://jvet-experts.org/doc_end_user/current_document.php?id=13317) AHG9: On source picture timing information SEI message [S. McCarthy, G. J. Sullivan, P. Yin (Dolby)]

[JVET-AF0097](https://jvet-experts.org/doc_end_user/current_document.php?id=13344) [AHG9] On description of picture 0 in source picture timing information SEI message [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

[JVET-AF0098](https://jvet-experts.org/doc_end_user/current_document.php?id=13345) [AHG9] On temporal reversal feature in source picture timing information SEI message [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

[JVET-AF0099](https://jvet-experts.org/doc_end_user/current_document.php?id=13346) [AHG9] On signaling of source picture interval scale factor and sublayer synthesized flag in source picture timing information SEI message [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

[JVET-AF0100](https://jvet-experts.org/doc_end_user/current_document.php?id=13347) [AHG9] On miscellaneous aspects of source picture timing information SEI message [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

* + 1. **AHG9: Encoder optimization information SEI message aspects (4)**

[JVET-AF0052](https://jvet-experts.org/doc_end_user/current_document.php?id=13300)AHG9: Proposed updates on the encoder optimization information SEI message **[**M. M. Hannuksela, F. Cricri (Nokia)]

[JVET-AF0068](https://jvet-experts.org/doc_end_user/current_document.php?id=13316) AHG8/AHG9: Signalling of encoder preprocessing information [W. Jia, Y. Li, Y.-K. Wang, K. Zhang, L. Zhang (Bytedance)]

[JVET-AF0105](https://jvet-experts.org/doc_end_user/current_document.php?id=13352) AHG9: On color vision deficiency optimization type for encoder optimization info SEI message [C. Kim, Hendry, D. Gwak, J. Lim, S. Kim (LGE)]

[JVET-AF0107](https://jvet-experts.org/doc_end_user/current_document.php?id=13354) AHG9: On miscellaneous aspects for encoder optimization info SEI message [C. Kim, Hendry, D. Gwak, J. Lim, S. Kim (LGE)]

* + 1. **AHG9: Object mask information SEI message aspects (2)**

[JVET-AF0087](https://jvet-experts.org/doc_end_user/current_document.php?id=13335) AHG9: Experimental results of object mask auxiliary picture coding [Z. Zhang, [J. Chen](mailto:jiechen.cj@alibaba-inc.com), Y. Ye, S. Wang (Alibaba)]

[JVET-AF0088](https://jvet-experts.org/doc_end_user/current_document.php?id=13336) AHG9: Software implementation and further fixes of the object mask information SEI message [J. Chen, Z. Zhang, Y. Ye, S. Wang (Alibaba)]

JVET-AF0088 also relates to the mandate to collect software and showcase information for SEI messages.

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

[JVET-AF0088](https://jvet-experts.org/doc_end_user/current_document.php?id=13336) AHG9: Software implementation and further fixes of the object mask information SEI message [J. Chen, Z. Zhang, Y. Ye, S. Wang (Alibaba)]

JVET-AF0088 also relates to the mandate to discuss the TuC for future extensions of VSEI

[JVET-AF0167](https://jvet-experts.org/doc_end_user/current_document.php?id=13425) AHG9: On the Proposed Multiplane Image Information SEI Message [T. Lu, P. Yin, S. Oh, S. McCarthy, W. Husak, G. J. Sullivan (Dolby)]

JVET-AF00167 also relates to the mandate to identify potential needs for additional SEI messages.

* 1. ***Identify potential needs for additional SEI messages***
     1. **AHG9: SEI messages related to generative video (4)**

[JVET-AF0048](https://jvet-experts.org/doc_end_user/current_document.php?id=13296) A Study on Decoder Interoperability of Generative Face Video Compression [B. Chen, S. Yin, J. Chen, Y. Ye (Alibaba), S. Wang (CityU HK)]

[JVET-AF0145](https://jvet-experts.org/doc_end_user/current_document.php?id=13403) AHG9: SEI message for text for generative AI [A. Hinds, S. Wenger (Tencent)]

[JVET-AF0146](https://jvet-experts.org/doc_end_user/current_document.php?id=13404) AHG9: On Face Motion Information for Generative Face Video [H.-B. Teo, J.-Y Thong, K. Jayashree, C.-S. Lim, K. Abe (Panasonic)]

[JVET-AF0234](https://jvet-experts.org/doc_end_user/current_document.php?id=13497) AHG9: Common text for proposed generative face video SEI message [B. Chen, S. Yin, J. Chen, Y. Ye (Alibaba), S. Wang (CityU HK), S. McCarthy, P. Yin, G.-M. Su, A. K. Choudrhury, W. Husak, G.J. Sullivan (Dolby)]

* + 1. **AHG9: Other SEI messages (9)**

[JVET-AF0138](https://jvet-experts.org/doc_end_user/current_document.php?id=13396) AHG8/AHG9: Truncated bit depth support SEI messages [D. Ding, X. Zhao, S. Liu, G. Teniou, S. Wenger (Tencent)]

[JVET-AF0140](https://jvet-experts.org/doc_end_user/current_document.php?id=13398) [AHG9] Inter Picture Dependency SEI message [G. Teniou, S. Wenger, S. Liu (Tencent)]

[JVET-AF0141](https://jvet-experts.org/doc_end_user/current_document.php?id=13399) AHG9: SEI messages for common image metadata formats [A. Hinds, S. Wenger (Tencent)]

[JVET-AF0142](https://jvet-experts.org/doc_end_user/current_document.php?id=13400) [AHG9] Local Film Grain Synthesis using Annotated Regions SEI [G. Teniou, S. Wenger (Tencent)]

JVET-AF0142 also relates to the mandate to study SEI messages in VSEI, VVC, HEVC and AVC.

[JVET-AF0144](https://jvet-experts.org/doc_end_user/current_document.php?id=13402) [AHG9] Adaptive Film grain synthesis using Alpha Channel Information SEI message [G. Teniou, S. Wenger (Tencent)]

JVET-AF0144 also relates to the mandate to study SEI messages in VSEI, VVC, HEVC and AVC.

[JVET-AF0147](https://jvet-experts.org/doc_end_user/current_document.php?id=13405) AHG8/AHG9: On Picture Modality Type [J. Gao, H.-B. Teo, C.-S. Lim, K. Abe (Panasonic)]

[JVET-AF0148](https://jvet-experts.org/doc_end_user/current_document.php?id=13406) [AHG9] Large SEI message signalling [G. Teniou, S. Wenger (Tencent)]

[JVET-AF0149](https://jvet-experts.org/doc_end_user/current_document.php?id=13407) [AHG9] Application-required SEI NAL Units [G. Teniou, S. Wenger (Tencent)]

[JVET-AF0167](https://jvet-experts.org/doc_end_user/current_document.php?id=13425) AHG9: On the Proposed Multiplane Image Information SEI Message [T. Lu, P. Yin, S. Oh, S. McCarthy, W. Husak, G. J. Sullivan (Dolby)]

JVET-AF00167 also relates to the mandate to collect software and showcase information for SEI messages.

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

[JVET-AF0045](https://jvet-experts.org/doc_end_user/current_document.php?id=13289) AHG2/AHG9: On the use of NNPF SEI messages in AVC [M. M. Hannuksela (Nokia)]

1. **Activities**

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

1. **Recommendations**

The AHG recommends to:

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

Topics that could possibly be discussed in BoG: 6.1-6.4, 6.6-6.8

[JVET-AF0010](https://jvet-experts.org/doc_end_user/current_document.php?id=13381) JVET AHG report: Encoding algorithm optimization (AHG10) [P. de Lagrange, A. Duenas, R. Sjöberg, A. Tourapis (AHG chairs)]

1. **Related contributions**

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

* 1. ***DMVR encoder control***
     1. **JVET-AF0057 – EE2-3.5: DMVR with robust MV derivation**

Though this contribution in an EE2 test, it contains an encoder-only aspect (test 3.5b), maxing out the RD cost of coding modes where DMVR is likely to create annoying artefacts. The BD-rate impact can be up to around 2% for high QPs (the normative solution seems to have both much less BD-rate impact and better subjective quality).

* 1. ***Adaptive resolution***
     1. **JVET-AF0058 – AHG12: GOP-based RPR encoder control for ECM**

This contribution proposes to port the GOP-base RPR encoder control feature available in the VTM to ECM. The gains seem to be reduced by more than half, but are still significant (around 1.5%; more than 3% in class A). It is said that subjective quality is also improved.

* 1. ***Local QP optimization***
     1. **JVET-AF0089 - AhG10: Lagrange multiplier and QP adaptation at CTU-level for VVC**

This contribution proposes a method that tries to give more quality weight to local areas that are likely to be reused in other picture (through temporal prediction). It does so by estimating a “distortion propagation factor” with a fast 1st inter pass. BD-Rate of around -4% is reported for LD and -0.25% for RA.

This is probably related to the “block importance mapping” feature present since VTM-16.0 (but not enable that default), that makes use of MCTF-derived metrics, and gives around -2% BD-rate for both RA and LD.

* 1. ***Fast methods***
     1. **JVET-AF0111 – AHG10: MTT split modes early termination**

This contribution further studies an early termination method that reduces encoding runtime for ECM and VTM (JVET-AE0057, adopted in ECM CTC, and VTM but not in CTC), by skipping MTT splits when the no-split cost of 64x64 CU is above a threshold. Other bit depths and larger QP range are tested, and some threshold adjustments are made. Encoding times of -3% to -4% are reported, with a BD-rate difference in the range of +0.02% to +0.06%, the tradeoff being improved compared to JVET-AE0057 for high QP values.

* 1. ***Lamba optimization***
     1. **JVET-AF0122 - AhG10: Lagrange multiplier optimization for chroma ALF and CCALF**

As in JVET-AD0136 and JVET-AE0123, this contribution proposes to adjust the lambda for the RDO of chroma ALF when luma ALF coefficients are to be transmitted, to account for a reduced overhead of the combined signaling, and also encourage combined signaling (resulting in more chroma filter diversity). Chroma gains are reported.

1. **Recommendation**

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

[JVET-AF0011](https://jvet-experts.org/doc_end_user/current_document.php?id=13382) JVET AHG report: Neural network-based video coding (AHG11) [E. Alshina, F. Galpin, S. Liu, A. Segall (co chairs), J. Li, R.-L. Liao, D. Rusanovskyy, T. Shao, M. Wien, P. Wu (vice chairs)]

1. **Activities**

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

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

The AHG released revised common test conditions as decided at the 31st meeting, including the following changes:

* Define the software anchor to be either NNVC-6.0 or the latest version of NNVC.
* Section 2 (Editorial): Change “a set” to “the set”.
* Section 3.3 (Editorial): Clarify how decoder runtime should be calculated when parallel processing is used. Specifically, clarify that the decoding run-time can be calculated in the same way as what’s defined for PSNR and MS-SSIM. That is, it can be calculated by either concatenated the bit-streams and performing a single decode or summing up the parallel decoding runtimes across segments.
* Other editorial improvements.
  + 1. **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>.

* 1. ***EE Coordination***

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

|  |  |  |
| --- | --- | --- |
| JVET-AF0023 | 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) |

* 1. ***Teleconferences***

The AHG conducted a joint teleconference with AHG14 and EE1 during the interim period. The teleconference was held on August 9, 2023. In this teleconference the progress of High Operation Point (HOP) NN-based filter training was discussed. Authors of JVET-AE0191 presented results of the third (last) stage training. Final results were reported in JVET-AF0041.

Low Operation Point (LOP) unified filter training status was also discussed.

On August 30, 2023, a second joint teleconference with AHG 14 was held to discuss further LOP training status. The training process and test results for each stage of training and inference testing are reported in JVET-AF0043.The meeting minutes are reported in JVET-AF0042. :

* 1. ***Performance Evaluation***

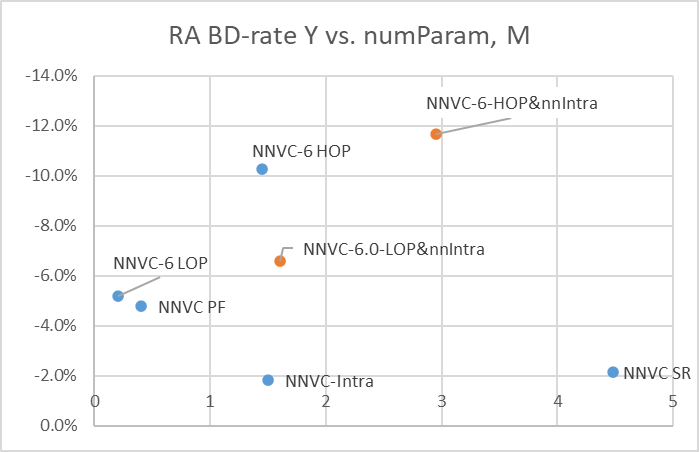
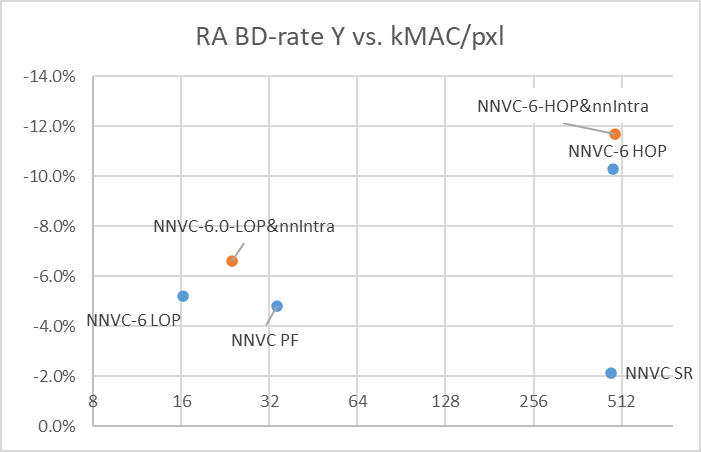
The performance of the NNVC-6.0 anchor compared to VTM anchor is reported below. By default, NN-Intra and Low Operation Point (LOP.1) filter are enabled in NNVC-6.0.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **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 | -7.2% | -4.9% | -6.0% | -8.7% | -4.1% | -5.1% | 134% | 7667% | 96% |
| Class A2 | -6.6% | -5.8% | -5.1% | -6.5% | -4.1% | -2.9% | 130% | 7183% | 97% |
| Class B | -6.3% | -8.2% | -7.5% | -6.0% | -6.1% | -5.2% | 132% | 7834% | 98% |
| Class C | -6.6% | -10.5% | -9.6% | -6.3% | -7.9% | -5.6% | 132% | 7422% | 98% |
| Class E |  |  |  |  |  |  |  |  |  |
| **Overall** | **-6.6%** | **-7.7%** | **-7.3%** | **-6.7%** | **-5.8%** | **-4.8%** | **132%** | **7557%** | **97%** |
| Class D | -7.6% | -9.4% | -9.7% | -5.8% | -6.8% | -5.9% | 129% | 8193% | 98% |
| Class F | -3.1% | -5.9% | -4.9% | -3.0% | -4.8% | -3.5% | 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 | -4.8% | -8.1% | -7.8% | -5.1% | -6.2% | -6.6% | 134% | 6812% | 99% |
| Class C | -5.2% | -11.1% | -9.0% | -5.6% | -10.7% | -4.0% | 126% | 6233% | 99% |
| Class E | -4.9% | -4.1% | -5.2% | -5.6% | -2.2% | -3.4% | 162% | 9877% | 98% |
| **Overall** | **-5.0%** | **-8.1%** | **-7.6%** | **-5.4%** | **-6.7%** | **-4.9%** | **138%** | **7257%** | **99%** |
| Class D | -6.5% | -10.8% | -9.8% | -5.8% | -9.5% | -6.8% | 125% | 6400% | 98% |
| Class F | -2.0% | -4.9% | -4.5% | -2.3% | -3.1% | -4.1% | 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 | -8.5% | -7.2% | -8.7% | -10.0% | -8.0% | -7.5% | 210% | 5134% | 97% |
| Class A2 | -6.6% | -8.0% | -7.9% | -6.9% | -6.4% | -5.4% | 204% | 4416% | 98% |
| Class B | -7.0% | -8.6% | -8.7% | -7.0% | -7.6% | -7.0% | 203% | 4493% | 98% |
| Class C | -7.3% | -9.8% | -9.9% | -7.6% | -8.4% | -7.6% | 190% | 3770% | 98% |
| Class E | -10.3% | -9.0% | -8.8% | -10.1% | -7.5% | -6.4% | 199% | 5149% | 96% |
| **Overall** | **-7.8%** | **-8.6%** | **-8.9%** | **-8.2%** | **-7.6%** | **-6.9%** | **201%** | **4507%** | **97%** |
| Class D | -7.6% | -8.1% | -9.3% | -7.2% | -6.8% | -7.6% | 180% | 3760% | 98% |
| Class F | -4.6% | -6.0% | -5.6% | -4.2% | -4.8% | -4.4% | 155% | 3961% | 98% |
| Class H |  |  |  |  |  |  |  |  |  |

Compression performance of other tools available in NNVC depending on two complexity metrics (kMAC/pxl and total member for parameters) is shown on plots below. Performance of tools under study in EE1 can be found in EE1 report.

Major trends in EE1 tests are

* complexity reduction of High Operation point (HOP) NN-filter, there is good chance to reduce complexity from 477 kMAC/pxl to lower than 400 kMAC/pxl almost w/o performance drop.
* More studies on training parameters, filter usage policy etc. from past filterset are studied in HOP/LOP.
* Performance improvement of Low Operation point (LOP) NN-filter, jointly trained new LOP provides 0.2% Luma gain and 4% Chroma gain over existing in NNVC-6 LOP filter, maintaining almost the same kMAC/pxl;
* Complexity reduction for NN-Intra, studied in EE1 version of NN-Intra reduces complexity from 7.8 KMAC/pxl to 4.8 kMAC/pxl;
* Content adaptation mechanisms of NN-filter are mostly useful in LDB configuration;
* NN-Inter studied in EE1 shows decoder run time increment comparable with other NNVC tools.

****

1. **Input contributions**

There are 35 input contributions related to the AHG mandates. The list of input contributions is provided below.

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

|  |  |  |
| --- | --- | --- |
| **Reporting** | | |
| [JVET-AF0023](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-AF0043](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13287) | AhG11/EE1: Status of the joint EE1-0 (LOP.2) training | [D. Rusanovskyy](mailto:dmytror@qti.qualcomm.com), [Y. Li (Qualcomm)](mailto:yli30@qti.qualcomm.com), [T. Shao](mailto:Tong.Shao@dolby.com), [P. Yin, S. McCarthy (Dolby)](mailto:pyin@dolby.com), [J. N. Shingala](mailto:jay.shingala@ittiam.com), [A. Shyam](mailto:ajayshyam@ittiam.com), [A. Suneja](mailto:ajat.suneja@ittiam.com), [S. P. Badya (Ittiam)](mailto:siddarth.badya@ittiam.com), [J. Li](mailto:lijunru@bytedance.com), [Y. Li](mailto:yue.li@bytedance.com), [C. Lin](mailto:linchaoyi.cy@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com), [R. Chang](mailto:renjiechang@tencent.com), [L. Wang](mailto:liqiangwang@tencent.com), [X. Xu](mailto:xiaozhongxu@tencent.com), [S. Liu (Tencent)](mailto:shanl@tencent.com) |
| [JVET-AF0056](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13304) | EE1-1.2.2: Content-adaptive LOP filter | [R. Yang](mailto:ruiying.yang@nokia.com), [M. Santamaria](mailto:maria.santamaria_gomez@nokia.com), [F. Cricri](mailto:francesco.cricri@nokia.com), [H. Zhang](mailto:honglei.1.zhang@nokia.com), [J. Lainema](mailto:jani.lainema@nokia.com), [M. M. Hannuksela](mailto:miska.hannuksela@nokia.com), [A. Hallapuro (Nokia)](mailto:antti.hallapuro@nokia.com) |
| [JVET-AF0071](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13319) | EE1-related: Further complexity reduction on the joint EE1-0 (LOP.2) unified filter | [T. Shao](mailto:tong.shao@dolby.com), [P. Yin](mailto:pyin@dolby.com), S. McCarthy (Dolby), [J. N. Shingala](mailto:jay.shingala@ittiam.com), [A. Shyam](mailto:ajayshyam@ittiam.com), A. Suneja, S. P. Badya (Ittiam) |
| [JVET-AF0085](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13333) | EE1-1.2.1: On residual adjustments for NNLF | [Z. Dai](mailto:daizhenyu@oppo.com), [Y. Yu](mailto:yue.yu@oppo.com), [H. Yu](mailto:v-yuhaoping@oppo.com), [D. Wang (OPPO)](mailto:wangdong7@oppo.com) |
| [JVET-AF0086](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13334) | EE1-1.2.6: On flipping of input and output of model in NNVC HOP filter | [Z. Xie](mailto:xiezhihuang@oppo.com), [Y. Yu](mailto:yue.yu@oppo.com), [H. Yu](mailto:v-yuhaoping@oppo.com), [D. Wang (OPPO)](mailto:wangdong7@oppo.com) |
| [JVET-AF0102](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13349) | EE1-1.1.2 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) |
| [JVET-AF0103](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13350) | EE1-1.1.3 Study on input feature set optimization | [D. Rusanovskyy](mailto:dmytror@qti.qualcomm.com), [Y. Li](mailto:yli30@qti.qualcomm.com), [M. Karczewicz (Qualcomm)](mailto:martak@qti.qualcomm.com) |
| [JVET-AF0139](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13397) | EE1-3.1: Neural network-based intra prediction with reduced complexity | [T. Dumas](mailto:thierry.dumas@interdigital.com), F. Galpin, P. Bordes (InterDigital) |
| [JVET-AF0153](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13411) | EE1-1.1.1: 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) |
| [JVET-AF0154](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13412) | EE1-1.2.3: Input and output rotation of model for NNVC in-loop filter | [R. Chang](mailto:renjiechang@tencent.com), [L. Wang](mailto:liqiangwang@tencent.com), [X. Xu](mailto:xiaozhongxu@tencent.com), [S. Liu (Tencent)](mailto:shanl@tencent.com) |
| [JVET-AF0181](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13439) | EE1-1.1.4.a: Simplified feature extraction for HOP | [Y. Li](mailto:yue.li@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com) |
| [JVET-AF0182](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13440) | EE1-1.1.4.b: Group convolution for HOP | [Y. Li](mailto:yue.li@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com) |
| [JVET-AF0183](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13441) | EE1-1.1.4.c/d: Separate models for HOP | [Y. Li](mailto:yue.li@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com) |
| [JVET-AF0192](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13453) | EE1-1.1.5: Combination test of EE1-1.1.1 and EE1-1.1.2 | [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), [D. Rusanovskyy](mailto:dmytror@qti.qualcomm.com), [Y. Li](mailto:yli30@qti.qualcomm.com), [M. Karczewicz (Qualcomm)](mailto:martak@qti.qualcomm.com) |
| [JVET-AF0205](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13466) | EE1-1.2.4: An improved inference design of NN-based loop-filters at high operation point | [J. Li](mailto:lijunru@bytedance.com), [Y. Li](mailto:yue.li@Bytedance.com), [C. Lin](mailto:linchaoyi.cy@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com) |
| [JVET-AF0208](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13471) | EE1-2.1: Deep Reference Frame Generation for Inter Prediction Enhancement | [W. Bao](mailto:baoweijie@whu.edu.cn), [X. Chen](mailto:cinched@whu.edu.cn), [J. Jia](mailto:jiajh2021@whu.edu.cn), [Y. Zhang](mailto:yuantongzhang@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) |
| Cross Checks | | |
| [JVET-AF0119](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13366) | Crosscheck of JVET-AF0056 (EE1-1.2.2: Content-adaptive LOP filter) | [J. N. Shingala](mailto:jay.shingala@ittiam.com), [A. Shyam](mailto:ajayshyam@ittiam.com), [S. P. Badya (Ittiam)](mailto:siddarth.badya@ittiam.com) |
| [JVET-AF0216](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13479) | Crosscheck of JVET-AF0056 (EE1-1.2.2: Content-adaptive LOP filter) | [Z. Dai (OPPO)](mailto:daizhenyu@oppo.com) |
| [JVET-AF0217](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13480) | Crosscheck of JVET-AF0139 (EE1-3.1: neural network-based intra prediction with reduced complexity) | [Z. Dai (OPPO)](mailto:daizhenyu@oppo.com) |
| [JVET-AF0254](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13517) | Cross-check of JVET-AF0085 (EE1-1.2.1: on residual adjustments of NNLF) | [T. Dumas (Interdigital)](mailto:thierry.dumas@interdigital.com) |
| [JVET-AF0257](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13520) | Crosscheck of JVET-AF0154 (EE1-1.2.3: Input and output rotation of model for NNVC in-loop filter)) | [D. Liu (Ericsson)](mailto:du.liu@ericsson.com) |
| [JVET-AF0258](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13521) | Crosscheck of JVET-AF0103 (EE1-1.1.3 Study on input feature set optimization) | F. Galpin (InterDigital) |
| [JVET-AF0260](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13523) | Crosscheck of JVET-AF0183 (EE1-1.1.4.c/d: Separate models for HOP) | F. Galpin (InterDigital) |

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

|  |  |  |
| --- | --- | --- |
| **Reporting** | | |
| [JVET-AF0011](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13382) | JVET AHG report: Neural network-based video coding (AHG11) | E. Alshina, F. Galpin, S. Liu, A. Segall (co chairs), J. Li, R.-L. Liao, D. Rusanovskyy, T. Shao, M. Wien, P. Wu (vice chairs) |
| [JVET-AF0041](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13285) | AhG11: HOP full results | [F. Galpin (InterDigital)](mailto:franck.galpin@interdigital.com), D. Rusanovskyy, Y. Li (Qualcomm), Y. Li, J. Li (ByteDance), L. Wang, R. Chang (Tencent), Z. Xie (OPPO), E. Alshina (Huawei) |
| [JVET-AF0042](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13286) | AhG11/AhG14 teleconference | [E. Alshina](mailto:elena.alshina@huawei.com), [F. Galpin](mailto:franck.galpin@interdigital.com) |
| Proposal | | |
| [JVET-AF0143](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13401) | AHG11: Unified CNN super resolution for resampling-based video coding | [C. Lin](mailto:linchaoyi.cy@bytedance.com), [Y. Li](mailto:yue.li@bytedance.com), [J. Li](mailto:lijunru@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com) |
| [JVET-AF0150](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13408) | AhG11: on HOP batch size | [F. Galpin](mailto:franck.galpin@interdigital.com), T. Dumas, P. Bordes (InterDigital) |
| [JVET-AF0155](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13413) | AhG11: on HOP luma/chroma balance | [F. Galpin](mailto:franck.galpin@interdigital.com), T. Dumas, P. Bordes (InterDigital) |
| [JVET-AF0158](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13416) | AHG11: HOP In-loop filter with transformer blocks | [Y. Li](mailto:yli30@qti.qualcomm.com), [D. Rusanovskyy](mailto:dmytror@qti.qualcomm.com), [T. Ryder](mailto:tryder@qti.qualcomm.com), [S. Eadie](mailto:seadie@qti.qualcomm.com), [M. Karczewicz (Qualcomm)](mailto:martak@qti.qualcomm.com) |
| [JVET-AF0180](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13438) | AHG11: HOP training adjustment to improve luma/chroma coding gain balance | [D. Rusanovskyy](mailto:dmytror@qti.qualcomm.com), [Y. Li](mailto:yli30@qti.qualcomm.com), [M. Karczewicz (Qualcomm)](mailto:martak@qti.qualcomm.com) |
| [JVET-AF0193](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13454) | AHG11: Decoder complexity optimization for NNVC in-loop filter | [R. Chang](mailto:renjiechang@tencent.com), [L. Wang](mailto:liqiangwang@tencent.com), [X. Xu](mailto:xiaozhongxu@tencent.com), [S. Liu (Tencent)](mailto:shanl@tencent.com) |
| [JVET-AF0206](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13467) | AHG11: Complexity reduction of NN-based loop-filters | [J. Li](mailto:lijunru@bytedance.com), [Y. Li](mailto:yue.li@Bytedance.com), [C. Lin](mailto:linchaoyi.cy@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com) |
| Cross Checks | | |
| [JVET-AF0256](file:///C:\Users\e00443164\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\8LUVY5N7\current_document.php%3fid=13519) | Crosscheck of JVET-AF0193 (AHG11: Decoder complexity optimization for NNVC in-loop filter) | [D. Liu (Ericsson)](mailto:du.liu@ericsson) |

1. **Recommendations**

The AHG recommends:

* Review all input contributions.
* Continue investigating neural network-based video coding tools, including coding performance and complexity.

It was suggested to also look into more combinations of tools, such as LF and inter pred., super resolution, in the upcoming EE.

[JVET-AF0012](https://jvet-experts.org/doc_end_user/current_document.php?id=13383) JVET AHG report: Enhanced compression beyond VVC capability (AHG12) [M. Karczewicz, Y. Ye, L. Zhang (co-chairs), B. Bross, R. Chernyak, X. Li, K. Naser, Y. Yu (vice-chairs)]

1. **Activities**

The Common Test Conditions were updated (JVET-AE2017). The peak memory usage has been added to the reporting template. The primary activity of the AHG was the “Exploration experiment on enhanced compression beyond VVC capability” (JVET-AE2024). The combined improvements of the ECM-10.0 over VTM-11.0ecm10.0 anchorfor AI, RA and LB configurations are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | All Intra Main10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -10.55% | -21.81% | -29.24% | 915.8% | 439.3% |
| Class A2 | -16.68% | -28.60% | -31.54% | 916.2% | 458.0% |
| Class B | -11.02% | -26.58% | -24.82% | 826.4% | 455.0% |
| Class C | -11.07% | -16.52% | -17.06% | 898.0% | 452.7% |
| Class E | -14.59% | -24.40% | -22.89% | 804.1% | 467.8% |
| Overall | **-12.49%** | **-23.52%** | **-24.63%** | **867.2%** | **454.4%** |
| Class D | -9.17% | -14.38% | -14.35% | 851.0% | 445.1% |
| Class F | -26.21% | -35.26% | -35.17% | 541.1% | 452.7% |
| Class TGM | -39.28% | -47.12% | -46.45% | 457.3% | 496.2% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Random Access Main 10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -22.25% | -27.01% | -36.07% | 882.3% | 784.9% |
| Class A2 | -25.50% | -35.13% | -38.75% | 791.3% | 889.0% |
| Class B | -20.50% | -34.42% | -32.13% | 673.3% | 731.6% |
| Class C | -21.91% | -26.29% | -26.07% | 708.2% | 777.7% |
| Class E |  |  |  |  |  |
| Overall | **-22.23%** | **-30.91%** | **-32.63%** | **744.0%** | **784.1%** |
| Class D | -22.87% | -27.80% | -28.10% | 738.6% | 851.2% |
| Class F | -28.30% | -36.71% | -37.10% | 602.2% | 482.5% |
| Class TGM | -36.63% | -43.02% | -43.08% | 570.5% | 415.0% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Low Delay B Main 10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -17.55% | -38.06% | -35.34% | 589.9% | 668.2% |
| Class C | -19.29% | -28.96% | -29.31% | 648.1% | 710.0% |
| Class E | -16.12% | -28.03% | -27.57% | 642.1% | 459.1% |
| Overall | **-17.77%** | **-32.52%** | **-31.39%** | **621.7%** | **620.8%** |
| Class D | -21.17% | -30.48% | -30.27% | 656.7% | 758.7% |
| Class F | -24.60% | -37.20% | -36.64% | 588.5% | 498.3% |
| Class TGM | -35.27% | -43.12% | -43.29% | 560.3% | 418.2% |

The rate reduction for natural sequences over VTM in RA configuration for {Y, U, V} increased from ECM-9.0’s {-21.03%, -28.42%, -29.93%} to ECM-10.0’s {-22.23%, -30.91%, -32.63%}. For SCC sequences (class TGM) the rate reduction for RA configuration increased from ECM-9.0’s {-35.17%, -41.66%, -41.35%} to ECM-10.0’s {-36.63%, -43.02%, -43.08%}.

1. **Contributions**

In addition to 24 EE2 contributions, 48 (comparing to 52 last meeting) EE2-related and AHG12-related contributions were received. There are also 7 contributions to AHG6 and AHG10 which address ECM encoder performance. The EE2-related and AHG12-related contributions can be subdivided as follows:

* 1. **Intra *(7)***

JVET-AF0059, "AHG12: Fix to interpolation filter for intra prediction", K. Andersson (Ericsson)

JVET-AF0083, "Non-EE2: Enhancements on CCP merge for chroma intra coding", H. Huang, Y. Yu, H. Yu, D. Wang (OPPO), Z. Deng, K. Zhang, L. Zhang (Bytedance)

JVET-AF0106, "EE2-related: Non-adjacent spatial candidates for DIMD merge", J. Huo, J. Fan, Z. Zhang, Y. Ma, F. Yang (Xidian Univ.), M. Li (OPPO)

JVET-AF0114, "AHG12: Local-Boosting on Cross-Component Merge Mode", H. Qin, K. Ding, Z. Xu (TCL)

JVET-AF0176, "Non-EE2: Decoder Derived Cross-Component Prediction", Y.-J. Chang, P.-H. Lin, V. Seregin, J.-L. Lin, M. Karczewicz (Qualcomm)

JVET-AF0199, "Non-EE2: Bilateral Filtering for Intra Prediction", W. Yin, K. Zhang, Y. Wang, L. Zhang (Bytedance)

JVET-AF0246, "AHG12: Harmonization of intra prediction mode derivation from neighboring blocks", D. Kim, K. Kim, J.-H. Son, J. -S. Kwak (WILUS)

* 1. **IntraTMP *and IntraBC (15)***

JVET-AF0053, "EE2-related: Test 2.5a and TIMD fusion improvement", P. Andrivon, M. Blestel (Ofinno), K. Naser, F. Le Léannec, P. Bordes, F. Galpin, A. Robert (InterDigital)

JVET-AF0072, "AHG12: FIBC flag inherits IntraTMP-FLM flag", M.-H. Jia, Y. Gao, Y. -X. Bai, S.-W. Xie, P. Wu, C. Huang (ZTE)

JVET-AF0084, "Non-EE2: Update on IBC-LIC Model Merge mode", L. Zhang, Y. Yu, H. Yu, D. Wang (OPPO)

JVET-AF0115, "Non-EE2: IBC with a further upward-extended reference area for screen content", Y. Kidani, H. Kato, K. Kawamura (KDDI)

JVET-AF0116, "Non-EE2: slope adjustment for IBC LIC", C. Ma, X. Xiu, W. Chen, H.-H.Jhu, C.-W. Kuo, N. Yan, X. Wang (Kwai)

JVET-AF0117, "Non-EE2: extensions of IBC GPM", C. Ma, X. Xiu, W. Chen, H.-H.Jhu, C.-W. Kuo, N. Yan, X. Wang (Kwai)

JVET-AF0129, "Non-EE2: Auto relocated block vector prediction", N. Zhang, K. Zhang, L. Zhang (Bytedance)

JVET-AF0132, "Non-EE2: Optimization of TIMD blending mode", C. Zhou, Z. Lv (vivo)

JVET-AF0137, "AHG12: IntraTMP with Merge Candidates", K. Naser, F. Le Léannec, A. Robert, F. Galpin (InterDigital)

JVET-AF0161, "AHG12: Harmonized IBC/ITMP search area with adaptive sampling", D. Ruiz Coll, B. Chen, (Ofinno), K. Naser, P. Bordes, F. Le Léannec, F. Galpin, (InterDigital)

JVET-AF0162, "AHG12: Fixes to template matching", J. Chen, R.-L. Liao, X. Li, Y. Ye (Alibaba)

JVET-AF0166, "Non-EE2: Intra TMP fusion probing", J.-L. Lin, P.-H. Lin, Y.-J. Chang, Z. Zhang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AF0184, "Non-EE2: Search range optimization for Intra TMP", G. Verba, Z. Zhang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AF0226, "AHG12: SGPM with IntraTMP and IBC", K. Naser, F. Le Léannec, P. Bordes, Y. Chen (InterDigital)

JVET-AF0248, "Non-EE2: TIMD fusion optimization", P. Andrivon, M. Blestel (Ofinno), C. Zhou, Zhuoyi Lv (vivo)

* 1. **Inter *(12)***

JVET-AF0113, "Non-EE2: Modifications of affine merge candidates", K. Kim, D. Kim, J.-H. Son, J.-S. Kwak (WILUS)

JVET-AF0118, "Non-EE2: Regression-based GPM blending", P. Bordes, F. Galpin, K. Naser, F. Urban, K. Reuze, F. Le Leannec, E. Francois (InterDigital)

JVET-AF0121, "AHG12: Adjusting out-of-boundary prediction samples", P. Astola, J. Lainema (Nokia)

JVET-AF0134, "AHG12: AMVP with SbTMVP mode", R.-L. Liao, J. Chen, Y. Ye, X. Li (Alibaba)

JVET-AF0135, "AHG12: Additional SbTMVP candidates", R.-L. Liao, J. Chen, Y. Ye, X. Li (Alibaba)

JVET-AF0160, "Non-EE2: On DMVR Extensions", M. Salehifar, Y. He, K. Zhang, L. Zhang (Bytedance)

JVET-AF0168, "Non-EE2: Geometric partitioning mode with affine prediction", K. Zhang, Z. Deng, L. Zhang (Bytedance)

JVET-AF0170, "Non-EE2: CIIP with subblock-based motion compensation", L. Zhao, K. Zhang, L. Zhang (Bytedance)

JVET-AF0171, "Non-EE2: Sample distance-based weighting for CIIP", J. Zhao, S. Kim (LGE)

JVET-AF0191, "Non-EE2: Enhancements on local illumination compensation", X. Xiu, C. Ma, N. Yan, H.-J. Jhu, C.-. Kuo, W. Chen, X. Wang (Kwai)

JVET-AF0194, "Non-EE2: On LIC flag in merge mode", Y. Zhang, V. Seregin, H. Wang, Z. Zhang, C.-C. Chen, H. Huang, M. Karczewicz (Qualcomm)

JVET-AF0200, "Non-EE2: Extension of local illumination compensation", Y. Wang, K. Zhang, Y. He, H. Liu, L. Zhang (Bytedance)

* 1. ***In* Loop *Filters (4)***

JVET-AF0169, "Non-EE2: Adaptive clipping with signaled lower and upper bounds", K. Cui, Z. Zhang, H. Huang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AF0178, "Non-EE2: Applying fixed filters to output of the Gaussian filter", N. Hu, M. Karczewicz, V. Seregin (Qualcomm)

JVET-AF0179, "Non-EE2: Fixed filter for Chroma ALF", N. Hu, M. Karczewicz, H. Wang, V. Seregin (Qualcomm)

JVET-AF0198, "Non-EE2: Coefficient Precision Adjustment for ALF", W. Yin, K. Zhang, L. Zhang (Bytedance)

* 1. **Transform *(2)***

JVET-AF0082, "Non-EE2: On LFNST/NSPT for inter coding", F. Wang, J. Gan, Y. Yue, H. Yu, D. Wang (OPPO)

JVET-AF0104, "Non-EE2: Direct Decoder-side Intra Mode Derivation for IntraTMP blocks", J. Huo, N. Qiu, Y. Ma, F. Yang (Xidian Univ.), Ming Li (OPPO)

* 1. ***Coefficients and Entropy Coding (4)***

JVET-AF0151, "EE2 related: CABAC parameters retraining", F. Galpin, F. Lo Bianco, C. Salmon-Legagneur, K. Naser (InterDigital)

JVET-AF0164, "AHG12: CABAC inter/intra model switch", F. Lo Bianco, F. Galpin, C. Salmon-Legagneur (InterDigital)

JVET-AF0185, "Non-EE2: Transform Coefficient Coding", P. Nikitin, M. Karczewicz, M. Coban, V. Seregin (Qualcomm)

JVET-AF0202, "EE2 related: CABAC context initialization retraining and slice type based window offsets", V. Seregin, M. Karczewicz (Qualcomm)

* 1. **Other *(4)***

JVET-AF0058, "AHG12: GOP-based RPR encoder control for ECM", K. Andersson, J. Ström, R. Yu, W. Ahmad, P. Wennersten (Ericsson)

JVET-AF0101, "AHG12: ECM software cleanup for decoder side intra predictions acceleration", Z. Fan, Y. Yasugi, T. Ikai (Sharp)

JVET-AF0165, "Ahg12: On VTM-11ecm anchor low-delay test conditions", F. Le Léannec, S. Puri, E. François, K. Naser (InterDigital)

JVET-AF0187, "Compression of Gaming Contents", Z. Lin, K. Cai, J. Sauer, Y. Zhao, E. Alshina (Huawei)

1. **Recommendations**

The AHG recommends to:

* To review all the related contributions.

[JVET-AF0013](https://jvet-experts.org/doc_end_user/current_document.php?id=13384) JVET AHG report: Film grain technologies (AHG13) [W. Husak, P. de Lagrange (co-chairs), A. Duenas, D. Grois, Y. He, X. Meng, M. Radosavljević, A. Segall, A. Tourapis, W. Zhang (vice-chairs)]

1. **Discussion**

For this AHG period, there were three major topics of conversation. ISO/IEC 23002-9 DTR was edited to reflect the National Body (NB) comments received in the CDTR period and style recommendations from SC29. The group also discussed content containing film grain for development, demonstrations, and testing of systems and processes. The final topic discussed was profiling and conformance for film grain synthesis.

There was substantial editing required to resolve the CDTR comments. Most were changes due to editorial comments. There were also some comments that required more detailed technical text additions. Other comments suggested reorganizing parts of the document for flow and appearance. Finally, SC29 advised redrawing the graphics to separate the text from the graphics for editing and translation for other languages.

An extensive amount of time was spent on content. Apple created a dataset using both content captured by Apple and content that is publicly available. The dataset is comprised of 20 sequences with 27 variants and is stored as 16-bit TIFFs yielding approximately 50TB. The dataset includes files that are considered “ground truth” files that are nominally noise free. A collection of representative still images was circulated for feedback and critique. Apple intends to distribute some or all of the content set in the future. Five other companies are looking at providing content.

Another major topic of discussion was profiling and conformance. The group felt profiling was out-of-scope for this AHG period however, the group recognized that conformance was worth exploring. This is because the group discussed conformance in the past and it was recognized that ISO/IEC/ITU could perform a valuable service by describing how an organization could approach conformance. This is not meant to exclude ISO/IEC/ITU as a conformance entity, just that at present, a common method (or methods) for film grain conformance could be explored and documented for other entities that may create conformance test plans and streams.

In the past, AHG13 discussed hard conformance, no conformance, and soft conformance. Due to a number of issues, soft conformance seems comfortable for discussion. The group enumerated several potential sources for inspiration when considering soft conformance including:

1. AOM’s philosophy for film grain conformance
2. Conformance for VPCC and/or V3C
3. Conformance for HDR – SL-HDR(x)

Most of the discussion was centered on AOM’s philosophy with comparable practices being used in V3C/VPCC and SL-HDR(x) for similar reasons. AOM offers two types of conformance methods: 1) hard conformance where a decoder can produce a bit-exact representation of the film grain, and 2) soft conformance that allows implementers to modify their video display path but still produces a “similar” film grain look with a variety of video processing pipelines. What is considered as “similar” is defined by distributors and not AOM.

There was also a discussion about a future version of the report that may contain an annex or even a section describing how a group could create a conformance program. In this way, application SDOs (or even SC29 groups) could have guidelines and tools for developing programs that could be consistent.

1. **Related contributions**

Eight contributions related to AHG13 were identified as of 10/10/2023.

* The one contribution was the AHG report:
  + JVET-AE0013 JVET AHG report: Film grain technologies (AHG13)
* Seven other contributions were added at the time of the report drafting:
  + JVET-AF0142 Local Film Grain Synthesis using Annotated Regions SEI
  + JVET-AF0144 Adaptive Film grain synthesis using Alpha Channel Information SEI message
  + JVET-AF0209 Frequency domain Film Grain Objective Metrics
  + JVET-AF0210 Exploration on the Capability of Frequency-Based Film Grain Synthesis
  + JVET-AF0262 New Film Grain Material based on a Ground Truth approach
  + JVET-AF0263 Suggested process for measuring Grain Fidelity using the Ground Truth test set
  + JVET-AF0274 film grain tuning tool
  1. **Contributions**

There were seven contributions uploaded other than the AHG report.

* + 1. **JVET-AF0142 Local Film Grain Synthesis using Annotated Regions SEI**

Film Grain Synthesis (aka FGS) is a tool aimed at maintaining the original film grain while preserving the encoding efficiency due to its random nature and complexity. While being used to maintain the artistic intent or to mitigate visual artefacts, film grain is not necessarily applicable to the entire picture. This contribution proposes to address the film grain synthesis applied locally on the picture by using the annotated regions as a mask.

The film grain characteristics SEI message has been defined in the VSEI specification to synthetize the grain of the content without encoding it. Today, it lacks the identification of the regions where it is only applicable in the image.

The proposed approach intends to define the labels "FGS-E" and "FGS-D" for FGS enabled and FGS disabled respectively. When one of these labels is identified in the list of labels from the Annotated Regions SEI message, the application of the film grain synthesis is limited to each object with the same label.

* + 1. **JVET-AF0144 Adaptive Film grain synthesis using Alpha Channel Information SEI message**

Film Grain Synthesis (aka FGS) is a tool aimed at maintaining the original film grain while preserving the encoding efficiency due to its random nature and complexity. While being used to maintain the artistic intent or to mitigate visual artefacts, film grain is not seen relevant for the entire sequence or even within an image. This contribution proposes to address the film grain synthesis applied locally on a per image basis by using the auxiliary data as a mask.

The film grain characteristics SEI message has been defined to synthetize the grain of the content without encoding it. Today, it lacks the spatial adaptation to the image.

The VSEI specification also specifies the Alpha Channel Information (ACI) SEI messages that defines how to combine the auxiliary data with the output image for alpha blending.

The ACI SEI message has 2 defined modes in alpha\_channel\_use\_idc: mode 0: multiplicative and mode 1: non multiplicative. A 3rd mode (mode 3, since mode 2 is set to "unspecified") can be defined as the multiplicative use of the auxiliary data with the corresponding simulated film grain, defined by the Film Grain characteristics SEI message.

While interpreting the film grain characteristics SEI message, if the ACI SEI message is also present for the same picture with alpha\_channel\_use\_idc equal to 3, then the associated auxiliary picture is used as a weighting factor of the film grain values.

* + 1. **JVET-AF0209 Frequency domain Film Grain Objective Metrics**

In this proposal, we propose a subband based film grain assessment method (SFGA), which shows promising results. SFGA, designed as a full-reference film grain quality assessment tool, initially transforms both reference and test images from the spatial domain to the frequency domain. It then selects several sub-bands within the frequency domain, compares their sub-band noise power spectra, and subsequently combines the sub-band results. The PLCC and SROCC between the proposed metric and MOS are both larger than 0.75, outperforming existing metrics like NSS (Natural Scene Statistics), FFT (Fast Fourier Transform), PSNR, and SSIM. This metric not only enhances our understanding of film grain technology, but also provides a valuable tool for fine-tuning film grain parameters in the modeling and synthesis process.

* + 1. **JVET-AF0210 Exploration on the Capability of Frequency-Based Film Grain Synthesis**

The state-of-the-art film grain synthesis framework cannot meet the needs of high-quality film grain synthesis. This deficiency may primarily stem from limitations in the denoiser, film grain modelling, or film grain synthesis. To isolate the denoiser's impact, we captured clean images and introduced film grain using Davinci Resolve. Then we evaluated the subjective results of film grain modelling and synthesis. We can see that for finer-grained film grain it is challenging for the frequency model to generate film grain effectively. For coarser-grained film grain, the AR model faces difficulties in generating film grain effectively. To further confirm the bottleneck of the film grain framework, we explored the capability of frequency-based film grain synthesis by generating 16470 film grain models and evaluating the 16470 synthesis results. Experimental results show that,

1. In most cases, we can find a suitable film grain model for film grain synthesis, which means the capability of film grain SEI message and synthesis is sufficient.
2. In the current film grain synthesis framework, the film grain modeling highly constraints the performance of film grain synthesis.
3. The proposed metric SFGA in JVET-AF0209 can help to fine-tune the film grain models.
   * 1. **JVET-AF0262 New Film Grain Material based on a Ground Truth approach**

This contribution presents the film grain content material that was generated using the ground truth method discussed in contributions JVET-AD0369 and JVET-AE0250. The material consists of either newly captured content or scenes taken from the OpenMovie project “Tears of Steel”. A total of 26 noise variants is created for each scene selected, with the noise variants exhibiting different noise characteristics. The proponents of the material suggest that the group study, and if desired use, such content in its different activities.

* + 1. **JVET-AF0263 Suggested process for measuring Grain Fidelity using the Ground Truth test set**

This contribution presents a suggested framework/method of how to potentially measure film grain fidelity using the Ground Truth test-set presented in contribution JVET-AF0262 for the evaluation of proposals and technologies relating to film grain analysis or synthesis. The idea is based on techniques that have already been provided to the group, e.g. in contributions JVET‑X0048 and JVET­‑X0103, for performing film grain analysis in the context of estimating film grain parameters using the Film Grain characteristics SEI message. The framework only provides suggestions of measuring the fidelity of noise and not of the final content with the added noise. No actual implementation is provided.

* + 1. **JVET-AF0274 film grain tuning tool**

This contribution describes a new tool to graphically display and edit an FGC SEI configuration file, available in the JVET gitlab (in “JVET AHG Film Grain Technologies” group).

As announced during the 30th JVET meeting, a group called “JVET AHG Film Grain Technologies” has been created on gitlab (https://vcgit.hhi.fraunhofer.de/jvet-ahg-fgt). This group contains software developments related to film grain that are either independent or not yet merged into reference software.

A recent addition is “FGC SEI designer”, a graphical interactive tool to edit an FCG SEI config file, that can be useful when automatic estimation is either not practical or needs to be refined. It can also be used to export a graphical representation of film grain parameters, for use e.g. in a document.

As it is implemented as a single python script, and supposed to be connected to a film grain synthesizer, it was just added to the “VFGS” software repository (https://vcgit.hhi.fraunhofer.de/jvet-ahg-fgt/vfgs), rather than creating a new repository. External dependencies are matplotlib and numpy.

The tool is also available from InterDigital github (https://github.com/InterDigitalInc/VersatileFilmGrain), which is in fact its primary source.

1. **Recommendations**

The AHG recommends:

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

It was commented that the definition of “soft conformance” is still rather vague.

[JVET-AF0014](https://jvet-experts.org/doc_end_user/current_document.php?id=13385) JVET AHG report: NNVC software development (AHG14) [F. Galpin (chair), Y. Li, Y. Li, J. Shingala, L. Wang, Z. Xie (vice chairs)]

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

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

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

NNVC-6.0 anchor at <https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/VVCSoftware_VTM> is used for NNVC performance evaluation.

* 1. ***Software changes***

Several commits were merged in the NNVC repository.

The following changes were integrated:

* HOP related
  + JVET-AE0191 HOP models
  + JVET-AE0289 HOP stage 3 training process
  + JVET-AE0291 HOP stage 3 dataset
  + JVET-AE0214 SADL v6
  + clarify documentation
  + naive quantization script
* fixes:
  + training hyper-parameters
  + 3 issues related to residual scaling in HOP (available in NNVC-6.1)
  + avoid log overwrite when resuming training
  + port JVET-AE0180 (memory reduction)
  + correct resume training hyper-parameters reloading
* cleaning:
  + remove dead code
  + remove dead features
  1. ***Software version***

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

NNVC-6.0 was tagged September 6th, 2023.

NNVC-5.1 was tagged July 19th, 2023.

NNVC-5.0 was tagged May 11th, 2023.

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

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

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

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

1. **CTC performance**

See configurations section for naming convention.

* 1. **Comparison *to VTM***
     1. ***Comparison* to previous version - NNVC-5.0 vs NNVC-6.0 VTM mode**

NNVC-6.0 in VTM mode performance are the same as the ones in NNVC-5.0 VTM (only changes are memory consumption).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over nnvc5.0\_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% | 99% | 99% |
| Class A2 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 99% | 99% |
| Class B | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 99% | 98% |
| Class C | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 99% | 98% |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 99% | 98% |
| Class D | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 99% | 98% |
| Class F | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 99% | 98% |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Low delay B Main10** | | | | | | | |
|  | **BD-rate Over nnvc5.0\_VTM-anchor** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 |  |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |  |
| Class B | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 99% | 99% |
| Class C | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 99% | 100% |
| Class E | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 99% | 100% |
| **Overall** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 99% | 100% |
| Class D | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 99% | 98% |
| Class F | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 97% | 98% |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over nnvc5.0\_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% | 99% | 100% |
| Class A2 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 99% | 100% |
| Class B | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 99% |
| Class C | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 99% | 98% |
| Class E | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 99% | 100% |
| **Overall** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 99% | 99% |
| 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% | 98% | 96% |

* + 1. **NNVC-6.0 VTM mode vs NNVC-6.0 anchor**

The NNVC-6.0 anchor includes LC filter and Intra Prediction tools activated. Results are the same as for NNVC-5.0 anchor (results are copied here for reference)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **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% |

* + 1. **NNVC-6.0 VTM mode vs NNVC-6.0 VTM mode + HOP (HOP only)**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over nnvc6.0\_VTM-anchor** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -10.52% | -20.62% | -25.85% | -11.91% | -22.85% | -27.10% | 176% | 49507% |
| Class A2 | -10.79% | -23.75% | -22.70% | -10.55% | -23.08% | -19.47% | 169% | 47472% |
| Class B | -9.65% | -25.88% | -24.33% | -9.17% | -25.28% | -24.84% | 171% | 49184% |
| Class C | -10.49% | -27.44% | -27.39% | -9.82% | -24.33% | -25.44% | 150% | 44716% |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -10.27% | -24.82% | -25.12% | -10.17% | -24.10% | -24.38% | 166% | 47674% |
| Class D | -11.88% | -26.75% | -27.45% | -9.11% | -24.06% | -23.76% | 145% | 41584% |
| Class F | -5.81% | -18.45% | -17.28% | -6.60% | -19.43% | -18.41% | 218% | 18932% |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Low delay B Main10** | | | | | | | |
|  | **BD-rate Over nnvc6.0\_VTM-anchor** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 |  |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |  |
| Class B | -9.46% | -20.43% | -14.83% | -9.63% | -21.13% | -17.32% | 166% | 48578% |
| Class C | -10.61% | -25.51% | -22.61% | -10.89% | -25.00% | -21.30% | 148% | 44637% |
| Class E | -11.79% | -15.78% | -6.88% | -12.19% | -16.97% | -7.79% | 266% | 54934% |
| **Overall** | -10.43% | -20.96% | -15.43% | -10.69% | -21.38% | -16.27% | 180% | 48701% |
| Class D | -11.81% | -23.44% | -16.11% | -10.59% | -23.09% | -14.53% | 141% | 41240% |
| Class F | -6.98% | -19.22% | -14.15% | -8.44% | -20.83% | -20.68% | 219% | 23962% |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over nnvc6.0\_VTM-anchor** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -7.20% | -18.27% | -22.13% | -8.54% | -20.42% | -23.39% | 180% | 34113% |
| Class A2 | -7.09% | -20.09% | -18.85% | -7.96% | -19.89% | -16.46% | 142% | 28458% |
| Class B | -6.98% | -19.22% | -19.78% | -7.01% | -20.82% | -21.74% | 134% | 26660% |
| Class C | -8.08% | -18.41% | -21.09% | -7.96% | -20.09% | -22.71% | 118% | 18245% |
| Class E | -10.45% | -20.65% | -21.76% | -10.85% | -20.11% | -23.27% | 140% | 29315% |
| **Overall** | -7.86% | -19.27% | -20.64% | -8.27% | -20.32% | -21.60% | 139% | 26224% |
| Class D | -7.92% | -17.77% | -21.44% | -7.11% | -21.03% | -22.49% | 114% | 16314% |
| Class F | -5.27% | -15.78% | -15.91% | -5.35% | -17.94% | -18.72% | 120% | 21783% |

* + 1. **NNVC-6.0 VTM mode vs NNVC-6.0 + HOP**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |  |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | |  |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU | DecT GPU |
| Class A1 | -12.25% | -21.90% | -26.99% | -13.64% | -23.87% | -28.04% | 192% | 49788% | #DIV/0! |
| Class A2 | -11.87% | -24.22% | -23.45% | -11.67% | -23.33% | -20.04% | 183% | 47404% | #DIV/0! |
| Class B | -11.02% | -26.65% | -25.11% | -10.45% | -25.97% | -24.94% | 187% | 49058% | #DIV/0! |
| Class C | -11.78% | -28.42% | -28.35% | -11.11% | -25.13% | -25.96% | 166% | 44554% | #DIV/0! |
| Class E |  |  |  |  |  |  |  |  |  |
| **Overall** | -11.64% | -25.69% | -26.02% | -11.51% | -24.80% | -24.85% | 181% | 47628% | #DIV/0! |
| Class D | -12.99% | -27.60% | -28.21% | -10.31% | -24.82% | -24.56% | 159% | 41349% | #DIV/0! |
| Class F | -6.64% | -18.95% | -17.80% | -7.35% | -19.89% | -18.68% | 234% | 18908% | #DIV/0! |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Low delay B Main10** | | | | | | | | | | | | | | | |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | | | | | | | | | |
|  | Y-PSNR | U-PSNR | | V-PSNR | | Y-MSIM | | U-MSIM | | V-MSIM | | EncT | | | DecT CPU | |
| Class A1 | #VALUE! | #VALUE! | | #VALUE! | | #VALUE! | | #VALUE! | | #VALUE! | | #DIV/0! | | | #DIV/0! | |
| Class A2 | #VALUE! | #VALUE! | | #VALUE! | | #VALUE! | | #VALUE! | | #VALUE! | | #DIV/0! | | | #DIV/0! | |
| Class B | -9.87% | -20.94% | | -15.27% | | -10.13% | | -21.91% | | -18.18% | | 174% | | | 48596% | |
| Class C | -10.84% | -25.68% | | -22.82% | | -11.11% | | -25.25% | | -21.39% | | 153% | | | 44563% | |
| Class E | -11.89% | -16.58% | | -6.94% | | -12.09% | | -17.05% | | -7.82% | | 269% | | | 54898% | |
| **Overall** | -10.70% | -21.43% | | -15.70% | | -10.94% | | -21.81% | | -16.66% | | 186% | | | 48674% | |
| Class D | -11.97% | -23.63% | | -15.96% | | -10.74% | | -22.83% | | -13.85% | | 147% | | | 41272% | |
| Class F | -7.11% | -19.37% | | -14.40% | | -8.62% | | -21.48% | | -21.24% | | 225% | | | 23976% | |
|  | **All Intra Main10** | | | | | | | | | | | | | | |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | | | | | | | | |
|  | Y-PSNR | | U-PSNR | | V-PSNR | | Y-MSIM | | U-MSIM | | V-MSIM | | EncT | DecT CPU | |
| Class A1 | -10.80% | | -21.61% | | -25.21% | | -12.42% | | -24.08% | | -26.37% | | 257% | 34651% | |
| Class A2 | -9.50% | | -22.21% | | -20.70% | | -10.28% | | -22.04% | | -18.20% | | 221% | 29092% | |
| Class B | -9.67% | | -21.93% | | -22.37% | | -9.68% | | -23.46% | | -23.99% | | 211% | 27305% | |
| Class C | -10.66% | | -21.34% | | -23.99% | | -10.48% | | -22.56% | | -25.23% | | 187% | 18934% | |
| Class E | -14.31% | | -24.62% | | -25.51% | | -14.72% | | -23.78% | | -26.51% | | 212% | 29758% | |
| **Overall** | -10.82% | | -22.24% | | -23.45% | | -11.25% | | -23.18% | | -24.12% | | 214% | 26851% | |
| Class D | -10.32% | | -20.91% | | -24.43% | | -9.53% | | -23.50% | | -25.27% | | 176% | 17077% | |
| Class F | -7.18% | | -17.41% | | -17.55% | | -7.14% | | -19.12% | | -20.80% | | 161% | 22168% | |

* 1. **Comparison *to NNVC-6.0 anchor***
     1. ***NNVC*-6.0 anchor vs NNVC-6.0 HOP**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 anchor** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -5.41% | -17.63% | -22.03% | -5.35% | -20.16% | -23.98% | 145% | 679% |
| Class A2 | -5.71% | -19.22% | -19.10% | -5.49% | -19.70% | -17.46% | 142% | 695% |
| Class B | -5.06% | -19.76% | -18.79% | -4.75% | -21.02% | -20.82% | 141% | 640% |
| Class C | -5.52% | -19.47% | -20.33% | -5.15% | -18.40% | -21.14% | 128% | 637% |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -5.38% | -19.15% | -19.91% | -5.12% | -19.88% | -20.86% | 138% | 658% |
| Class D | -5.84% | -19.48% | -19.97% | -4.76% | -18.71% | -19.54% | 129% | 567% |
| Class F | -3.74% | -13.98% | -13.51% | -4.46% | -16.07% | -15.83% | 158% | 207% |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Low delay B Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 anchor** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! |
| Class A2 | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! |
| Class B | -5.29% | -13.80% | -7.93% | -5.25% | -16.62% | -12.00% | 128% | 606% |
| Class C | -5.91% | -16.15% | -15.00% | -5.80% | -16.41% | -18.24% | 120% | 617% |
| Class E | -7.28% | -12.89% | -2.29% | -6.77% | -15.05% | -5.00% | 163% | 434% |
| **Overall** | -5.99% | -14.36% | -8.88% | -5.81% | -16.16% | -12.33% | 133% | 561% |
| Class D | -5.79% | -14.12% | -6.95% | -5.27% | -14.49% | -7.18% | 117% | 556% |
| Class F | -5.19% | -15.13% | -10.53% | -6.60% | -18.51% | -17.78% | 149% | 209% |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 anchor** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -2.58% | -15.45% | -18.02% | -2.61% | -17.91% | -20.40% | 119% | 592% |
| Class A2 | -3.12% | -15.30% | -13.68% | -3.64% | -16.63% | -13.39% | 106% | 602% |
| Class B | -2.88% | -14.39% | -14.82% | -2.83% | -17.15% | -18.16% | 102% | 540% |
| Class C | -3.63% | -12.62% | -15.24% | -3.12% | -15.61% | -18.81% | 96% | 464% |
| Class E | -4.52% | -16.74% | -18.11% | -5.00% | -17.14% | -21.34% | 105% | 529% |
| **Overall** | -3.31% | -14.72% | -15.81% | -3.35% | -16.85% | -18.41% | 104% | 538% |
| Class D | -2.98% | -13.49% | -16.04% | -2.54% | -17.60% | -18.62% | 94% | 393% |
| Class F | -2.82% | -12.02% | -12.45% | -2.94% | -14.92% | -16.77% | 100% | 495% |

* 1. ***Comparison to NNVC-6.0 HOP***
     1. **NNVC-6.0 anchor NNVC-6.1 HOP**

A bug fix was performed on top of NNVC 6.0 and tagged 6.1. It only affects the HOP results. Please note that timing is not reliable.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -0.08% | -0.57% | -0.40% | 0.01% | -0.42% | -0.17% | 140% | 153% |
| Class A2 | -0.04% | -0.15% | 0.07% | -0.08% | -0.08% | 0.02% | 139% | #NUM! |
| Class B | -0.05% | 0.00% | -0.53% | -0.14% | -0.36% | -0.62% | 142% | 159% |
| Class C | 0.03% | -0.24% | -0.32% | 0.02% | -0.29% | -0.42% | 143% | 170% |
| Class E |  |  |  |  |  |  |  |  |
| **Overall** | -0.03% | -0.21% | -0.33% | -0.06% | -0.30% | -0.35% | 141% | #NUM! |
| Class D | 0.08% | -0.33% | -0.45% | 0.12% | -0.26% | -0.22% | 141% | 170% |
| Class F | -0.16% | -0.71% | -0.50% | 0.17% | -0.48% | -0.44% | 143% | 175% |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Low delay B Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
| Class A2 | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
| Class B | 0.03% | -0.38% | -3.75% | 0.13% | -0.96% | -3.75% | 138% | #NUM! |
| Class C | -0.01% | -0.17% | -0.89% | 0.03% | -0.11% | -1.18% | 139% | 166% |
| Class E | 0.06% | -0.82% | -0.70% | -0.07% | -0.61% | -0.61% | 145% | 150% |
| **Overall** | 0.03% | -0.42% | -2.04% | 0.05% | -0.59% | -2.11% | 140% | #NUM! |
| Class D | -0.06% | -2.23% | -2.57% | 0.13% | -3.51% | -2.84% | 138% | 170% |
| Class F | -0.07% | -0.06% | -0.74% | 0.06% | -0.56% | -1.02% | 143% | #NUM! |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | | | | |
|  | **BD-rate Over NNVC-6.0 VTM** | | | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-MSIM | U-MSIM | V-MSIM | EncT | DecT CPU |
| Class A1 | -0.14% | -0.17% | 0.75% | -0.16% | 0.02% | 0.52% | 144% | 147% |
| Class A2 | -0.02% | -0.13% | -0.04% | -0.04% | 0.19% | -0.07% | 143% | 146% |
| Class B | -0.06% | 0.04% | -0.09% | -0.07% | -0.04% | -0.18% | 141% | 149% |
| Class C | -0.09% | -0.92% | -0.90% | -0.12% | -0.82% | -0.86% | 140% | 163% |
| Class E | -0.14% | 0.78% | 0.56% | -0.08% | 0.56% | 0.37% | 142% | 152% |
| **Overall** | -0.09% | -0.11% | -0.01% | -0.09% | -0.07% | -0.10% | 142% | 152% |
| Class D | -0.05% | -0.12% | -0.31% | 0.01% | -0.11% | -0.58% | 137% | 166% |
| Class F | -0.23% | -0.47% | 1.12% | -0.08% | -0.72% | 1.86% | 131% | 165% |

* 1. ***Other tools***

Other results remain the same as tools were not changed.

1. **Discussions**
   1. ***Previous discussions***

The following points were discussed:

* Final training strategy for HOP model: final training and models were decided.
* Some aspects of encoder optimization (eg RDO models) are not yet aligned with HOP model.

1. **Configurations**

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

The column “tested” is read as follow:

* Y: the configuration has been tested using the new NNVC-6.0 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-6.0 VTM mode | none | encoder\_xxx\_vtm.cfg | Y | Y |
| NNVC 6.0 Anchor/EE1 | Intra Pred + LC filterset | encoder\_xxx\_nnvc.cfg | Y | Y |
| NNVC-6.0. HOP | Intra Pred + HOP | encoder\_xxx\_nnvc.cfg + nn-based/HOP.cfg | Y |  |
| 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 |

1. **Recommendations**

The AHG recommends to:

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

It was commented that at some point it might again be useful to test combinations of ECM and NNVC tools.

# Project development (23)

## AHG1: Deployment and advertisement of standards (3)

Contributions in this area were discussed at 1710–1750 on Thursday 19 Oct. 2023 (chaired by JRO).

[JVET-AF0020](https://jvet-experts.org/doc_end_user/current_document.php?id=13447) Deployment status of the HEVC standard [G. J. Sullivan (SC 29 chair & VCEG rapporteur)]

This information contribution contains a survey of deployed products and services using the HEVC standard and the formal specifications in which it is supported, along with a brief introduction to the standard written for broad readership. Revision marking is included to show changes relative to JVET-AD0020-v1 of April 2023.

1. As of April–June 2023, ScientiaMobile reported statistics for video capabilities of mobile devices, based on a sampling of billions of device events, reporting that 93% of smartphone usage was from devices with hardware support for HEVC decoding, most typically supporting level 5 for 4096×2160 @ 30.0 fps or level 5.1 for 4096×2160 @ 60.0 fps [41].
2. In August 2023, Telestream reported that HEVC usage had been rising as a percentage of streaming video bitstream usage, reaching 23% usage with a steady increase since 2018 when it had about half that share [42]. They reported “a shift towards the adoption of HEVC as the preferred codec in various applications” and reported that “its utilization is likely to further grow” [42].

[JVET-AF0021](https://jvet-experts.org/doc_end_user/current_document.php?id=13448) Deployment status of the VVC standard [G. J. Sullivan (SC 29 chair & VCEG rapporteur)]

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

Revision marking is included to show changes relative to JVET-AD0021-v7 of April 2023.

1. Version 1.9.0 rc2 of VVenC and version 2.1.0 of VVdeC were released in July 2023 [36]. The main changes for VVenC v1.9.0 rc2 since version 1.8.0 included rate capping for rate control encoding (capping at minimum 1.5x target rate), first pass speedups by the means of temporal and spatial subsampling (the latter feature being experimental), and initial experimental support for frames-in-flight parallelism [36]. The two main changes for VVdec since version 1.6.1 were initial implementation of error resilience and initial native support for ARMv8 architecture [36].
2. **NHK** released software providing multilayer support on top of Fraunhofer HHI’s VVenC v1.6.1 encoder to GitHub in July 2023 [42].
3. **Electronic Team, Inc.** released a version 8.16 of its video player for macOS called the **Elmedia Player** in July 2023 [58][59][60]. It was based on the VVdeC software and could achieve 8K video playback.
4. Updated information about the Ali266 encoder was reported in October 2023 [87]. Compared to x265 (v3.5), the updated Ali266 encoder reportedly provided over 60% bit rate savings relative to x265’s “veryslow” preset at similar encoding speed for non-real-time applications, and over 40% bit rate savings for 1080p sequences relative to x265’s “medium” preset at similar encoding speed for real-time low-latency applications such as live streaming. The average encoding speed of 1080p sequences was reportedly 60 fps on a mainstream PC for the “faster” preset.
5. The **Youku** video streaming service had been commercially deployed to serve video to televisions, including playing video to more than 30 recent television models as of October 2023 [87]. With a bit rate savings of 35% to 55% below that of HEVC encoding for movies or TV drama series encoded by the Ali266 encoder at 1080p and 720p, the reduction in bit rate reportedly enabled a reduction in the playback stall rate of approximately 50%. The deployed Youku service had reportedly delivered VVC encoded video to tens of thousands of devices, with a cumulative number of playback events in the millions.
6. At IBC in September 2023, Spin Digital demonstrated live 8K 60 fps 10-bit HDR VVC encoding, streaming, and playback [94].
7. **MC-IF** published its first release candidate (v1.0-rc) of VVC technical guidelines for broadcast & streaming applications on 15 September 2023, with planned finalization on 10 November 2023 [108][109]. MC-IF is an industry forum with the purpose of furthering the adoption of MPEG standards and has been initially focusing on VVC. The MC-IF VVC technical guidelines document provides a general overview of VVC standard and its usage and configuration aspects including, pre-processing, encoding and post-decoding processes, as well as providing more specific information on VVC usage in broadcast and streaming applications with an -application-centric perspective based on the VVC Main 10 profile.

[JVET-AF0175](https://jvet-experts.org/doc_end_user/current_document.php?id=13433) MC-IF VVC technical guidelines [L. Litwic (Ericsson), S. McCarthy (Dolby), S. Wenger, J. Ridge (Nokia), B. Bross (Fraunhofer HHI), D. Rusanovskyy (Qualcomm), A. Stein (InterDigital), J. Outters (Ateme), T. Suzuki (Sony), Y.-J. Chiu (Intel), J. Lemotheux (Orange)]

Media Coding Industry Forum (MC-IF) Interoperability WG has published first release candidate (v1.0-rc) of the VVC technical guidelines for broadcast & streaming applications which aim to:

* cover best practices of VVC configuration in terms of functionality and compression performance for industry relevant VVC-based profiles and interoperability points,
* provide guidance and up to date information on VVC operating bitrate ranges for selected applications,
* provide information on the usage of VVC with accompanying technologies e.g., such as Versatile Supplemental Enhancement Information messages,
* advocate interoperability and seek commonality of VVC usage across transport and application layer standards and relevant industry guidelines.

The guidelines are available for download at <https://www.mc-if.org/wp-content/uploads/2023/09/MC-IF-VVC-guidelines-v1.0-rc.pdf> (no registration required).

This release candidate version (v1.0-rc) is available for community review. Comments and error reports against the technical guidelines can be submitted to: [vvc-guidelines@lists.mc-if.org](mailto:vvc-guidelines@lists.mc-if.org) by 10th November 2023.

## AHG2: Text development and errata reporting (1+2)

Contributions in this area were discussed at 1625–XXXX on Monday XX Oct. 2023 (chaired by XXX).

[JVET-AF0045](https://jvet-experts.org/doc_end_user/current_document.php?id=13289) AHG2/AHG9: On the use of NNPF SEI messages in AVC [M. M. Hannuksela (Nokia)]

See section 6.3

[JVET-AF0063](https://jvet-experts.org/doc_end_user/current_document.php?id=13311) On MV-HEVC profiles [Y.-K. Wang, H. Liu, L. Zhang, S. Jiao, C. Hu, J. Cui, G. Xu (Bytedance)]

It is asserted that, although the syntax of MV-HEVC allows for an output layer set (OLS) in an MV-HEVC bitstream to have both texture layers and depth layers, the profile definitions do not allow this. Recently, applications have come up, e.g., virtual reality, that benefit from all the following three aspects simultaneously: 1) using inter-view prediction between different views, 2) having depths associated with the coded views to be available to the decoder and the renderer, and 3) using existing HEVC hardware encoder/decoder designs. It is also asserted that in such applications it is desirable that the bit depth of the depth component can be of up to 16 bits.

This contribution proposes the following:

1. Update the definitions of the four new profiles, namely the Multiview Main 10, Multiview Monochrome, Multiview Monochrome 10, and Multiview Monochrome 12 profiles, to allow for an OLS containing a layer conforming to such a profile to have both texture layers and depth layers.
2. Add a new profile, named the Multiview Monochrome 16 profile, where it is also allowed for an OLS containing a layer conforming to this profile to have both texture layers and depth layers.

The proposed changes, including some minor editorial changes not summarized above, are included in an attachment, with all the proposed changes marked relative to JVET-AE1006-v1.

It was reported that the existing definition of the new Multiview profiles does not support carriage of depth maps. A text is provided to release the constraint in the profile definition. In that context, the bit depth of the depth map can be larger than the bit depth of the texture. There are also a number of additional editorial improvements.

The question was raised whether also a new 8-bit Multiview profile should be defined that supports carriage of depth. It was later found that this is not necessary, as the wexisting profile already allows depth (see below).

After offline discussion between A. Tourapis, I. Moccagatta, and Y.-K. Wang, a new version of the contribution was produced which was presented on Thu. 19 Oct. 1230.

The following aspects were newly observed:

1. The multilayer extensions of HEVC allow two ways of representing depth:
   1. Via DepthLayerFlag  = =  1 (*depth*), and/or
   2. Via AuxId  = =  AUX\_DEPTH (*auxiliary depth*)

And both of the following combinations are allowed:

1. DepthLayerFlag  = =  1  &&  AuxId  = =  AUX\_DEPTH
2. DepthLayerFlag  = =  0  &&  AuxId  = =  AUX\_DEPTH
3. Although the current definitions of the existing Multiview Main profile and the being-developed Multiview Main 10 profile do not allow having both texture layers and *depth layers* in an OLS, they do allow having both *auxiliary depth layers* with chroma\_format\_idc equal to 1 and texture layers in an OLS but do not allow *having* *both auxiliary depth layers with chroma\_format\_idc equal to 0 and texture layers in an OLS* (referred to as “*the wanted combination*” below for simplicity).
4. Identified another asserted bug for the profile definitions in the CDAM text, on the Multiview Main 10 profile: missing the constraint of colour\_mapping\_enabled\_flag equal to 0 only.

In JVET-AF0063-v2, instead of proposing to allow having both texture layers and *depth layers* in an OLS, it is proposed to allow *the wanted combination*. Two options summarized as follows are proposed:

1. Update the being-developed Multiview Main 10 profile to allow *the wanted combination*.
2. Keep the being-developed Multiview Main 10 profile disallowing *the wanted combination*, but add a second Multiview Main 10 profile and also a second Multiview Main profile that both allow *the wanted combination*.

Besides, both options also include the addition of the Multiview Monochrome 16 profile and updates to the being-developed Multiview Monochrome, Multiview Monochrome 10, and Multiview Monochrome 12 profiles, and all these monochrome profiles allow *the wanted combination*.

It was commented that the design of the existing 8-bit Multiview profile was done as it is (coding depth in a 4:2:0 fashion without filling the chroma channels with any information), because existing hardware could right away be used for decoding texture and depth.

It was argued by the proponent that in the 10 bit domain, most decoding hardware (consumer sector) would support monochrome. This however does not seem to be the case, according to other experts’ opinion. There are so many HEVC Main10 decoder chips in the market, that it is difficult to assess, anyway.

For 12 bit, this situation would be different, as in range extensions monochrome is implicit.

Defining two different 10 bit Multiview profiles is inappropriate. Furthermore, the unnecessary processing of not-used chroma channels in 4:2:0 is not a heavy burden, even less if depth maps are downsampled in practical texture+depth applications

It was concluded to keep the definition for the Multiview main 10 profile as it is in terms of encoding depth maps. This also retains the same concept as in the 8 bit profile.

[JVET-AF0064](https://jvet-experts.org/doc_end_user/current_document.php?id=13312) AHG2: Editor commentary on some editorial issues [Y.-K. Wang, G. J. Sullivan (Editors)]

Presented Wed. 18 Oct. 1350

The following observations are reported:

* The AVC, HEVC, VVC, and VSEI specs define the descriptor “b(8)”, and use it in a few places. Its interpretation is basically the same as u(8).

It is defined as follows:

– b(8): byte having any pattern of bit string (8 bits). The parsing process for this descriptor is specified by the return value of the function read\_bits( 8 ).

The usage of b(8) in the four specs is as follows:

* 3 times in VVC, for rbsp\_byte[ ] only.
* 9 times in HEVC, for rbsp\_byte[ ], itu\_t\_t35\_country\_code, itu\_t\_t35\_country\_code\_extension\_byte, itu\_t\_t35\_payload\_byte, user\_data\_payload\_byte, picture\_md5[ cIdx ][ I ], and reserved\_sei\_message\_payload\_byte.
* 9 times in AVC, similarly as in HEVC, but replacing picture\_md5[ cIdx ][ I ] with priority\_id\_setting\_uri[ ].
* Uses that are in HEVC but not in VVC are basically all in VSEI.

Should some remark be added to the definition of b(8) that explicitly says it is functionally the same as u(8) but is used in contexts where there is a desire to emphasize that the interpretation of the syntax element involves the processing of bytes and each byte is not necessarily interpreted as an 8-bit unsigned integer (e.g., it could be interpreted as a character)?

* In VVC, f(1) is usually used for a bit that has a fixed value and that is not reserved, e.g., forbidden\_zero\_bit, vps\_ptl\_alignment\_zero\_bit, sps\_vui\_alignment\_zero\_bit, rbsp\_stop\_one\_bit, rbsp\_alignment\_zero\_bit, gci\_alignment\_zero\_bit, sli\_alignment\_zero\_bit. But not consistently, e.g., we use u(1) for sn\_zero\_bit (this is the only exception we found in VVC).

In VSEI, currently, u(1) is used for nnpfc\_alignment\_zero\_bit\_a, nnpfc\_alignment\_zero\_bit\_b (as far as we know these are the only two cases of fix-valued, non-reserved bits in VSEI v3, none in v1 or v2).

Would it be better to consistently use f(1) for all fix-valued, non-reserved bits? The same approach could apply for all fix-valued, non-reserved bytes, two-bytes, etc.

* The colour transform information (CTI) SEI message in VSEI is for replacement of the colour remapping information (CRI) SEI message in HEVC. However, while a new CRI SEI message with a particular colour\_remap\_id value does not cancel the persistence of a previous CRI SEI message with a different colour\_remap\_id value, a new CTI SEI message with a particular colour\_transform\_id value does cancel the persistence a previous CTI SEI message with a different colour\_transform\_id value.

The authors don’t remember that a change of this behavior was discussed when the CTI SEI message was adopted, and we suggest the change may have been unintentional.

Is this a bug?

* In subclause 7.3.11.10 (Transform unit syntax) of VVC, there is an instance of “if(IntraSubPartitionsSplitType != ISP\_NO\_SPLIT )”, where a space is missing after ‘(’.

The item of CTI SEI is the only one with technical impact. It was commented that there might be a problem with splicing when no cancellation would be done.

All other aspects are editorial.

Include in JVET-AF1004.

## AHG3: Test conditions (0)

This section is kept as a template for future use.

## AHG3: Software development (0)

This section is kept as a template for future use.

## AHG4: Subjective quality testing and verification testing (4)

Contributions in this area were discussed at 1130–XXXX on Wednesday 18 Oct. 2023 in a joint meeting with WG 4, AG 5 and VCEG (chaired by JRO and MW) – see notes in section xx.

[JVET-AF0044](https://jvet-experts.org/doc_end_user/current_document.php?id=13288) AHG4: Report on AhG meeting on verification testing for VVC multilayer coding [M. Wien]

No need for presentation – included in JVET-AF0004.

[JVET-AF0075](https://jvet-experts.org/doc_end_user/current_document.php?id=13323) Evaluation results of Low Complexity Enhancement Video Codec (LCEVC) with HM and VTM on 4K content [O. Chubach, Y.-L. Hsiao, C.-Y. Chen, C.-W. Hsu, T.-D. Chuang, Y.-W. Chen, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-AF0224](https://jvet-experts.org/doc_end_user/current_document.php?id=13487) Crosscheck of JVET-AF0075 (Evaluation results of LCEVC) [C. Lehmann (HHI)] [late]

[JVET-AF0228](https://jvet-experts.org/doc_end_user/current_document.php?id=13491) Crosscheck of JVET-AF0075 (Evaluation results of Low Complexity Enhancement Video Codec with HM and VTM on 4K content) [P. Chen, M. Zhou, W. Wan (Broadcom)] [late]

[JVET-AF0277](https://jvet-experts.org/doc_end_user/current_document.php?id=13541) Crosscheck of JVET-AF0075 (Evaluation results of Low Complexity Enhancement Video Codec (LCEVC) with HM and VTM on 4K content) [V. Seregin (Qualcomm)] [late]

[JVET-AF0292](https://jvet-experts.org/doc_end_user/current_document.php?id=13556) Considerations on LCEVC performances in relation to input document JVET-AF0075/m65560 [S. Ferrara, X. Ciccarelli (V-Nova)] [late]

[JVET-AF0186](https://jvet-experts.org/doc_end_user/current_document.php?id=13444) AHG4: experiments related to VVC spatial scalability visual testing [P. de Lagrange, F. Urban (InterDigital)]

[JVET-AF0311](https://jvet-experts.org/doc_end_user/current_document.php?id=13575) Results of expert viewing at AF meeting for the spatial scalability category of the VVC multilayer VT M. Wien (RWTH Aachen University)

TBP

See section 7.4 for the initial discussion in a joint session.This document reports the results of an on-site expert viewing performed at the 32nd JVET meeting in Hannover, DE.

The evaluation targets at the spatial scalability category of the test plan document JVET-AE2021, comparing the compression performance of dual layer (DL) coding with a downscaled base layer at scaling ratio 2 to single layer coding at the low resolution with upscaling to the target resolution using the VVC interpolation filters (UP) as well as single layer coding at the target resolution (SL).

A number of 19 experts participated in the expert viewing. With a comparably low number of valid scores, the results should be considered as an indication of tendencies. They are not regarded as a necessarily conclusive result.

With respect to the goals of this experiment, the following observations are made:

* Since the UP sequences represent the extracted base layer of the DL streams, the impact of the enhancement layer can be studied. It can be seen that in most cases, the MOS only slightly increases with the inclusion of the enhancement layer. This might be attributed to the fact that other effects overlay the benefit of enhanced detail provided with the DL stream.
* The result show a similar relation of the SL, DL, and UP results compared to JVET-AE0288 for the sequences CorsaireJoy. For TallBuildings2, the results seem to indicate a visual improvement of the DL configuration compared to the UP configuration. This might be attributed to the encoder modifications reported in JVET-AF0186.
* The MOS results for the DrivingPOVLogo sequence seem to indicate that there might be a range of benefit for the DL scheme at rates at or below 1Mbit/s.
* The MOS results for the NxBoxeLogo sequence seem to indicate bitrate ranges where UP, DL, and SL have benefits, similar to what might be chosen in a bitrate ladder setting.

It is recommended to take these results into account by reconsiderering the test setup with respect to bitrate ranges to be tested for certain resolutions and encoder configurations. The test sequence CorsaireJoy might not include sufficient detail to provide a suitable demonstration case for the targeted application.

Initial version had been presented during joint meeting XXX, final version was presented in JVET on Friday Oct. 20 1115-1140.

It is notable that the upscaled base layer has a very high quality when presented to viewers without compressing it. This may suggest the conclusion that the ranges of rates where spatially scalable coding with conveying information in the enhancement layer need to be carefully selected.

It was however commented that the conclusion that the uncoded upsampled base layer would be almost transparent might be questionable, as it had been tested along with coded sequences with visible distortions. Another test methodology would be necessary to investigate whether it is transparent compared to the uncoded full resolution.

It was commented that in some cases of low rate, dual layer coding becomes better than single layer at full resolution. It was pointed out that when using RPR in an optimized this might change (as also upsampled base layer is better in those cases). Unavailability of an optimized encoder criterion might be a problem here.

However, the purpose of this verification test is not to demonstrate superiority of multi-layer coding over single layer, but rather demonstrating that multi-layer functionality works reasonably well with VVC.

It is concluded to take the results into account in developing the new version of the multi-layer VVC test plan JVET-AF2021.

## AHG4: Test material (2+1)

Contributions in this area were discussed at 1405–1440 on Wednesday 18 Oct. 2023 (chaired by JRO).

[JVET-AF0253](https://jvet-experts.org/doc_end_user/current_document.php?id=13516) [AHG-4] On the proposed gaming sequences from InterDigital and class gaming [S. Puri, T. Poirier, C. Bonnineau, I. Marzuki, R. Utida, E. Francois (InterDigital)] [late]

The contribution provides the update on the gaming sequences initially proposed in JVET-Y0041. The proposed sequences are submitted after fixing some of the issues raised in the past (e.g., flickering and object motion). As a reminder, these gaming sequences were created using Unity game engine and included depth and optical flow information. The compression results along with example bitstreams using VTM-11ecm10.0, ECM-10.0 are provided with this contribution.

In the past, it was agreed to create an “optional” class for gaming (Class “G”). Here, it is proposed to study the above gaming sequences as potential candidates to be added to Class “G”. It is also proposed to include ArenaOfValor from Class F in Class “G”.

The contribution further reports few issues noticed when encoding the proposed gaming sequences with ECM-10.0. First, the worst-case runtime for gaming sequence (HD) is much higher than worst-case runtime for Class B sequences (~19 days) in LDB Low Latency Controlled Complexity (LLCC) configuration. Second, for one gaming sequence, QP27 encoder runtime is higher than QP22. Third, the BD-rate vs decoder time trade-off in gaming sequence is worse than CTC when comparing VTM11ecm10 vs ECM-10.

To address the above issues and to possibly reduce the encoding/decoding time of gaming sequences further, it is suggested to study the low delay configuration and the impact of tools on/off on Class “G” possibly under a new AHG.

Currently, only 3 gaming class sequences are available (including one from class F). More would be needed for establishing a reasonable class.

Encoder run time with LB was even larger (>30 days)

It was commented that changing class F would be undesirable, and having a sequence in two classes is also inappropriate.

[JVET-AF0187](https://jvet-experts.org/doc_end_user/current_document.php?id=13445) Compression of Gaming Contents [Z. Lin, K. Cai, J. Sauer, Y. Zhao, E. Alshina (Huawei)]

With the wide adoption of cloud gaming, efficient compression of gaming content becomes more and more important. The characteristics of gaming content are quite different from those of the TGM sequences. Performance on several gaming sequences for ECM, VTM and HM are tested. This contribution proposes to extend test condition by adding gaming contents and develop specific compression solution for gaming contents.

It was commented that finding gaming sequences that are suitable for testing (including visual testing) appears difficult.

Establish AHG (J. Sauer, S. Puri, R. Chernyak)

Primary task would be to identify suitable sequences, and test conditions relevant for gaming.

It was also commented that defining coding tools specifically for this application would require strong evidence about benefit.

[JVET-AF0262](https://jvet-experts.org/doc_end_user/current_document.php?id=13526) New Film Grain Material based on a Ground Truth approach [D. Ugur, D. Podborski, A. M. Tourapis (Apple Inc)] [late]

See section 4.12

## Codec performance with alternative test materials (1)

Contributions in this area were discussed at 1440–XXXX on Wednesday 18 Oct. 2023 (chaired by JRO).

[JVET-AF0047](https://jvet-experts.org/doc_end_user/current_document.php?id=13294) [AHG4] ECM10.0 evaluation on V3C test content [S. Schwarz, M.M. Hannuksela (Nokia)]

This document provides performance evaluations of the latest ECM version on volumetric video-based coding (V3C) video test sequences. When compared to ECM8.0, ECM10.0 shows improvements for MPEG immersive video (MIV) content, whereas there is a moderate loss in RA conditions and a moderate gain in AI conditions for video-based point cloud coding (V-PCC) content. Comparing the BD-rate of ECM against VTM, a smaller bitrate reduction is obtained with V3C content than with the content classes of the common test conditions.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ECM10.0 over VTM22.0** | | | | |
| **RA10** | **BD-Rate [%]** | | |  |
| **Y** | **U** | **V** | **Occ Y** |
| **V-PCC texture** | -12.90% | -27.63% | -7.76% | -7.76% |
| **V-PCC geometry** | -13.94% |  |  | -6.66% |
| **MIV texture** | -16.08% | -32.52% | -32.52% |  |
| **MIV geometry** | -14.47% |  |  |  |
| **Average** | -14.53% | -30.56% | -30.56% | -7.21% |
| **AI10** | **BD-Rate [%]** | | |  |
| **Y** | **U** | **V** | **Occ Y** |
| **V-PCC texture** | -7.66% | -22.98% | -22.98% | -12.80% |
| **V-PCC geometry** | -4.20% |  |  | -2.68% |
| **MIV texture** | -11.57% | -31.28% | -31.28% |  |
| **MIV geometry** | -16.50% |  |  |  |
| **Average** | -10.79% | -27.96% | -27.96% | -7.74% |

It is noted that the gain of ECM vs. VTM is comparably lower for V-PCC content than for CTC classes. However, the same is the case in a comparison VTM vs. HM.

## AHG5: Conformance test development (0)

This section is kept as a template for future use.

## AHG7: ECM tool assessment (1)

Contributions in this area were discussed at 1610–1610 on Tuesday 17 Oct. 2023 (chaired by Y. Ye).

[JVET-AF0165](https://jvet-experts.org/doc_end_user/current_document.php?id=13423) Ahg7/Ahg12: On VTM-11ecm anchor low-delay test conditions [F. Le Léannec, S. Puri, E. François, K. Naser (InterDigital)]

See notes in section 5.2.5.

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

Contributions in this area were discussed at 1415–1510 on Tuesday 17 Oct. 2023 (chaired by JRO).

### Non-SEI contributions (2)

[JVET-AF0060](https://jvet-experts.org/doc_end_user/current_document.php?id=13308) AHG8: An exemplar software implementation for temporal resampling for VCM [S. Wang, J. Chen, Y. Ye (Alibaba), S. Wang (CityU)]

Sections on using temporal resampling to improve coding efficiency for video coding for machines have been included in JVET-AD2030 “Optimization of encoders and receiving systems for machine analysis of coded video content (draft 2)” at the conceptual level, but without providing any detail. This contribution proposes a temporal resampling software implementation for inclusion in JVET-AD2030. The proposed implementation can achieve 22.12%, 30.03% and 75.68% average Pareto BD-rate improvements for random access (RA), low delay (LD) and all intra (AI) configurations for SFU-HW dataset under the common test conditions for optimization of encoders and receiving systems for machine analysis of coded video content as defined in JVET-AE2031, when VTM version is 20.0 and temporal resampling ratio is 4. For TVD dataset, 45.39%, 35.50% and 73.72% average Pareto can be achieved for RA, LD and AI configurations, respectively.

When VTM version is 12.0, it is reported that the following performance could be achieved:

For object tracking

temporal resampling ratio 2: -30.32% (RA), -23.41% (LD)

temporal resampling ratio 4: -44.20% (RA), -34.38% (LD), -77.35% (AI)

For object detection

temporal resampling ratio 2: -19.80% (RA), -20.52% (LD), -52.79% (AI)

temporal resampling ratio 4: -21.34% (RA), -29.67% (LD), -75.05% (AI)

It was asked what the benefit of the frame interpolation actually is? How would it compare against a VTM where just the same frame skipping is done at the encoder, and the temporally subsampled video is fed into the machine vision task? How would it compare to a simpler conventional upsampling?

What is the complexity of the additional processing at the decoder? Roughly 20-30x runtime increase.

It was commented that in video surveillance frame dropping might not be appropriate.

It was pointed out that when temporal subsampling is done, the access period should still be retained as 1 second in RA case. In the experiments, the original period of 64 frames (which would be relevant for 60 fps) was used, but with a subsampling by factor 4 it should have been 16. This might significantly increase the rate.

[JVET-AF0308](https://jvet-experts.org/doc_end_user/current_document.php?id=13572) Crosscheck of JVET-AF0060 (AHG8: An exemplar software implementation for temporal resampling for VCM) [Qi Zhang (Peking University)] [late]

[JVET-AF0157](https://jvet-experts.org/doc_end_user/current_document.php?id=13415) AHG8: A temporal resampling algorithm [D. Ding, X. Zhao, Z. Liu, S. Liu (Tencent)]

In JVET-AE2030, a pre-processing method based on temporal resampling is proposed, which can be used to optimize video encoders and receiving systems for machine consumption. By skipping certain frames before encoding, the bit rate can be significantly reduced without a strong negative impact on the machine consumption performance. In this contribution, a temporal resampling algorithm is detailed and its software implementation is provided. Based on the common test conditions described in JVET-AE2031, it is reported that the following coding performance can be achieved:

For VTM 20:

For object tracking

2x temporal resampling rate: -51.02% (AI), -15.88% (LD), -8.32% (RA)

4x temporal resampling rate: -75.93% (AI), -30.53% (LD), -3.84% (RA)

For object detection

2x temporal resampling rate: -49.91% (AI), -13.18% (LD), -6.72% (RA)

4x temporal resampling rate: -69.53% (AI), 5.53% (LD), 31.91% (RA)

For VTM 12:

For object tracking

2x temporal resampling rate: -50.69% (AI), -9.11% (LD), -4.04% (RA)

4x temporal resampling rate: -75.99% (AI), -24.00% (LD), -3.69%(RA)

For object detection

2x temporal resampling rate: -50.50% (AI), -7.76% (LD), 4.94% (RA)

4x temporal resampling rate: -69.61% (AI), 10.69% (LD), 29.1% (RA)

In the V2 verison the results of RA, as well as the comparison with VTM 12 anchor are updated. In addition, it is also discussed about the configuration of RA to fair compare with anchor. It is shown that the solution is an encoder decision coding method.

In this V3 version, the VTM12 Object Track, RA results are updated

The results for RA are more realistic than for JVET-AF0060, as the RA period is kept as 1 s.

Complexity? Not exactly known, run on GPU.

It was commented that run time may not be sufficient for assessing complexity.

Comments from crosschecker:

Generally results matched. In case of LB, there is an outlier park scene with considerably worse results.

Further study was recommended for both JVET-AF0060 and JVET-AF0157

* To investigate the benefit of sophisticated temporal upsampling algorithms, by comparing them against a solution where the machine vision task is just operated at the downsampled sequence, or against a simpler interpolation method, e.g. based on optical flow.
* To better document the complexity, e.g. model size, kMAC/px, CPU runtime, etc.

General for AHG8:

The CTC requires a change on the aspect of RA period when subsampling is used.

[JVET-AF0241](https://jvet-experts.org/doc_end_user/current_document.php?id=13504) Cross-check of JVET-AF0157 (AHG8: A temporal resampling algorithm) [C. Hollmann (Ericsson)] [late]

[JVET-AF0309](https://jvet-experts.org/doc_end_user/current_document.php?id=13573) AHG8: Ground Truth data for SFU-HW dataset [S. Keating (Sony), R. Nguyen (Canon)] [late]

(include abstract)

For information – no need for presentation.

### Contributions related to SEI messages (see section 6.2) (3)

[JVET-AF0068](https://jvet-experts.org/doc_end_user/current_document.php?id=13316) AHG8/AHG9: Signalling of encoder preprocessing information [W. Jia, Y. Li, Y.-K. Wang, K. Zhang, L. Zhang (Bytedance)]

[JVET-AF0138](https://jvet-experts.org/doc_end_user/current_document.php?id=13396) AHG8/AHG9: Truncated bit depth support SEI messages [D. Ding, X. Zhao, S. Liu, G. Teniou, S. Wenger (Tencent)]

[JVET-AF0147](https://jvet-experts.org/doc_end_user/current_document.php?id=13405) AHG8/AHG9: On Picture Modality Type [J. Gao, H.-B. Teo, C.-S. Lim, K. Abe (Panasonic)]

## AHG10: Encoding algorithm optimization (3)

Contributions in this area were discussed at 1610–1750 on Tuesday 17 Oct. 2023 (chaired by Y. Ye).

[JVET-AF0089](https://jvet-experts.org/doc_end_user/current_document.php?id=13293) AhG10: Lagrange multiplier and QP adaptation at CTU-level for VVC [H. Guo, C. Zhu, L. Luo, J. Chen (UESTC), Y. Huo, Y. Liu (Transsion)]

This contribution proposes a Lagrange multiplier and QP adaptation at CTU-level for VVC. The proposed Lagrange multiplier and QP adaptation algorithm was implemented on VTM-21.2. Compared to the original VTM-21.2, the reported overall PSNR-Y, U, V BD-rate and {EncT, DecT} variations in low delay B, low delay P, random access with GOP size of 32, and random access with GOP size of 16 configurations are shown below, respectively.

Over VTM-21.2:

* LDB: -4.27%, -5.26%, -3.91% {101%, 100%}
* LDP: -3.64%, -4.32%, -2.49% {101%, 99%}
* RA (GOP 32): -0.26%, -0.05%, -0.01% {101%, 99%}
* RA (GOP 16): -0.25%, 0.17%, 0.18% {101%, 99%}

It was commented that VTM has a similar algorithm already, which could be enabled by turning on “bim” (from JVET-Y0077, disabled by default). The proponent commented that “bim” in VTM could only be applied when MCTF is turned on, because it requires motion information from the MCTF process. It was reported by the proponent of this contribution that the performance of “bim” in VTM is approximately 2% for both RA and LB.

This proposal relies on a light pre-encoding scheme. The proponent suggests that the proposed method can be combined with “bim”, especially for the RA config.

Currently there is no cross checker for this contribution.

The proponent reported that several variations of the distortion propagation factor (DPF) strategy have been experimented, and the one proposed in JVET-AF0089 provides the best performance especially for LDB.

It was commented that the sequence-by-sequence performance reported in the xls has large variations, though gains can be obtained for most sequences, loss of 7+% is observed for one sequence.

Several experts commented that the contribution is interesting, and further optimization could be applied to reduce loss, optimize the interaction with “bim”, etc.

It was suggested for the proponents to also examine the impact on subjective quality. recommended.

[JVET-AF0111](https://jvet-experts.org/doc_end_user/current_document.php?id=13358) AHG10 MTT split modes early termination [W. Ahmad, P. Wennersten, K. Andersson (Ericsson)]

This contribution presents a further study on the early termination method presented in JVET-AE0057 which was adopted for ECM and VTM. The contribution JVET-AE0057 was enabled for ECM CTC and to enable it in VTM CTC, it was requested to investigate the performance of the proposed scheme for different input bit depths and for low QP values.

Additionally, the proposal is also tested for large QP values and a modification is proposed in the threshold calculation to obtain a better trade-off between encoding time reduction and coding loss for a large range of QP values.

As in JVET-AE0057, the cost of coding a CU without splitting the CU at all is compared to a threshold. If the cost is higher than the threshold, MTT split will not be considered, which is claimed to reduce encoding runtime. The proposal is implemented on top of VTM 20.2 software and compared to VTM 20.2, the average BDR loss and runtimes reported in this contribution are as follows:

* Test 1.1 (VTM SDR, Proposed condition is applied to CTC QPs, i.e., 22, 27, 32, 37)
  + JVET-AF0111 (New Threshold)

Y 0.02%, U 0.00%, V -0.01%, EncTime 96.4%, DecTime 99.7% for the RA configuration,

Y 0.04%, U 0.12%, V 0.03%, EncTime 96.2%, DecTime 100.0% for the LDB configuration,

* + JVET-AE0057 (Old Threshold)

Y 0.02%, U 0.04%, V 0.05%, EncTime 95.8%, DecTime 100.1% for the RA configuration,

Y 0.05%, U 0.50%, V 0.19%, EncTime 95.5%, DecTime 99.8% for the LDB configuration,

* Test 1.2 (VTM SDR, Proposed condition is applied to low QP values, i.e., 12, 17, 19, 22)
  + JVET-AF0111 (New Threshold)

Y 0.06%, U 0.16%, V 0.16%, EncTime 90.7%, DecTime 100.1% for the RA configuration,

* + JVET-AE0057 (Old Threshold)

Y 0.06%, U 0.14%, V 0.14%, EncTime 91.3%, DecTime 99.9% for the RA configuration,

* Test 1.3 (VTM SDR, Proposed condition is applied to high QP values, i.e., 37,42,47,52)
  + JVET-AF0111 (New Threshold)

Y 0.00%, U -0.03%, V -0.07%, EncTime 99.7%, DecTime 99.7% for the RA configuration,

* + JVET-AE0057 (Old Threshold)

Y 0.24%, U 0.43%, V 0.38%, EncTime 96.6%, DecTime 99.9% for the RA configuration,

* Test 2 (VTM HDR, Proposed condition is applied to HDR CTC QPs, i.e., 22, 27, 32, 37)
  + JVET-AF0111 (New Threshold)

Y 0.01%, U 0.08%, V 0.09% in PSNR, 0.02% 0.12% 0.08% in wPSNR, EncTime 98.3%, DecTime 99.7% for the RA configuration,

* + JVET-AE0057 (Old Threshold)

Y 0.02%, U -0.04%, V 0.04% in PSNR, 0.02% -0.02% 0.09% in wPSNR, EncTime 98.1%, DecTime 100.0% for the RA configuration,

* Test 3 (VTM HBD, Proposed condition is applied to high bit depth normal QP range)
  + JVET-AF0111 (New Threshold)

Y 0.03%, U 0.09%, V 0.08% in PSNR, 0.01% 0.06% 012% in wPSNR, EncTime 97.8%, DecTime 100.1% for the RA configuration,

* + JVET-AE0057 (Old Threshold)

Y 0.03%, U 0.09%, V 0.14% in PSNR, 0.01% -0.02% 011% in wPSNR, EncTime 97.2%, DecTime 99.7% for the RA configuration,

* Test 4.1 (ECM SDR, Proposed condition is applied to CTC QPs, i.e., 22, 27, 32, 37)

Y -0.01%, U 0.00%, V -0.01%, EncTime 100.6%, DecTime 99.8% for the RA configuration,

* Test 4.2 (ECM SDR, Proposed condition is applied to high QP values, i.e., 37,42,47,52)

Y -0.41%, U -0.30%, V -0.40%, EncTime 102.9%, DecTime 99.8% for the RA configuration,

Compared to JVET-AE0057, this contribution proposes a new threshold for non-split cost of 64x64 luma block, with the main motivation being to reduce the coding loss for high QP ranges (37-52). Without the new threshold (i.e., if the existing threshold is used), there would be a 0.24% performance penalty with about 3-4% encoding time reduction for these QPs. With the new threshold, the proposed method would be effectively “turned off” for these higher QP ranges, resulting in little change in performance and encoding time.

The new threshold is more “aggressive” for lower QP values, in that it would result in more coding time reduction. But for these lower QP values, the new threshold wouldn’t incur more coding performance loss.

Comparing to the existing threshold in ECM, using the new threshold would result in 0.41% luma gain but with 3% encoding time increase for high QPs. Several experts commented that this is not desirable tradeoff.

However, it was also commented that a little encoder runtime increase for high QPs should be tolerable, as simulations are much faster for high QPs anyway, and that performance loss incurred by the existing threshold isn’t desirable.

It was also commented that the new threshold is just a refinement of an existing algorithm.commented that if we take the new threshold, then we should apply it to VTM and ECM, as well as VTM-11ecm.

Decision: adopt the new threshold in JVET-AF0111 to the next releases of ECM, VTM and VTM-11ecm.

Cross checker reported that results are matched for VTM, low QP, SDR. Cross check of other combination of QP ranges and SDR/HDR/HDP has not been performed.

The main reasons for not turning the proposed speedup algorithm on for VTM were: 1) VTM results were not complete, 2) it was desirable to check the algorithm’s behavior for low QP values and other types of content (high bit-depth, and HDR).

The results in this contribution provided answers to the above questions from last meeting.

It was commented that we should keep CTC of VTM and VTM-11ecm (which serves as the anchor for ECM) consistent.

Decision (CTC): enable JVET-AF0111 in VTM CTC. Also turn it on for VTM-11ecm.

[JVET-AF0247](https://jvet-experts.org/doc_end_user/current_document.php?id=13510) Cross-check of JVET-AF0111 (AHG10: MTT split modes early termination) [T. Dumas (Interdigital)] [late]

[JVET-AF0122](https://jvet-experts.org/doc_end_user/current_document.php?id=13291) AhG10: Lagrange multiplier optimization for chroma ALF and CCALF [S.-W. Xie, Y. Gao, M.-H. Jia, Y.-X. Bai, C. Huang, P. Wu (ZTE)]

This contribution reports further results for the RDO Optimization method in ALF, which was proposed in JVET-AD0136 and JVET-AE0123. Regarding the proposed method, the chroma Lagrange multiplier will be adjusted by a scaling factor when the luma ALF APS transmission flag is ON and the number of history chroma ALF APS is insufficient. With VTM-22.0 and ECM-10.0 as anchors and test platforms, the reported overall PSNR-Y, U, V BD-rate and {EncT, DecT} variations of this proposal are as follows:

On top of VTM-22.0:

* LDB: 0.05%, -0.80%, -1.18% {100%, 100%}
* LDP: 0.02%, -0.89%, -1.03% {100%, 101%}

On top of ECM-10.0:

* LDB: 0.04%, -1.89%, -1.60% {100.0%, 100.7%}
* LDP: 0.11%, -1.60%, -1.45% {100.2%, 100.3%}

During the discussion of JVET-AE0123, it was suggested that the multiplier might need to be adjusted for ECM compared to VTM. The proponent reported that experimentations had been conducted to adjust the multiplier, and results show that using the same value for VTM and ECM achieves the most balanced performance among luma and chroma components.

Cross checker reported that VTM results have been completed, but ECM results are still partial at the time of presesntation. VTM results and partial ECM results are matched with that provided by the proponent.

It was asked why this has no effect on RA config. This could be because for RA config parallel encoding is performed.

Whereas the performance for VTM is considered to be overall beneficial (the chroma gain seems to justify the small loss in luma), the ECM performance is less clear. This is also the first time we see results for ECM. Further tuning might provide more balanced performane for ECM.

Several experts supported including this into VTM, but keeping it turned off for CTC.

Decision (SW): adopt JVET-AF0122 to the next release of VTM, turned off by default.

[JVET-AF0304](https://jvet-experts.org/doc_end_user/current_document.php?id=13568) Crosscheck of JVET-AF0122 (AhG10: Lagrange multiplier optimization for chroma ALF and CCALF) [H. Zhang (Tencent)] [late]

## AHG13: Film grain synthesis (5+3)

Contributions in this area were discussed at 1615–1835 on Wednesday 18 Oct. 2023 (chaired by JRO).

[JVET-AF0142](https://jvet-experts.org/doc_end_user/current_document.php?id=13400) [AHG9] Local Film Grain Synthesis using Annotated Regions SEI [G. Teniou, S. Wenger (Tencent)]

Film Grain Synthesis (aka FGS) is a tool aimed at maintaining the original film grain while preserving the encoding efficiency due to its random nature and complexity. While being used to maintain the artistic intent or to mitigate visual artefacts, film grain is not necessarily applicable to the entire picture. This contribution proposes to address the film grain synthesis applied locally on the picture by using the annotated regions as a mask.

It is proposed to define labels for enabling/disabling FGS in certain regions.

The proposal is to send an ASCII string as payload in the annotated regions SEI (without changing its syntax or semantics). A change is necessary in FGC, expressing that, if such a label is detected, the film grain synthasis should be performed on the region or not.

It was commented by several experts that this would not be supported by existing devices, and implementation-wise is an additional processing step. The typical use case would be full-frame synthesis of FG.

If the approach would be applied such that it is generated full frame and then using a mask in overlaying, it might by OK, according to one expert.

It was commented that by using content layering the same functionality could also be achieved in legacy devices.

On the other hand, legacy devices may ignore it, anyway.

It was commented that the SEI message might simplify the synthesis when the properties of the two regions are very different.

The proposal would require a change of an existing SEI message, and also introduces an uncommon relationship between two SEI messages, deviating from the original purpose of the annotated regions SEI. Furthermore, there are other ways of achieving this functionality which are more consistent and have even additional flexibility (content layering, alpha map as per subsequent contribution).

No action.

[JVET-AF0144](https://jvet-experts.org/doc_end_user/current_document.php?id=13402) [AHG9] Adaptive Film grain synthesis using Alpha Channel Information SEI message [G. Teniou, S. Wenger (Tencent)]

Film Grain Synthesis (aka FGS) is a tool aimed at maintaining the original film grain while preserving the encoding efficiency due to its random nature and complexity. While being used to maintain the artistic intent or to mitigate visual artefacts, film grain is not seen relevant for the entire sequence or even within an image. This contribution proposes to address the film grain synthesis applied locally on a per image basis by using the auxiliary data as a mask.

The proposal would require a change of FGC and a new mode for aux pictures.

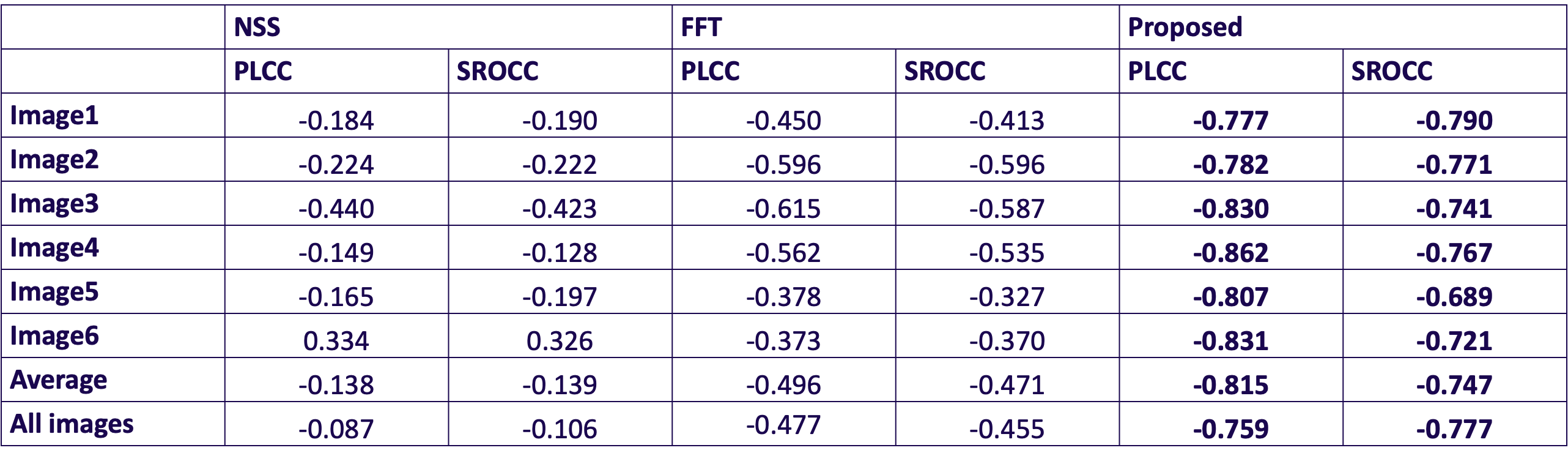
It was suggested that a better solution would be to define a new SEI message for this purpose that, if present, would point to a certain aux picture in alpha mode and combine it with the blending of film grain.

This would also be in line with not changing the operation that a legacy would execute on a previously existing SEI message. Also extensions of existing SEI messages are usually avoided.

The functionality as such is assessed to be useful. Proponents are asked to study the possibility of defining a new SEI message.

[JVET-AF0209](https://jvet-experts.org/doc_end_user/current_document.php?id=13472) AHG13: Frequency domain Film Grain Objective Metrics X. Meng, W. Zhang, S. Labrozzi (Disney Streaming) [late]

In this proposal, a subband based film grain assessment method (SFGA) is proposed, which shows promising results. SFGA, designed as a full-reference film grain quality assessment tool, initially transforms both reference and test images from the spatial domain to the frequency domain. It then selects several sub-bands within the frequency domain, compares their sub-band noise power spectra, and subsequently combines the sub-band results. The PLCC and SROCC between the proposed metric and MOS are both larger than 0.75, outperforming existing metrics like NSS (Natural Scene Statistics), FFT (Fast Fourier Transform), PSNR, and SSIM. This metric not only enhances our understanding of film grain technology, but also provides a valuable tool for fine-tuning film grain parameters in the modeling and synthesis process.



Subbands with 4 orientations and 4 scales.

Availability of objective metric would be highly desirable. This is a very attractive approach.

Software available? Some legal issues would need to be resolved, hopefully by next meeting. If data are generated by JVET, the analysis could be made by proponents.

Is the analysis done only in luma or in RGB domain? Only luma, even though the noise is coloured.

It was commented that evaluation when compression artifacts are coming in would be desirable

[JVET-AF0210](https://jvet-experts.org/doc_end_user/current_document.php?id=13473) AHG13: Exploration on the Capability of Frequency-Based Film Grain Synthesis [X. Meng, W. Zhang, S. Labrozzi (Disney Streaming)] [late]

The state-of-the-art film grain synthesis framework cannot meet the needs of high-quality film grain synthesis. This deficiency may primarily stem from limitations in the denoiser, film grain modelling, or film grain synthesis. To isolate the denoiser's impact, we captured clean images and introduced film grain using Davinci Resolve. Then we evaluated the subjective results of film grain modelling and synthesis. We can see that for finer-grained film grain it is challenging for the frequency model to generate film grain effectively. For coarser-grained film grain, the AR model faces difficulties in generating film grain effectively. To further confirm the bottleneck of the film grain framework, we explored the capability of frequency-based film grain synthesis by generating 16470 film grain models and evaluating the 16470 synthesis results. Experimental results show that,

1. In most cases, we can find a suitable film grain model for film grain synthesis, which means the capability of film grain SEI message and synthesis is sufficient.
2. In the current film grain synthesis framework, the film grain modeling highly constraints the performance of film grain synthesis.
3. The proposed metric SFGA in JVET-AF0209 can help to fine-tune the film grain models.

Synthesis parameters were derived from GT images, using SFGA metric.

As the metric is not perfect, would it be viable by this investigation to identify cases where the metric fails and could be improved.

It was suggested that by computing the metric in the RGB domain might be more appropriate.

It was also suggested that perhaps also looking at cases which are close to the best (in terms of the metric) might be useful.

Several experts expressed that this contribution provides very interesting information, also about the capability of the frequency domain method from FGC SEI.

[JVET-AF0262](https://jvet-experts.org/doc_end_user/current_document.php?id=13526) New Film Grain Material based on a Ground Truth approach [D. Ugur, D. Podborski, A. M. Tourapis (Apple Inc)] [late]

This document presents the film grain content material that was generated using the ground truth method discussed in contributions JVET-AD0369 and JVET-AE0250. The material consists of either newly captured content or scenes taken from the OpenMovie project “Tears of Steel”. A total of 26 noise variants is created for each scene selected, with the noise variants exhibiting different noise characteristics. It is suggested that the group study, and if desired use, such content in its different activities.

The extension of the number of variants is very helpful. Next step would be to select which ones are most relevant.

Coordination necessary for distribution of the content. Mathias/Alexis to sort out.

CC3 applies, copyright should be ok.

[JVET-AF0263](https://jvet-experts.org/doc_end_user/current_document.php?id=13527) Suggested process for measuring Grain Fidelity using the Ground Truth test set [A. M. Tourapis, J. Kim, S. Paluri, D. Podborski (Apple Inc)] [late]

This document presents a suggested framework/method of how to potentially measure film grain fidelity using the Ground Truth test-set presented in contribution JVET-AF0262 for the evaluation of proposals and technologies relating to film grain analysis or synthesis. The idea is based on techniques that have already been provided to the group, e.g. in contributions JVET‑X0048 and JVET­‑X0103, for performing film grain analysis in the context of estimating film grain parameters using the Film Grain characteristics SEI message. The framework only provides suggestions of measuring the fidelity of noise and not of the final content with the added noise. No actual implementation is provided.

The approach could be combined with using the metric of JVET-AF0209.

[JVET-AF0274](https://jvet-experts.org/doc_end_user/current_document.php?id=13538) AHG13/AHG3: film grain tuning tool [P. de Lagrange (InterDigital)] [late]

This contribution describes a new tool to graphically display and edit an FGC SEI configuration file, available in the JVET gitlab (in “JVET AHG Film Grain Technologies” group).

A recent addition is “FGC SEI designer”, a graphical interactive tool to edit an FCG SEI config file, that can be useful when automatic estimation is either not practical or needs to be refined. It can also be used to export a graphical representation of film grain parameters, for use e.g. in a document.

As it is implemented as a single python script, and supposed to be connected to a film grain synthesizer, it was just added to the “VFGS” software repository (https://vcgit.hhi.fraunhofer.de/jvet-ahg-fgt/vfgs), rather than creating a new repository. External dependencies are matplotlib and numpy.

The tool is also available from InterDigital github (https://github.com/InterDigitalInc/VersatileFilmGrain), which is in fact its primary source.

An impressive demonstration was given.

[JVET-AF0306](https://jvet-experts.org/doc_end_user/current_document.php?id=13570) Essentiality of manifest SEI for film grain [A. Tourapis, D. Podborski (Apple)] [late]

This document proposes a clarification regarding the significance of the SEI message for film grain. It offers a recommendation to use the SEI manifest SEI as well as the SEI prefix indication SEI message for signaling the essentiality of film grain synthesis. This recommendation is suggested to be incorporated into the technical report that JVET has been developing, which addresses practices for film grain technology.

Was presented on Thu Oct. 19 at 1215.

Agreed to be included in a future version of the TR.

## Implementation studies (1)

Contributions in this area were discussed at 1750–1800 on Thursday 19 Oct. 2023 (chaired by JRO).

[JVET-AF0211](https://jvet-experts.org/doc_end_user/current_document.php?id=13474) Latest update on Ali266: performance and deployment S. Fang, Z. Huang, S. Xu, L. Yu, J. Chen, R.-L. Liao, Y. Ye (Alibaba), X. Zhai, Y. Jia, D. Fan, Y. Zhang, C. Dou, X. Fu, F. Hu, R. Li (Youku) [late]

This contribution reports the latest update on performance and deployment status of Ali266, an optimized VVC software encoder and decoder developed by Alibaba. In JVET-W0127 and JVET-X0104, the Ali266 encoder was reported to respectively achieve 8-bit 720p and 1080p real-time encoding on a mainstream PC. In JVET-Y0066, the Ali266 encoder was further accelerated to achieve 10-bit 1080p real-time encoding.

In addition to our long-term efforts to improve coding efficiency and runtime performance, we are also committed to applying Ali266 to Alibaba’s various video businesses, starting with our collaborative effort with the Youku team as reported in JVET-Y0122.

This contribution reports the latest update on performance and deployment status of Ali266. Firstly, we report that compared to x265 (v3.5), the widely used open-source HEVC software encoder, Ali266 provides over 60% bitrate savings over x265’s veryslow preset at similar encoding speed for non-real-time applications, and over 40% bitrate savings for 1080p sequences over x265’s medium preset at similar encoding speed for real-time low-latency applications such as live streaming. The average encoding speed of 1080p sequences achieves 60 fps on a mainstream PC at the faster preset. Secondly, we report that Youku has implemented the playback of Ali266 content on more than 30 TV models, covering tens of thousands of devices. The cumulative number of playback times has exceeded millions.

## Profile/tier/level specification (1)

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

[JVET-AF0063](https://jvet-experts.org/doc_end_user/current_document.php?id=13311) On MV-HEVC profiles [Y.-K. Wang, H. Liu, L. Zhang, S. Jiao, C. Hu, J. Cui, G. Xu (Bytedance)]

See notes in 4.2.

## General aspects of standards development and applications of standards (1)

Contributions in this area were discussed at 1735–1745 on Thursday 19 Oct. 2023 (chaired by JRO).

[JVET-AF0022](https://jvet-experts.org/doc_end_user/current_document.php?id=13450) Summary of time spans for NB comment availability [G. J. Sullivan (SC 29 chair)]

To assist convenors and NBs in understanding basic aspects of the approval process, this document is intended to provide a high-level summary of the period of time between when an editor submits acceptable text for circulation and when the results of the NB comment period are available in ISO/IEC JTC 1/SC 29.

For information – short presentation for raising awareness.

# Low-level tool technology proposals (114)

## AHG11/AHG14: Neural network-based video coding (28+2)

### Summary, BoG reports, and information documents (2)

Contributions in this area were discussed at 0840–1125 on Saturday 14 Oct. 2023 (chaired by JRO).

[JVET-AF0023](https://jvet-experts.org/doc_end_user/current_document.php?id=13469) 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]

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

This round of EE1 tests included:

1. EE1-0.x: Unified LOP (based on [JVET-AE0281](https://jvet-experts.org/doc_end_user/current_document.php?id=13244))
2. EE1-1.1.x category: architectural changes for HOP (based on [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)).
3. EE1-1.2.x category: filter usage aspects (based on [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)).
4. EE1-2.x NN-Inter (based on [JVET-AE0112](https://jvet-experts.org/doc_end_user/current_document.php?id=13060) investigating complexity reduction possibilities).
5. EE1-3.x NN-Intra (based on [JVET-AE0144](https://jvet-experts.org/doc_end_user/current_document.php?id=13060), awaiting training cross-check).

All tests in EE1 were advised to use NNVC-6.0 as code base (unless it is not possible). The anchor for EE1 test is the default configuration of NNVC-6.0 as defined by AhG11/AhG14 (NN-intra and low complexity NN-filter enabled by default) in JVET-AE2016. Anchor performance and reference point for HOP NN-filters will be provided by AhG14.

Tests targeting to LOP replacement in NNVC (EE1-0) should report results vs NNVC-6.0 default configuration (LOP-1 filter and NN-Intra are enabled by default). Tests targeting improvement of HOP (EE1-1 category) should report results in comparison to NNVC-6.0 with HOP-1 (instead LOP-1) and NN-Intra enabled. HOP-1 architecture and reproducible training procedure description can be found in JVET-AE2019 and [readme.md](https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/VVCSoftware_VTM/-/blob/VTM-11.0_nnvc/training/training_scripts/NN_Filtering_HOP/readme.md).

NNVC has two NN-filters, so can be configures for Low Operation Point (LOP) and High Operation Point (HOP), in both NN-Intra is enabled. Performance of those two configurations relatively to VTM is summarized in Table 1. Run time is not always very reliable and accurate, but can give group extra information about complexity, so run time was added to the tables.

Table 1 Two configuration of NNVC tested vs VTM, NN-Intra enabled in both.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Test | Doc# | # Par M | kMAC /pxl | Random Access cfg. | | | | | All Intra cfg. | | | | |
| Y | U | V | Enc | Dec | Y | U | V | Enc | Dec |
| LOP | [JVET-AE0014](https://jvet-experts.org/doc_end_user/current_document.php?id=13090) | 0.2+1.3 | 17+8 | **-6.6%** | -7.7% | -7.3% | 1.3 | 76 | **-7.8%** | -8.6% | -8.9% | 2.0 | 45 |
| HOP | [JVET-AE0014](https://jvet-experts.org/doc_end_user/current_document.php?id=13090) | 1.5+1.3 | 477+8 | **-11.6%** | -25.7% | -26.0% | 1.8 | 476 | **-10.8%** | -22.2% | -23.4% | 2.1 | 269 |

By default in NNVC-6.0 NN-Intra and Low Operation Point (LOP) NN-filter are enabled. Proposals tested against NNVC-6.0 are listed in Table 2.

Table 2 Proposals tested against NNVC default configuration (LOP and NN-Intra are enabled)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Test | Doc# | # Par M | kMAC /pxl | Random Access cfg. | | | | | All Intra cfg. | | | | | |
| Y | U | V | Enc. | Dec. | Y | U | V | Enc. | Dec. |
| NNVC-6-LOP | [JVET-AE0014](https://jvet-experts.org/doc_end_user/current_document.php?id=13090) | 0.2 | 17 | **0.0%** | 0.0% | 0.0% | 100% | 100% | **0.0%** | 0.0% | 0.0% | 100% | 100% |
| Targeting replacement of LOP filter in NNVC | | | | | | | | | | | | | | |
| EE1-0 | [JVET-AF0043](https://jvet-experts.org/doc_end_user/current_document.php?id=13287) | 0.05 | 17 | **-0.3%** | -4.3% | -4.3% | 96% | 114% | **-0.2%** | -4.6% | -4.6% | 97% | 119% |
| Content adaptation with overfitting at encoder side | | | | | | | | | | | | | | |
| EE1-1.2.2 | [JVET-AF0056](https://jvet-experts.org/doc_end_user/current_document.php?id=13304) | 0.2 | 17 | **-1.2%** | -6.4% | -5.5% | 99% | 105% |  |  |  |  |  |
| NN-Inter (performance on top of LOP & NN-Intra) | | | | | | | | | | | | | | |
| EE1-2.1 | [JVET-AF0208](https://jvet-experts.org/doc_end_user/current_document.php?id=13471) | 3.8 | 504 | **-2.3%** | -4.2% | -3.8% | 133% | 1204% |  |  |  |  |  |
| Targeting replacement of NN-Intra filter in NNVC | | | | | | | | | | | | | | |
| EE1-3.1 | [JVET-AF0139](https://jvet-experts.org/doc_end_user/current_document.php?id=13397) | 1.3 | 4.8 | **0.2%** | 0.1% | 0.2% | 99% | 99% | **0.5%** | 0.6% | 0.7% | 95% | 94% |

Majority of EE1 tests are focusing on improvement of High Operation Point (HOP) complexity-performance trade-off. The natural anchor for those tests is NNVC-6.0 with HOP and NN-Intra enabled. Test results relatively to this anchor are summarized in Table 3.

Table 3 Proposals tested against NNVC- high configuration (HOP and NN-Intra are enabled)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Test | Doc# | #Par M | kMAC /pxl | Random Access cfg. | | | | | All Intra cfg. | | | | |
| Y | U | V | Enc. | Dec. | Y | U | V | Enc. | Dec. |
| NNVC-6-HOP | [JVET-AF0014](https://jvet-experts.org/doc_end_user/current_document.php?id=13385) | 1.45 | 477 | **0.0%** | 0.0% | 0.0% | 100% | 100% | **0.0%** | 0.0% | 0.0% | 100% | 100% |
| Group convolutions in BBB (Back Bone Block) | | | | | | | | | | | | | |
| EE1-1.1.1-t1 | [JVET-AF0153](https://jvet-experts.org/doc_end_user/current_document.php?id=13411) | 1.28 | 422 | **0.3%** | 2.0% | 1.5% |  |  | **0.1%** | 1.1% | 1.0% |  |  |
| EE1-1.1.4b1 | [JVET-AF0182](https://jvet-experts.org/doc_end_user/current_document.php?id=13440) | 1.23 | 406 | **0.3%** | 2.7% | 1.0% | 120% | 155% | **0.1%** | 0.7% | 0.7% | 111% | 157% |
| EE1-1.1.4b2 | [JVET-AF0182](https://jvet-experts.org/doc_end_user/current_document.php?id=13440) | 1.27 | 419 | **0.3%** | 1.1% | 1.2% | 124% | 172% | **0.2%** | 0.9% | 0.9% | 112% | 175% |
| EE1-1.1.4b3 | [JVET-AF0182](https://jvet-experts.org/doc_end_user/current_document.php?id=13440) | 1.05 | 349 | **0.6%** | 2.6% | 2.4% | 135% | 204% | **0.3%** | 1.3% | 1.1% | 118% | 210% |
| Number of channels reduction in BBB (Back Bone Block) | | | | | | | | | | | | | |
| EE1-1.1.2b | [JVET-AF0102](https://jvet-experts.org/doc_end_user/current_document.php?id=13349) | 1.3 | 426 | **0.1%** | 1.0% | 0.4% | 82% | 71% | **0.1%** | 0.2% | 0.4% | 83% | 66% |
| EE1-1.1.1-t2 | [JVET-AF0153](https://jvet-experts.org/doc_end_user/current_document.php?id=13411) | 1.35 | 442 | **0.3%** | 1.3% | 0.9% |  |  | **0.2%** | 0.8% | 1.1% |  |  |
| EE1-1.1.5 | [JVET-AF0192](https://jvet-experts.org/doc_end_user/current_document.php?id=13453) | 1.2 | 392 | **0.4%** | 1.7% | 1.2% | 72% | 56% | **0.2%** | 1.0% | 0.8% | 77% | 53% |
| “head” block simplifications | | | | | | | | | | | | | |
| EE1-1.1.4a | [JVET-AF0181](https://jvet-experts.org/doc_end_user/current_document.php?id=13439) | 1.45 | 457 | **0.2%** | 0.9% | 0.5% | 95% | 86% | **0.1%** | 0.2% | 0.4% | 98% | 89% |
| EE1-1.1.3 | [JVET-AF0103](https://jvet-experts.org/doc_end_user/current_document.php?id=13350) | 1.45 | 472 | **0.0%** | 1.0% | 0.3% | 98% | 96% | **0.1%** | 0.3% | 0.3% | 95% | 92% |
| Separation of NN-filters for Luma and Chroma | | | | | | | | | | | | | |
| EE1-1.1.4c | [JVET-AF0183](https://jvet-experts.org/doc_end_user/current_document.php?id=13441) | 1.42 | 473 | **-0.3%** | 1.5% | 1.9% | 129% | 202% | **-0.4%** | 0.3% | 0.5% | 112% | 206% |
| EE1-1.1.4d | [JVET-AF0183](https://jvet-experts.org/doc_end_user/current_document.php?id=13441) | 1.23 | 390 | **0.0%** | 1.8% | 2.0% | 137% | 221% | **-0.2%** | -0.1% | 0.2% | 116% | 232% |
| Filter usage aspects (also applicable to LOP) | | | | | | | | | | | | | |
| EE1-1.2.1 | [JVET-AF0085](https://jvet-experts.org/doc_end_user/current_document.php?id=13333) | 1.45 | 477 | **0.0%** | -0.2% | -0.5% | 100% | 101% | **0.0%** | -0.1% | 0.0% | 100% | 99% |
| EE1-1.2.3 | [JVET-AF0154](https://jvet-experts.org/doc_end_user/current_document.php?id=13412) | 1.45 | 477 | **-0.1%** | -0.6% | -0.8% | 97% | 99% |  |  |  |  |  |
| EE1-1.2.4 | [JVET-AF0205](https://jvet-experts.org/doc_end_user/current_document.php?id=13466) | 1.45 | 477 | **0.0%** | -0.1% | -0.1% | 100% | 99% | **0.0%** | -0.1% | -0.1% | 100% | 100% |
| EE1-1.2.6 | [JVET-AF0086](https://jvet-experts.org/doc_end_user/current_document.php?id=13334) | 1.45 | 477 | **-0.1%** | -0.5% | -0.7% | 98% | 99% |  |  |  |  |  |

Some proponents managed to complete LDB test, results are summarized in Table 4.

Table 4 Proposals tested against NNVC- high configuration (HOP and NN-Intra are enabled) in LDB configuration.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Test | Doc# | #Par M | kMAC /pxl | Low Delay-B cfg. | | | | |
| **Y** | U | V | Enc. | Dec. |
| NNVC-6-HOP | [JVET-AF0014](https://jvet-experts.org/doc_end_user/current_document.php?id=13385) | 1.45 | 477 | **0.0%** | 0.0% | 0.0% | 100% | 100% |
| EE1-1.1.1-t1 | [JVET-AF0153](https://jvet-experts.org/doc_end_user/current_document.php?id=13411) | 1.28 | 422 | **0.2%** | -0.1% | -6.2% |  |  |
| EE1-1.1.4b1 | [JVET-AF0182](https://jvet-experts.org/doc_end_user/current_document.php?id=13440) | 1.23 | 406 | **0.2%** | 0.0% | -5.8% | 121% | 157% |
| EE1-1.1.4b2 | [JVET-AF0182](https://jvet-experts.org/doc_end_user/current_document.php?id=13440) | 1.27 | 419 | **0.2%** | -0.4% | -7.4% | 127% | 174% |
| EE1-1.1.1-t2 | [JVET-AF0153](https://jvet-experts.org/doc_end_user/current_document.php?id=13411) | 1.35 | 442 | **0.1%** | -0.5% | -8.2% |  |  |
| EE1-1.1.5 | [JVET-AF0192](https://jvet-experts.org/doc_end_user/current_document.php?id=13453) | 1.2 | 392 | **0.3%** | -0.8% | -7.2% |  |  |
| EE1-1.1.4a | [JVET-AF0181](https://jvet-experts.org/doc_end_user/current_document.php?id=13439) | 1.45 | 457 | 0.3% | -1.6% | -8.4% | 94% | 85% |
| EE1-1.2.1 | [JVET-AF0085](https://jvet-experts.org/doc_end_user/current_document.php?id=13333) | 1.45 | 477 | **-0.3%** | -1.8% | -4.9% | 100% | 105% |
| EE1-1.2.3 | [JVET-AF0154](https://jvet-experts.org/doc_end_user/current_document.php?id=13412) | 1.45 | 477 | **-0.4%** | -2.6% | -2.2% | 98% | 100% |
| EE1-1.2.6 | [JVET-AF0086](https://jvet-experts.org/doc_end_user/current_document.php?id=13334) | 1.45 | 477 | **-0.4%** | -2.1% | -1.4% | 98% | 101% |

Observations of those test results are as follows:

* Reduction of kMAC/pxl not necessarily leads to reduction of encoding/decoding run-time (often causes opposite effect).
* Strangely almost all proposals tested in LDB cfg demonstrated significant gain in PSNR-V. Likely this indicates some problem with NNVC-HOP training or usage.

Observations on tests:

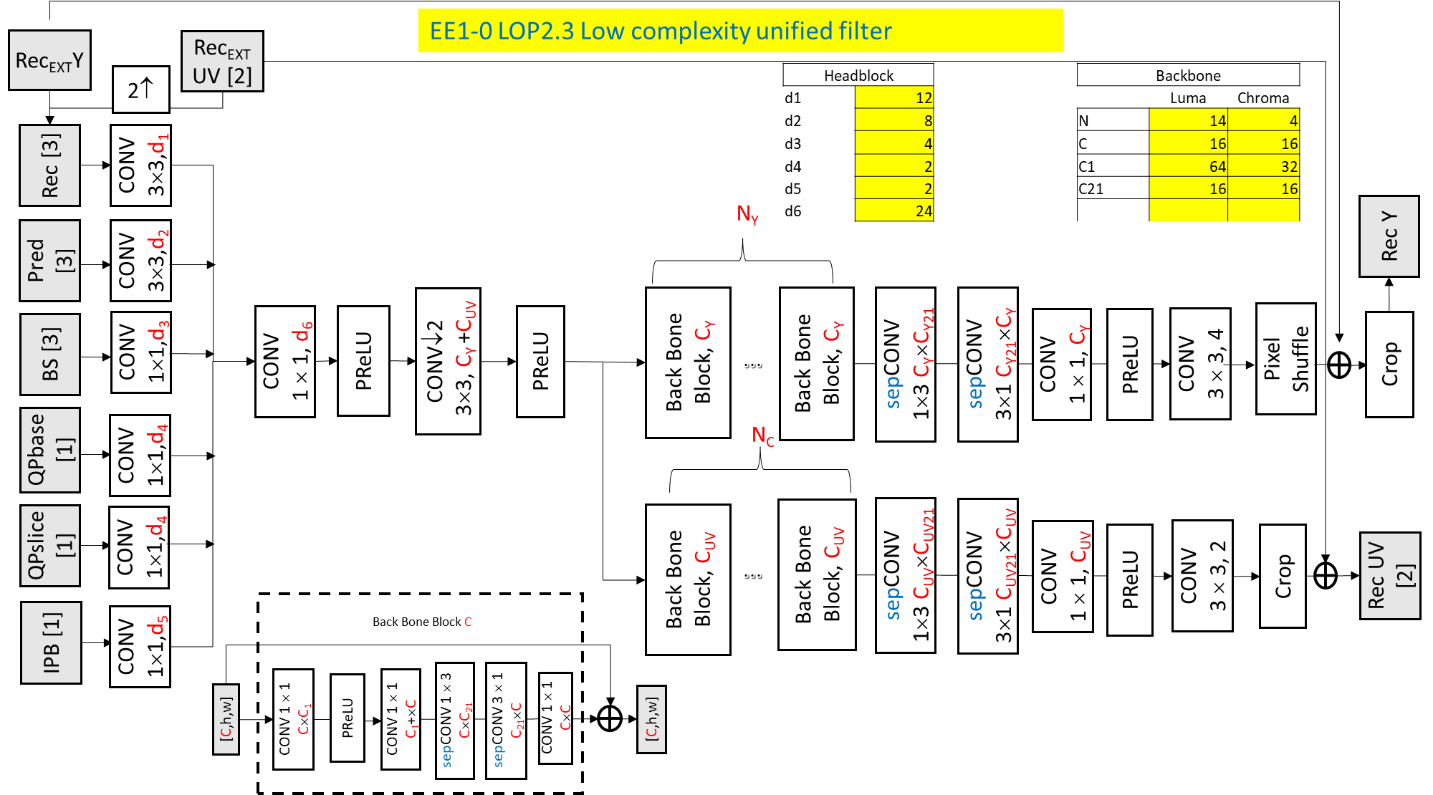
* Majority of proposed HOP NN-filter simplifications do not contradict to each other, potentially can be used together (the only exception is group convolution size, group should choose between group size 2 or 4).
* Among filter usage aspects (flipping, rotating, residual adjustment…) noticeable (0.3-0.4%) gain is observed in LDB cfg.

Mode details about EE1 tests can be found in presentation attached to this contribution.

**Tests description and results**

***EE1-0 tests 0.1 to 0.4***

Input contribution: [JVET-AF0043](https://jvet-experts.org/doc_end_user/current_document.php?id=13287) **AhG11/EE1: Status of the joint EE1-0 (LOP.2) training** [D. Rusanovskyy](mailto:dmytror@qti.qualcomm.com)**,**[Y. Li (Qualcomm)](mailto:yli30@qti.qualcomm.com)**,**[T. Shao](mailto:Tong.Shao@dolby.com)**,**[P. Yin (Dolby)](mailto:pyin@dolby.com)**,**[J. N. Shingala](mailto:jay.shingala@ittiam.com)**,**[A. Shyam](mailto:ajayshyam@ittiam.com)**,**[A. Suneja](mailto:ajat.suneja@ittiam.com)**,**[S. P. Badya (Ittiam)](mailto:siddarth.badya@ittiam.com)**,**[J. Li](mailto:lijunru@bytedance.com)**,**[Y. Li](mailto:yue.li@bytedance.com)**,**[C. Lin](mailto:linchaoyi.cy@bytedance.com)**,**[K. Zhang](mailto:zhangkai.video@bytedance.com)**,**[L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com)**,**[R. Chang](mailto:renjiechang@tencent.com)**,**[L. Wang](mailto:liqiangwang@tencent.com)**,**[X. Xu](mailto:xiaozhongxu@tencent.com)**,**[S. Liu (Tencent)](mailto:shanl@tencent.com)



*Figure* *1 Proposed unified architecture for LOP EE1 test*

Training and inference cross-checked since training was conducting by multiple parties: Qualcomm, Ittiam, Dolby (no separate documents needed).

After Stage 3 training, selected LOP 2.3 candidate model demonstrated BD-rate change of: {-4.7%, -9.8, -10.1%} and {-5.3%, -10.9%,-10.4%} vs VTM for AI and RA configurations, respectively.

Comparing to NNVC anchor (NN-Intra ON), LOP2.3 with enabled NN-Intra demonstrated BD-rate change of {-0.2%, -4.6%, -4.6%} and {-0.3%,-4.3%,-4.3%} for AI and RA, respectively. In this test, filter LOP2.3 utilizes a single model with 0.05M parameters, whereas the Anchor uses 4 models, total size of 0.2M.

Additional bd-rate gain of {-0.0%, -0.5%, -0.2%} and {-0.0%,-1.3%,-0.3%} for AI and RA, respectively, is reported in the sub-test EE1-0.4, targeting improved filter usage (interface).Aspects of common interest:

* finding in training strategy,
* layer quantization

Recommendation: to be presented in details as strong candidate for adoption and source of useful information.

***EE1-1.1.1***

Input contribution: [JVET-AF0153](https://jvet-experts.org/doc_end_user/current_document.php?id=13411) EE1-1.1.1: 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)

A diagram of a computer

Description automatically generated

Filter architecture: HOP-based

Training strategy: HOP

Training conducted: Tencent

Cross-check: Inference and training cross-check of test 1 are conducted, cross-checker: OPPO ([JVET-AF0266](https://jvet-experts.org/doc_end_user/current_document.php?id=13530)).

Notes: This test includes subtests targeting to optimize the complexity-performance trade-off of residual block structures and network architecture. The improved residual blocks (HOP residual block with depth-wise separable convolution and HOP residual block with group convolution) and split architecture from JVET-AE0160 are to be tested. This test may include the subtests by combining the elements from JVET-AE0160 and the subtests planned in test EE1-1.5 from JVET-AD2023.

***EE1-1.1.2***

Input contribution: [JVET-AF0102](https://jvet-experts.org/doc_end_user/current_document.php?id=13349) EE1-1.1.2 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)

A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated

Filter architecture: HOP-based.

Training strategy: HOP

Training conducted: Qualcomm

Cross-check: Inference and training supposed to be cross-checked by Byetedance, cross-check is missed.

This test targets to optimize the complexity-performance trade-off of the convolution decomposition and residual block structures. Reduced rank decomposition and decomposition order alternation of JVET-AE0164 to be tested. The test will include sub-tests with across architecture complexity-performance optimization.

***EE1-1.1.3***

Input contribution: [JVET-AF0103](https://jvet-experts.org/doc_end_user/current_document.php?id=13350) **EE1-1.1.3 Study on input feature set optimization,**[Y. Li](mailto:yli30@qti.qualcomm.com)**,**[M. Karczewicz (Qualcomm)](mailto:martak@qti.qualcomm.com)

A diagram of a block diagram

Description automatically generated

Filter architecture: HOP-based

Training strategy: HOP

Training conducted: Qualcomm

Cross-check: Training cross-checked by InterDigital ([JVET-AF0258](https://jvet-experts.org/doc_end_user/current_document.php?id=13521))

This test targets to optimize input feature set extraction, along methods of EE1-1.3.2 (JVET-AD0205). In particular, the headblock design optimization toward improving performance-complexity trade-off and hardware friendliness.

***EE1-1.1.4***

Input contribution(s):

[JVET-AF0181](https://jvet-experts.org/doc_end_user/current_document.php?id=13439) **EE1-1.1.4.a: Simplified feature extraction for HOP** [Y. Li](mailto:yue.li@bytedance.com)**,**[K. Zhang](mailto:zhangkai.video@bytedance.com)**,**[L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com)

[JVET-AF0182](https://jvet-experts.org/doc_end_user/current_document.php?id=13440) **EE1-1.1.4.b: Group convolution for HOP** [Y. Li](mailto:yue.li@bytedance.com)**,**[K. Zhang](mailto:zhangkai.video@bytedance.com)**,**[L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com)

[JVET-AF0183](https://jvet-experts.org/doc_end_user/current_document.php?id=13441) **EE1-1.1.4.c/d: Separate models for HOP** [Y. Li](mailto:yue.li@bytedance.com)**,**[K. Zhang](mailto:zhangkai.video@bytedance.com)**,**[L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com)

A diagram of a flowchart

Description automatically generatedA diagram of a system

Description automatically generated

Filter architecture: HOP-based.

Training strategy: HOP

Training conducted: Bytedance

This test was planned to include sub-tests targeting at optimizing the complexity-performance trade-off of the HOP model, e.g. separate models for luma and chroma (as described in EE1-1.1 in JVET-AD2023), network architecture optimization (as described in EE1-1.4 in JVET-AD2023), along methods in JVET-AD0106, JVET-AD0237.

Cross-check:

* 1.1.4.c/d: Separate luma/chroma model: InterDigital ([JVET-AF0260](https://jvet-experts.org/doc_end_user/current_document.php?id=13523))
* 1.1.4.a/b: Reduced complexity headblock and group convolution usage: Qualcomm ([JVET-AF0282 and JVET-AF0283)](https://jvet-experts.org/doc_end_user/current_document.php?id=13523)

***EE1-1.1.5: Combination of EE1-1.1.1 and EE1-1.1.2***

Input contribution: [JVET-AF0192](https://jvet-experts.org/doc_end_user/current_document.php?id=13453) **EE1-1.1.5: Combination test of EE1-1.1.1 and EE1-1.1.2** [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)**,**[D. Rusanovskyy](mailto:dmytror@qti.qualcomm.com)**,**[Y. Li](mailto:yli30@qti.qualcomm.com)**,**[M. Karczewicz (Qualcomm)](mailto:martak@qti.qualcomm.com)

A diagram of a block diagram

Description automatically generated

Filter architecture: HOP-based.

Training strategy: HOP

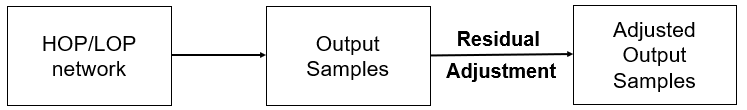
Training conducted: Qualcomm and Tencent

Cross-check: Inference and training cross-check are planned, cross-checker: OPPO ([JVET-AF0267](https://jvet-experts.org/doc_end_user/current_document.php?id=13531)).

This test study a direct combination of tests EE1-1.1.1 and EE1-1.1.2. Group convolution of EE1-1.1.1 to be introduced in the multi-scale branch of the Backbone Blocks (T1/T2) with a reduced rank decomposition of EE1-1.1.2 being used in the main branch of the T1/T2.

***EE1-1.2.1***

Input contribution: [JVET-AF0085](https://jvet-experts.org/doc_end_user/current_document.php?id=13333) EE1-1.2.1: On residual adjustments for NNLF [Z. Dai](mailto:daizhenyu@oppo.com), [Y. Yu](mailto:yue.yu@oppo.com), [H. Yu](mailto:v-yuhaoping@oppo.com), [D. Wang (OPPO)](mailto:wangdong7@oppo.com)



Filter architecture: HOP and LOP

Tester: OPPO

Cross-check of test: Inference cross-check is conducted, cross-checker: InterDigital([JVET-AF0254](https://jvet-experts.org/doc_end_user/current_document.php?id=13517)).

Notes: This test evaluates the residual offset adjustment and combination with chroma order adjustment on top of HOP/LOP. A frame level two-step residual adjustment method is proposed to adjust the output of the NNLF in the inference stage. Firstly, the residual is adjusted by reducing the magnitude of the residual at each pixel by a small offset value, and the offset candidates are {1, 2}. Secondly, the residual is decremented by the average residual value which is derived by computing the average of the residual for all pixels in each CTU. Also, a frame level chroma order adjustment method can be used to allow the input/output order switch between the U and V components of the neural network-based loop filters in the inference stage.

***EE1-1.2.2 Content-adaptive LOP filter***

Input contribution: [JVET-AF0056](https://jvet-experts.org/doc_end_user/current_document.php?id=13304) **EE1-1.2.2: Content-adaptive LOP filter** [R. Yang](mailto:ruiying.yang@nokia.com), [M. Santamaria](mailto:maria.santamaria_gomez@nokia.com), [F. Cricri](mailto:francesco.cricri@nokia.com), [H. Zhang](mailto:honglei.1.zhang@nokia.com), [J. Lainema](mailto:jani.lainema@nokia.com), [M. M. Hannuksela](mailto:miska.hannuksela@nokia.com), [A. Hallapuro (Nokia)](mailto:antti.hallapuro@nokia.com)

3x3 conv 3x3x10xM

Leaky ReLu

3x3 conv 3x3xKxL

1x1 conv 1x1xMxM

Leaky ReLu

3x3 conv 3x3xKxK

1x1 conv 1x1xMxK

1x1 conv 1x1xKxM

Leaky ReLu

3x3 conv 3x3xKxK

1x1 conv 1x1xMxK

BS info

Qstep

Yx4+U+V 72x72x10

NxN

1x1 conv 1x1xKxM

Leaky ReLu

3x3 conv 3x3xKxK

1x1 conv 1x1xMxK

Y’x4+U’+V‘ 64x64x6

N’xN’

***n*** Hidden Layers

1x1 conv 1x1xKxR

3x1 Sep conv 3x1xRxR

3x1xRxR

1x1 conv 1x1xRxK

1x3 Sep conv 1x3xRxR

CP decomposition

Layers approximated with CP decomposition

Separable Convolution layers

1x1 conv 1x1xMxR

1x1 conv 1x1xRxM

Filter architecture: NNVC 5.0 LOP with multiplier terms

Training strategy: no re-training, but the filter is overfitted on each test sequence, see JVET-AE0093.

Cross-check of overfitting: cross-checkers: OPPO([JVET-AF0216](https://jvet-experts.org/doc_end_user/current_document.php?id=13479)) & Ittiam([JVET-AF0119](https://jvet-experts.org/doc_end_user/current_document.php?id=13366))

Notes: Study content-adaptation on top of NNVC 5.0 LOP filter & signaling mechanism for the adaptation parameters.

***EE1-1.2.3***

Input contribution: [JVET-AF0154](https://jvet-experts.org/doc_end_user/current_document.php?id=13412) **EE1-1.2.3: Input and output rotation of model for NNVC in-loop filter**

[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)

A close-up of a white card

Description automatically generated

Filter architecture: HOP and LOP

Tester: Tencent

Cross-check of test: Inference cross-check is conducted, cross-checker: Ericsson([JVET-AF0257](https://jvet-experts.org/doc_end_user/current_document.php?id=13520)).

Notes: Rotation operations are applied to the inputs and outputs of NNVC HOP/LOP network to further improve the coding performance. All the input samples including reconstructed samples, prediction samples, BS, BaseQP, SliceQP and IPB are rotated and then fed into network to perform the inference. After that, the output of network is rotated back to restore the original order of filtered samples.

***EE1-1.2.4***

Input contribution: [JVET-AF0205](https://jvet-experts.org/doc_end_user/current_document.php?id=13466) EE1-1.2.4: An improved inference design of NN-based loop-filters at high operation point [J. Li](mailto:lijunru@bytedance.com), [Y. Li](mailto:yue.li@Bytedance.com), [C. Lin](mailto:linchaoyi.cy@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com)

Filter architecture: HOP-based.

Cross-check: Inference cross-check are planned, cross-checker TBA

Notes: This EE test study the unified filter usage of HOP and LOP, along methods in JVET-AE0238.

***EE1-1.2.6***

Input contribution: [JVET-AF0086](https://jvet-experts.org/doc_end_user/current_document.php?id=13334) **EE1-1.2.6: On flipping of input and output of model in NNVC HOP filter** [Z. Xie](mailto:xiezhihuang@oppo.com)**,**[Y. Yu](mailto:yue.yu@oppo.com)**,**[H. Yu](mailto:v-yuhaoping@oppo.com)**,**[D. Wang (OPPO)](mailto:wangdong7@oppo.com)



Filter architecture: HOP

Tester: OPPO

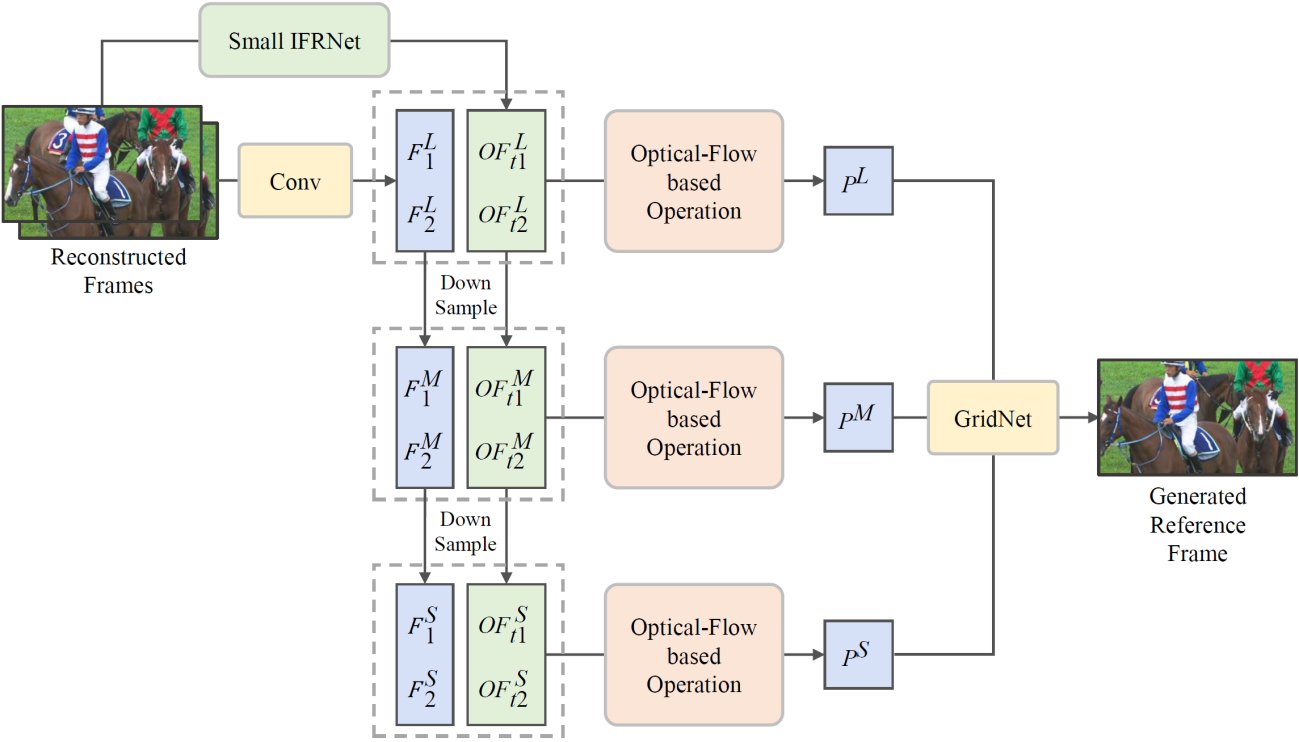
Cross-check of test: Inference cross-check is conducted, cross-checker: Tencent [JVET-AF0272](https://jvet-experts.org/doc_end_user/current_document.php?id=13536).

Notes: Flipping operations are applied to the inputs and outputs of NNVC HOP network to further improve the coding performance. All the input samples including reconstructed samples, prediction samples, BS, BaseQP, SliceQP and IPB are flipped horizontally or vertically, then fed into network to perform the inference. After inference, the output of network is flipped back to restore the original order of filtered samples.

***EE1-2.1***

Input contribution: [JVET-AF0208](https://jvet-experts.org/doc_end_user/current_document.php?id=13471) **EE1-2.1: Deep Reference Frame Generation for Inter Prediction Enhancemen** [W. Bao](mailto:baoweijie@whu.edu.cn)**,**[X. Chen](mailto:cinched@whu.edu.cn)**,**[J. Jia](mailto:jiajh2021@whu.edu.cn)**,**[Y. Zhang](mailto:yuantongzhang@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)

Filter architecture: Figure 2 illustrates the network architecture of the inter frame generation network, which is based on the model described in JVET-AE0112.



*The network architecture of the inter frame generation method*

Training strategy:The model can be trained using compressed BVI and TVD datasets or other datasets that can be used for NNVC standardization.The pictures in the dataset are compressed by VTM, with QP value randomly selected from 22, 27, 32, 37, and 42. The loss function is L1 loss, which is minimized using the Adam optimizer. A pre-trained small IFRNet network is used as our flow estimator, with a learning rate set at . The other parts of the network trained at a learning rate of . The training epochs are set to 200.

Training conducted: Wuhan University, Tencent.

Notes: This test aims to accelerate the float inference and implement integer inference. This test will analyze the compression performance of this method with LOP filter enabled, while also exploring the potential for reducing complexity.

Inference cross-check was conducted, cross-checker: OPPO ([JVET-AF0268](https://jvet-experts.org/doc_end_user/current_document.php?id=13532)).

Cross-checker(s) comments: Results match.

***EE1-3.1***

Input contribution: [JVET-AF0139](https://jvet-experts.org/doc_end_user/current_document.php?id=13397) EE1-3.1: 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 the table below. 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.

Diagram

Description automatically generated

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

|  |  |  |
| --- | --- | --- |
| **architecture** | **NN-Intra in common SW base** | **Proposal** |
| 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 . |

Training cross-check was conducted by Qualcomm (JVET-AF0281) and OPPO ([JVET-AF0217](https://jvet-experts.org/doc_end_user/current_document.php?id=13480))

Cross-checker(s) comments:

***Recommendations***:

* Present JVET-AF0043 (EE1-0) in details and adopt proposed LOP design to NNVC-7
* Agree on Next HOP design, combining elements from different proponents and train HOP jointly during next meeting period (BoG is needed)
* Adopt (EE1-3.1) and so reduce computational complexity of NN-Intra in NNVC

Remarks from presentation in JVET:

For EE1-1.1.4c/d (separate filters luma chroma), some uma gain, but chroma has losses 1.5%..2%, and the decoder run time more than doubles relative to NNVC HOP (may be an implementation issue? Group convolution in SADL?), encoder run time also increases. Other variants such as 1.1.2/1.1.5 have reductions in run time.

The loss in chroma may be an issue of training, or due to the fact that luma-to-chroma benefit is lost. Chroma has higher gain in NNVC HOP anyway, and balance is more difficult to control in single model.

New unified LOP architecture from LOP (EE1-0.x) also has separate branches for luma/chroma output, but giving gain for chroma relative to old LOP.

From comments of experts in the discussion, some more preference could be found for separation of luma and chroma filters, due to higher flexibility, better control in achieving luma/chroma balance, faster (parallel) training. On the other hand, it was argued that a single model is better to understand.

At the current meeting, there is anyway no candidate to change the HOP filter into separate models, as training crosscheck would still need to be performed for proposals available.

EE1-1.2.4 has slight improvement of HOP by modification at inference (boundary processing), but crosscheck not finished. Was later adopted – see BoG report JVET-AF0308.

EE1-1.2.2 (content adaptive LOP): It was suggested to provide the option of sending filter parameters once per RA period (rather than once per sequence as done currently)

EE1-2.1 (deep reference frame generation): Only inference cross-checked so far – training cross-check to be conducted towards next EE, also combination with HOP filter (gain is shown over NNVC6.0 which has LOP and intra pred. in it). Training was done based on NNVC5 reference picture output (which should be similar to NNVC6).

EE1-3.1 (reduced complexity intra pred.): None of the two cross-checks was finished yet. The training consists of 4 stages, and the script may not be simple to understand, according to cross-checkers. The training cross-check could not be finished before the meeting ended. In case that the training crosscheck would be successfully confirmed until Nov. 10 the AHG11 which is planned to meet on Nov. 16 is tasked to consider and recommend an adoption to NNVC7.1 SW (non CTC) .

It was also reported verbally by proponents that by an extended training the performance of the reduced-complexity intra prediction couldbe improved (no loss relative to current version).

It was also commented that at the current stage of exploration, there is no urgency of including lower complexity intra pred.

BoG (E. Alshina, F. Galpin) on assessment of HOP proposals, starting from Monday.

Status of EE proposals:

* Suggest potential adoptions for HOP in NNVC7 (likely only EE1-1.2.4)
* Identify promising approaches for which a training crosscheck should be performed in the next EE
* Identify promising new proposals on HOP for next EE
* Discuss aspects of luma/chroma balance, and aspects of single filter vs. separate luma/chroma filters.

[JVET-AF0042](https://jvet-experts.org/doc_end_user/current_document.php?id=13286) AhG11/AhG14 teleconference [E. Alshina, F. Galpin]

### EE1 contributions: Neural network-based video coding (16)

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

[JVET-AF0041](https://jvet-experts.org/doc_end_user/current_document.php?id=13285) AhG11: HOP full results [F. Galpin (InterDigital), D. Rusanovskyy, Y. Li (Qualcomm), Y. Li, J. Li (Bytedance), L. Wang, R. Chang (Tencent), Z. Xie (Oppo), E. Alshina (Huawei)]

[JVET-AF0043](https://jvet-experts.org/doc_end_user/current_document.php?id=13287) AhG11/EE1: Status of the joint EE1-0 (LOP.2) training [D. Rusanovskyy, Y. Li (Qualcomm), J. N. Shingala, A. Shyam, A. Suneja, S. P. Badya (Ittiam), T. Shao, P. Yin (Dolby)]

A Unified Filter Architecture for the Low-performance Operation Point (LOP) was proposed in JVET-AE0281 to meet LOP complexity constraints (LOP.2), following HOP based training process. This contribution reports on the progress of LOP.2 joint training and presents full set of LOP2.0, LOP2.3 results. Information presented in this document was partly presented during the AhG11/14 Telcos on 08/09/2023 and 08/30/2023.

After Stage 3 training, selected LOP 2.3 candidate model demonstrated BD-rate change of: {-4.7%, -9.8, -10.1%} and {-5.3%, -10.9%,-10.4%} vs VTM for AI and RA configurations, respectively. Comparing to NNVC anchor (NN-Intra ON), LOP2.3 with enabled NN-Intra demonstrated BD-rate change of {-0.2%, -4.6%, -4.6%} and {-0.3%,-4.3%,-4.3%} for AI and RA, respectively. In this test, filter LOP2.3 utilizes a single model with 0.05M parameters, whereas the Anchor uses 4 models, total size of 0.2M.

Additional bd-rate gain of {-0.0%, -0.5%, -0.2%} and {-0.0%,-1.3%,-0.3%} for AI and RA, respectively, is reported in the sub-test EE1-0.4, targeting improved filter usage (interface).

It is suggested to adopt filter of EE1-0 into NNVC common software and enable it by default for LOP. It is also suggested to make training strategy defined for EE1-0 a part of NNVC SW. Some aspects of LOP model training, such as training batch size, complexity-performance trade-off optimization and fast Stage 3 training can be further studied in EE1.

Was presented in JVET Saturday 14 Oct., 1125

It was asked if a better luma/chroma balance could be achieved (currently, higher gain in chroma). According to proponents likely yes, pointing to some contributions on this topic.

It is noted that the boundary padding (similar to EE1-1.2.4 for HOP) is already part of the package. It was asked if zero padding (instead of boundary sample values padding) might have subjective impact?

Influence of parameters such as batch size, learning rate? Some of these are selected different than for HOP.

Decision: Adopt JVET-AF0043 LOP 2.3 architecture, and inference interface. This will become part of NNVC7.0 anchor. Configuration files for training shall also become part of NNVC7.0.

Further study in EE: Fast stage 3 training as proposed in JVET-AF0043, influence of parameters such as batch size, learning rate, improved luma/chroma balance (bits from C to L), continuation of previous subtest EE1-0.2.

[JVET-AF0056](https://jvet-experts.org/doc_end_user/current_document.php?id=13304) EE1-1.2.2: Content-adaptive LOP filter [R. Yang, M. Santamaria, F. Cricri, H. Zhang, J. Lainema, M. M. Hannuksela, A. Hallapuro (Nokia)]

This contribution reports the results for the EE1-1.2.2 test which studies an integrated signalling mechanism for the content-adaptive neural network loop-filter in JVET-AE0093, and uses as basis the LOP1 loop-filter. It is reported that a new type of adaptation parameter set (APS), called neural network filter update APS, was designed and implemented. In JVET-AE0093, the inference was conducted on top of NNVC 5.0, using SADL library with half rounding enabled. The attained coding gains were -1.21% (Y), -6.43% (Cb) and -5.52% (Cr). Nevertheless, it was observed that the deactivation of half rounding in SADL library led to a reduction in coding gains by approximately 1%. Therefore, this contribution also studies data-driven quantization as means to improve the inference performance with SADL library and int16 precision. It is reported that, the average coding gains of this approach are -1.19% (Y), -6.40% (Cb) and -5.49% (Cr) with respect to the NNVC 5.0 anchor (NN intra tool ON and LOP1 tool ON). Furthermore, a simulation of signalling the update for each intra frame was performed and it is reported that the average coding gains are -0.34% (Y), -5.52% (Cb) and -4.75% (Cr) with respect to the NNVC 5.0 anchor.

From EE report: Continue EE, see under JVET-AF0023

Was further presented Tue 17 Oct. 1330 (chaired by JRO)

In the new version, signalling was enabled for each RA points (I picture), which caused a drop of gain down to 0.3% due to additional bit overhead. However, the same filter parameters were signalled for each I period, trained for the entire sequence. The more realistic scenario would be training for each I period separately, which would probably improve the performance.

Investigate in EE.

New LOP should be used, training crosscheck only for chunks, not the base model.

[JVET-AF0119](https://jvet-experts.org/doc_end_user/current_document.php?id=13366) Crosscheck of JVET-AF0056 (EE1-1.2.2: Content-adaptive LOP filter) [J. N. Shingala, A. Shyam, S. P. Badya (Ittiam)] [late]

[JVET-AF0216](https://jvet-experts.org/doc_end_user/current_document.php?id=13479) Crosscheck of JVET-AF0056 (EE1-1.2.2: Content-adaptive LOP filter) [Z. Dai (OPPO)] [late]

[JVET-AF0085](https://jvet-experts.org/doc_end_user/current_document.php?id=13333) EE1-1.2.1: On residual adjustments for NNLF [Z. Dai, Y. Yu, H. Yu, D. Wang (OPPO)]

[JVET-AF0254](https://jvet-experts.org/doc_end_user/current_document.php?id=13517) Cross-check of JVET-AF0085 (EE1-1.2.1: on residual adjustments of NNLF) [T. Dumas (Interdigital)] [late]

[JVET-AF0086](https://jvet-experts.org/doc_end_user/current_document.php?id=13334) EE1-1.2.6: On flipping of input and output of model in NNVC HOP filter [Z. Xie, Y. Yu, H. Yu, D. Wang (OPPO)]

[JVET-AF0272](https://jvet-experts.org/doc_end_user/current_document.php?id=13536) Crosscheck of JVET-AF0086 (EE1-1.2.6: On flipping of input and output of model in NNVC HOP filter) [R. Chang (Tencent)] [late]

[JVET-AF0102](https://jvet-experts.org/doc_end_user/current_document.php?id=13349) EE1-1.1.2 Complexity-performance tradeoff of decomposition [D. Rusanovskyy, Y. Li, M. Karczewicz (Qualcomm)]

[JVET-AF0285](https://jvet-experts.org/doc_end_user/current_document.php?id=13549) Crosscheck of JVET-AF0102 (EE1-1.1.2 Complexity-performance trade off of decomposition) [Y. Li (Bytedance)] [late]

[JVET-AF0103](https://jvet-experts.org/doc_end_user/current_document.php?id=13350) EE1-1.1.3 Study on input feature set optimization [D. Rusanovskyy, Y. Li, M. Karczewicz (Qualcomm)]

[JVET-AF0258](https://jvet-experts.org/doc_end_user/current_document.php?id=13521) Crosscheck of JVET-AF0103 (EE1-1.1.3 Study on input feature set optimization) [F. Galpin (InterDigital)] [late]

[JVET-AF0139](https://jvet-experts.org/doc_end_user/current_document.php?id=13397) EE1-3.1: neural network-based intra prediction with reduced complexity [T. Dumas, F. Galpin, P. Bordes (InterDigital)]

[JVET-AF0217](https://jvet-experts.org/doc_end_user/current_document.php?id=13480) Crosscheck of JVET-AF0139 (EE1-3.1: neural network-based intra prediction with reduced complexity) [Z. Dai (OPPO)] [late]

[JVET-AF0281](https://jvet-experts.org/doc_end_user/current_document.php?id=13545) Crosscheck of JVET-AF0139 (EE1-3.1: neural network-based intra prediction with reduced complexity) [P. Garus, D. Rusanovskyy (Qualcomm)] [late]

[JVET-AF0153](https://jvet-experts.org/doc_end_user/current_document.php?id=13411) EE1-1.1.1: Optimization for complexity-performance trade-off of HOP network [R. Chang, L. Wang, X. Xu, S. Liu (Tencent)]

[JVET-AF0266](https://jvet-experts.org/doc_end_user/current_document.php?id=13530) Crosscheck of JVET-AF0153 (EE1-1.1.1: Optimization for complexity-performance trade-off of HOP network) test1 [Z. Xie (OPPO)] [late]

[JVET-AF0154](https://jvet-experts.org/doc_end_user/current_document.php?id=13412) EE1-1.2.3: Input and output rotation of model for NNVC in-loop filter [R. Chang, L. Wang, X. Xu, S. Liu (Tencent)]

[JVET-AF0257](https://jvet-experts.org/doc_end_user/current_document.php?id=13520) Crosscheck of JVET-AF0154 (EE1-1.2.3: Input and output rotation of model for NNVC in-loop filter) [D. Liu (Ericsson)] [late]

[JVET-AF0181](https://jvet-experts.org/doc_end_user/current_document.php?id=13439) EE1-1.1.4.a: Simplified feature extraction for HOP [Y. Li, K. Zhang, L. Zhang (Bytedance)]

[JVET-AF0282](https://jvet-experts.org/doc_end_user/current_document.php?id=13546) Crosscheck of JVET-AF0181 (EE1-1.1.4.a: Simplified feature extraction for HOP) [D. Rusanovskyy (Qualcomm)] [late]

[JVET-AF0182](https://jvet-experts.org/doc_end_user/current_document.php?id=13440) EE1-1.1.4.b: Group convolution for HOP [Y. Li, K. Zhang, L. Zhang (Bytedance)]

[JVET-AF0283](https://jvet-experts.org/doc_end_user/current_document.php?id=13547) Crosscheck of JVET-AF0182 (EE1-1.1.4.b: Group convolution for HOP) [D. Rusanovskyy (Qualcomm)] [late]

[JVET-AF0183](https://jvet-experts.org/doc_end_user/current_document.php?id=13441) EE1-1.1.4.c/d: Separate models for HOP [Y. Li, K. Zhang, L. Zhang (Bytedance)]

[JVET-AF0260](https://jvet-experts.org/doc_end_user/current_document.php?id=13523) Crosscheck of JVET-AF0183 (EE1-1.1.4.c/d: Separate models for HOP) [F. Galpin (InterDigital)] [late]

[JVET-AF0192](https://jvet-experts.org/doc_end_user/current_document.php?id=13453) EE1-1.1.5: Combination test of EE1-1.1.1 and EE1-1.1.2 [R. Chang, L. Wang, X. Xu, S. Liu (Tencent), D. Rusanovskyy, Y. Li, M. Karczewicz (Qualcomm)]

[JVET-AF0267](https://jvet-experts.org/doc_end_user/current_document.php?id=13531) Crosscheck of JVET-AF0192 (EE1-1.1.5: Combination test of EE1-1.1.1 and EE1-1.1.2) [Z. Xie (OPPO)] [late]

[JVET-AF0205](https://jvet-experts.org/doc_end_user/current_document.php?id=13466) EE1-1.2.4: An improved inference design of NN-based loop-filters at high operation point [J. Li, Y. Li, C. Lin, K. Zhang, L. Zhang (Bytedance)]

[JVET-AF0294](https://jvet-experts.org/doc_end_user/current_document.php?id=13558) Crosscheck of JVET-AF0205 (EE1-1.2.4: An improved inference design of NN-based loop-filters at high operation point) [T. Shao (Dolby)] [late]

[JVET-AF0208](https://jvet-experts.org/doc_end_user/current_document.php?id=13471) EE1-2.1: Deep Reference Frame Generation for Inter Prediction Enhancement W. Bao, X. Chen, J. Jia, Y. Zhang, Z. Chen (Wuhan Univ.), Z. Liu, X. Xu, S. Liu (Tencent) [late]

[JVET-AF0268](https://jvet-experts.org/doc_end_user/current_document.php?id=13532) Crosscheck of JVET-AF0208 (EE1-2.1: Deep Reference Frame Generation for Inter Prediction Enhancement) [Z. Xie (OPPO)] [late]

### EE1 related contributions: Neural network-based video coding (1)

Contributions in this area were discussed at 1330–XXXX on Tuesday 17 Oct. 2023 (chaired by JRO).

[JVET-AF0071](https://jvet-experts.org/doc_end_user/current_document.php?id=13319) EE1-related: Further complexity reduction on the joint EE1-0 (LOP.2) unified filter [[T. Shao](mailto:tong.shao@dolby.com), [P. Yin](mailto:pyin@dolby.com), S. McCarthy (Dolby), [J. N. Shingala](mailto:jay.shingala@ittiam.com), [A. Shyam](mailto:ajayshyam@ittiam.com), A. Suneja, S. P. Badya (Ittiam)]

A Low-complexity Operation Point (LOP) in-loop filter (LOP1.0) proposed in JVET-AD0156, with complexity of 16.2 kMAC/pixel (block-wise), was adopted in JVET-AD meeting to the NNVC-5.0 software. LOP1.0 served as an anchor for further LOP development. A unified filter architecture for LOP (LOP2.x), was subsequently proposed in JVET-AE0281 to meet LOP complexity constraints while following the HOP training process. The detailed training and inference results of the unified LOP filter are reported in JVET-AF0043.

To further reduce the complexity of the unified LOP2 filter, this contribution proposes to adjust the skip connections in the backbone which enables the fusion of adjacent 1x1 convolutions. This skip-fusion design could reduce the complexity of the unified LOP2.3 from 17.05 kMac/Pixel to 16.56 kMac/Pixel and reduce the number of convolutional layers by 13%. The fast stage 3 training results show that the BD-Rate for fp32 is {0.00%, 0.71%, 0.99%} under AI and {-0.09%, 1.55%, 1.49%} under RA, while for int16 is {0.00%, 0.78%, 1.06%} under AI and {-0.08%, 1.76%, 2.04%} under RA, compared to the best LOP2.3 training results in JVET-AF0043. Compared to NNVC-6.0 VTM anchor, the BD-Rate for fp32 is {-4.77%, -9.14%, -9.13%} under AI and {-5.46%, -9.31%, -9.20%} under RA, while for int16 is {-4.74%, -9.09%, -9.17%} under AI and {-5.38%, -9.41%, -8.57%} under RA. The actual decoding time is reduced by about 5% compared to the LOP2.3.

Study in EE

[JVET-AF0284](https://jvet-experts.org/doc_end_user/current_document.php?id=13548) Crosscheck of JVET-AF0071 (EE1-related: Further complexity reduction on the joint EE1-0 (LOP.2) unified filter) [Y. Li, D. Rusanovskyy (Qualcomm)] [late]

### Improvements of NNVC software beyond EE1 (11)

Contributions in this area were discussed at 1240–1330 on Tuesday 17 Oct. 2023 (chaired by JRO), unless noted otherwise.

[JVET-AF0143](https://jvet-experts.org/doc_end_user/current_document.php?id=13401) AHG11: Unified CNN super resolution for resampling-based video coding [C. Lin, Y. Li, J. Li, K. Zhang, L. Zhang (Bytedance)]

The unified filter architecture for the in-loop filter has been adopted into NNVC. However, for super resolution (SR) technology, there is no corresponding unified SR model. This contribution proposes a unified SR architecture, which is designed to be aligned to the unified in-loop filter architecture and complexity. The experimental results are reported as:

For LOP SR (20.8 kMAC/pixel), the simulation results reportedly show {-3.86%, -1.19%, 0.56%} and {-3.20%, -6.15%, -0.11%} BD rate savings for {Y, Cb, Cr} under RA and AI configurations, respectively.

For HOP SR (469 kMAC/pixel), the simulation results reportedly show {-5.70%, -2.25%, -0.75%} and {-6.16%, -8.7%, -4.78%} BD rate savings for {Y, Cb, Cr} under RA and AI configurations, respectively.

Architecture similar to those of LOP/HOP loop filers. Performance of LOP worse than SR of current NNVC with significantly reduced complexity, HOP is better with half decrease in number of parameters.

Investigate in EE. Also test in combination with loop filters. It was also suggested to investigate whether the gain of the combination can be improved when SR is trained with LF output.

Further reduction of model size would be desirable. It was commented that other SR approaches are known with much less parameters than LOP.

[JVET-AF0150](https://jvet-experts.org/doc_end_user/current_document.php?id=13408) AhG11: on HOP batch size [F. Galpin, T. Dumas, P. Bordes (InterDigital)]

[JVET-AF0152](https://jvet-experts.org/doc_end_user/current_document.php?id=13410) AhG14: SADL update [F. Galpin, T. Dumas, P. Bordes, E. Francois (InterDigital)]

This contribution presents updates in the Small AdHoc Deep Learning (SADL) library. The update contains both some fixes and new features.

New layers

The following new layers were added (thanks to W. Bao for high quality MR):

- Grid-sample layer: float and integer support (e.g. support for motion compensation)

- Resize layer

Improvements and fixes

- Fix for slicing (support for LOP)

- Extend Conv2Dtranspose to new sizes

- Corrected leaky ReLU saturation for |alpha|>1

- Corrected sample programs

[JVET-AF0155](https://jvet-experts.org/doc_end_user/current_document.php?id=13413) AhG11: on HOP luma/chroma balance [F. Galpin, T. Dumas, P. Bordes (InterDigital)]

[JVET-AF0158](https://jvet-experts.org/doc_end_user/current_document.php?id=13416) AHG11: HOP In-loop filter with transformer blocks [Y. Li, D. Rusanovskyy, T. Ryder, S. Eadie, M. Karczewicz (Qualcomm)]

A Unified Filter Architecture for the High-performance Operation Point (HOP) was defined at JVET-AD0380. Training strategy and intermediate results were presented in JVET-AE0191 and complete inference results for Stage3 HOP training are reported in JVET-AF0041.

In this proposal, transformer blocks are integrated into the Residual Block (RB) of the HOP backbone. Proposed filter architecture was constrained to the complexity and memory requirements of the HOP architecture. Training was conducted by the HOP training guidelines.

Complexity of the model in kMAC/pixel: 478 (block-wise) with 1.45M parameters.

With a short training cycle (13 epochs), the proposed filter (floating model) provides: {-8.9%, -26.4%, -27.5%} for AI test configuration, and {-11.1%, -31.2%, -31.5%} for RA test configuration, versus the VTM. Compared to the HOP anchor (nnIntra OFF), the proposed filter (floating model) provides: {-1.2%, -9.4%, -9.0%} for AI configuration, and {-0.9%, -8.8%, -8.6%} for RA configuration.

Which kind of transformer? Using self attention, block transformer.

It was commented that transformers may have problems in providing bit exact behaviour as required for LF.

It was further commented that kMAC measure is not fully reflecting complexity, due to more complicated nonlinearities.

Block size of transformer if not the same as in remaining LF. Larger block size is used, such that attention mechanism is better supported. This has also impact on memory usage, which should be investigated as well.

Investigate in EE.

[JVET-AF0172](https://jvet-experts.org/doc_end_user/current_document.php?id=13430) AHG14: Cleanup on scale flag signalling for HOP [Y. Li, K. Zhang, L. Zhang (Bytedance)]

[JVET-AF0259](https://jvet-experts.org/doc_end_user/current_document.php?id=13522) Crosscheck of JVET-AF0172 (AHG14: Cleanup on scale flag signalling for HOP) [F. Galpin (InterDigital)] [late]

[JVET-AF0305](https://jvet-experts.org/doc_end_user/current_document.php?id=13569) Crosscheck of JVET-AF0172 (AHG14: Cleanup on scale flag signalling for HOP) [J. N. Shingala, A. Shyam (Ittiam)] [late]

[JVET-AF0313](https://jvet-experts.org/doc_end_user/current_document.php?id=13577) Crosscheck of JVET-AF0172 (AHG14: Cleanup on scale signaling for HOP) [Y. Li (Qualcomm)] [late]

[JVET-AF0180](https://jvet-experts.org/doc_end_user/current_document.php?id=13438) AHG11: HOP training adjustment to improve luma/chroma coding gain balance [D. Rusanovskyy, Y. Li, M. Karczewicz (Qualcomm)]

[JVET-AF0193](https://jvet-experts.org/doc_end_user/current_document.php?id=13454) AHG11: Decoder complexity optimization for NNVC in-loop filter [R. Chang, L. Wang, X. Xu, S. Liu (Tencent)]

[JVET-AF0256](https://jvet-experts.org/doc_end_user/current_document.php?id=13519) Crosscheck of JVET-AF0193 (AHG11: Decoder complexity optimization for NNVC in-loop filter) [D. Liu (Ericsson)] [late]

[JVET-AF0278](https://jvet-experts.org/doc_end_user/current_document.php?id=13542) Cross-check of JVET-AF0193 (AHG11: Decoder complexity optimization for NNVC in-loop filter) [M. Santamaria (Nokia)] [late]

[JVET-AF0206](https://jvet-experts.org/doc_end_user/current_document.php?id=13467) AHG11: Complexity reduction of NN-based loop-filters [J. Li, Y. Li, C. Lin, K. Zhang, L. Zhang (Bytedance)]

Objection was raised against presentation of this document on Wednesday Oct. 18, as the first version with reasonable description and results only was made available by Oct. 17.

This contribution proposes a unified filter architecture, with reported complexity of 5.0 and 7.3 kmac/pixel in two tests, which is even lower than 17.0 kmax/pixel of the lower operation point (LOP) in NNVC. The training and inference are kept same as the existing NN-based filters in NNVC.

Compared with NNVC-6.0 VTM anchor the BD-rate changes are shown as below reportedly:

Test #1: AI: -3.14%, -6.34%, -5.29%, RA: -3.09%, -6.22%, -4.04% (On top of NNVC-6.0 VTM anchor, 7.3 kmac/pixel)

Test #2: AI: -2.47%, -4.63%, -4.21%, RA: -1.96%, -3.98%, -2.93% (On top of NNVC-6.0 VTM anchor, 5.0 kmac/pixel)

Was presented Thu. Oct. 18, 1155.

It was commented that it is questionable whether even lower complexity operation point would be needed at this stage of the exploration, and whether otherwise this contribution provides more insight.

It was commented that RA has a significant drop compared to LOP, and that according to another experts’ experiments, better performance could be achieved.

Further study recommended.

[JVET-AF0236](https://jvet-experts.org/doc_end_user/current_document.php?id=13499) AHG14: The extension of SADL library [W. Bao, Y. Cai, Y. Zhang, Z. Chen (Wuhan Univ.)] [late]

This contribution presents the extensions in the Small AdHoc Deep Learning (SADL) library from Wuhan University. These extensions include layers such as Conv2DTranspose, GridSample, and Resize, along with a quantization test script. All of these extensions have been verified and merged.

[JVET-AF0296](https://jvet-experts.org/doc_end_user/current_document.php?id=13560) AhG11: on HOP learning rate [F. Galpin, T. Dumas, P. Bordes (InterDigital)] [late]

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

Section kept as a template for future use.

## AHG6/AHG12: Enhanced compression beyond VVC capability (78+1)

### Summary and BoG reports (1)

Contributions in this area were discussed at 1400–1915 on Friday 13 Oct. 2023 (chaired by JRO) except otherwise noted.

[JVET-AF0024](https://jvet-experts.org/doc_end_user/current_document.php?id=13470) 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]

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-10.0, released at <https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tags/ECM-10.0>. ECM-10.0 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-ae-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-ae-ee2/simulation-results> if cross-check reports are not submitted as they are optional for EE tests.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Tests** | **Tester** | **Cross-checker** |
| **1 Partitioning** | | | |
| 1.1a | Non-square quadtree partitioning | LGE  Y. Ahn  [JVET-AF0108](https://jvet-experts.org/doc_end_user/current_document.php?id=13355) | InterDigital  R. Utida |
| 1.1b | ECM with maximum MTT depth increments | LGE  Y. Ahn  [JVET-AF0108](https://jvet-experts.org/doc_end_user/current_document.php?id=13355) | InterDigital  R. Utida |
| **2 Intra prediction** | | | |
| 2.1a | DIMD merge | Nokia  S. Blasi  [JVET-AF0120](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0120-v1.zip) | Ofinno  P. Andrivon  [JVET-AF0124](https://jvet-experts.org/doc_end_user/current_document.php?id=13369) |
| 2.1b | DIMD merge with reduced storage | Nokia  S. Blasi  [JVET-AF0120](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0120-v1.zip) | Ofinno  P. Andrivon  [JVET-AF0124](https://jvet-experts.org/doc_end_user/current_document.php?id=13369) |
| 2.2a | DIMD with filtered template | vivo  C. Zhou  [JVET-AF0131](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0131-v1.zip) | Alibaba  J. Chen |
| 2.2b | DIMD with filtered template (without modification of gradient operators) | vivo  C. Zhou  [JVET-AF0131](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0131-v1.zip) | Alibaba  J. Chen |
| 2.3a | Test 2.1a + Test 2.2a | Nokia  S. Blasi  vivo  C. Zhou  [JVET-AF0126](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0126-v1.zip) | Bytedance  W. Yin  [JVET-AF0204](https://jvet-experts.org/doc_end_user/current_document.php?id=13465) |
| 2.3b | Test 2.1a + Test 2.2b | Nokia  S. Blasi  vivo  C. Zhou  [JVET-AF0126](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0126-v1.zip) | Bytedance  W. Yin  [JVET-AF0204](https://jvet-experts.org/doc_end_user/current_document.php?id=13465) |
| 2.3c | Test 2.1b + Test 2.2a | Nokia  S. Blasi  vivo  C. Zhou  [JVET-AF0126](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0126-v1.zip) | InterDigital  K. Naser |
| 2.3d | Test 2.1b + Test 2.2b | Nokia  S. Blasi  vivo  C. Zhou  [JVET-AF0126](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0126-v1.zip) | InterDigital  K. Naser |
| 2.4 | IntraCIIP as additional mode of IntraTMP | InterDigital  K. Naser  [JVET-AF0130](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0130-v1.zip) | OPPO  F. Wang  [JVET-AF0214](https://jvet-experts.org/doc_end_user/current_document.php?id=13477) |
| 2.5a | TIMD with IntraTMP/IBC | InterDigital  K. Naser  [JVET-AF0136](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0136-v1.zip) | Ofinno  P. Andrivon  [JVET-AF0125](https://jvet-experts.org/doc_end_user/current_document.php?id=13370) |
| 2.5b | Test 2.5a + Test 2.4 | InterDigital  K. Naser  [JVET-AF0136](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0136-v1.zip) | OPPO  F. Wang  [JVET-AF0215](https://jvet-experts.org/doc_end_user/current_document.php?id=13478)  Nokia  S. Blasi  [JVET-AF0243](https://jvet-experts.org/doc_end_user/current_document.php?id=13506) |
| 2.6a | Fractional-pel intraTMP BVs are stored in block vector buffer | OPPO  Y. Yu  Qualcomm  P.-H. Lin  [JVET-AF0079](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0079-v1.zip) | ETRI  W. Lim  [JVET-AF0227](https://jvet-experts.org/doc_end_user/current_document.php?id=13490) |
| 2.6b | IntraTMP BVs are stored for HMVP | OPPO  Y. Yu  Qualcomm  P.-H. Lin  [JVET-AF0079](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0079-v1.zip) | ETRI  W. Lim  [JVET-AF0227](https://jvet-experts.org/doc_end_user/current_document.php?id=13490) |
| 2.6c | Test 2.6a + Test 2.6b | OPPO  Y. Yu  Qualcomm  P.-H. Lin  [JVET-AF0079](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0079-v1.zip) | ETRI  W. Lim  [JVET-AF0227](https://jvet-experts.org/doc_end_user/current_document.php?id=13490) |
| 2.7 | Extrapolation filter-based intra prediction mode | OPPO  L. Xu  [JVET-AF0080](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0080-v1.zip) | Kwai  H. -J. Jhu  [JVET-AF0230](https://jvet-experts.org/doc_end_user/current_document.php?id=13493)  Alibaba  X. Li  [JVET-AF0221](https://jvet-experts.org/doc_end_user/current_document.php?id=13484) |
| 2.8 | DBV signalling modification | OPPO  L. Xu  [JVET-AF0081](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0081-v1.zip) | Qualcomm  H. Huang  JVET-AF0219 |
| 2.9a | Enable DBV in single tree under CTC | Qualcomm  H. Huang  [JVET-AF0066](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0066-v1.zip) | OPPO  L. Xu  [JVET-AF0076](https://jvet-experts.org/doc_end_user/current_document.php?id=13324) |
| 2.9b | Enable DBV in single tree when DualITree is set to 0 | Qualcomm  H. Huang  [JVET-AF0066](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0066-v1.zip) | OPPO  L. Xu  [JVET-AF0076](https://jvet-experts.org/doc_end_user/current_document.php?id=13324) |
| 2.10 | Combination test of Test 2.8 and Test 2.9 | Qualcomm  H. Huang  OPPO  L. Xu  [JVET-AF0207](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0207-v1.zip) | VIVO  C. Zhou |
| 2.11a | Test 2.1b + Test 2.5b | InterDigital  K. Naser  Nokia  S. Blasi  vivo  C. Zhou  [JVET-AF0229](https://jvet-experts.org/doc_end_user/current_document.php?id=13492) | Xiaomi  R.G. Youvalari  M. Abdoli  [JVET-AF0244](https://jvet-experts.org/doc_end_user/current_document.php?id=13507) |
| 2.11b | Test 2.2b + Test 2.5b | InterDigital  K. Naser  Nokia  S. Blasi  vivo  C. Zhou  [JVET-AF0229](https://jvet-experts.org/doc_end_user/current_document.php?id=13492) | Xiaomi  R.G. Youvalari  [JVET-AF0244](https://jvet-experts.org/doc_end_user/current_document.php?id=13507) |
| 2.11c | Test 2.3c + Test 2.5b | InterDigital  K. Naser  Nokia  S. Blasi  vivo  C. Zhou  [JVET-AF0229](https://jvet-experts.org/doc_end_user/current_document.php?id=13492) | Xiaomi  R.G. Youvalari  [JVET-AF0244](https://jvet-experts.org/doc_end_user/current_document.php?id=13507) |
| 2.11d | Test 2.3d + Test 2.5b | InterDigital  K. Naser  Nokia  S. Blasi  vivo  C. Zhou  [JVET-AF0229](https://jvet-experts.org/doc_end_user/current_document.php?id=13492) | Xiaomi  R.G. Youvalari  [JVET-AF0244](https://jvet-experts.org/doc_end_user/current_document.php?id=13507) |
| **3** **Inter prediction** | | | |
| 3.1a | CCP merge mode for chroma inter coding | MediaTek  M.-S. Chiang  Bytedance  Z. Deng  [JVET-AF0073](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0073-v1.zip) | InterDigital K. Naser  JVET-AF0240 |
| 3.1b | CCP merge mode for chroma inter coding without the additional second type of shifted temporal candidates | MediaTek  M.-S. Chiang  Bytedance  Z. Deng  [JVET-AF0073](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0073-v1.zip) | Qualcomm  Y.-J. Chang  JVET-AF0233 |
| 3.2 | LIC flag derivation of merge candidates with template costs | Bytedance  N. Zhang  [JVET-AF0128](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0128-v1.zip) | OPPO  Z. Xie |
| ~~3.3a~~ | ~~Multi-model CCRM~~ | ~~Bytedance~~  ~~Z. Deng~~ |  |
| 3.3b | CCP merge mode with the derived candidates with 4 derived candidates | Bytedance  Z. Deng  MediaTek  M.-S. Chiang  [JVET-AF0073](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0073-v1.zip) | vivo  Z. Lv  JVET-AF0242 |
| 3.1d | Test 3.1a + Test 3.3b (with 1 derived candidate for low-delay pictures) | MediaTek  M.-S. Chiang  Bytedance  Z. Deng  [JVET-AF0073](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0073-v1.zip) | OPPO  F. Wang  [JVET-AF0213](https://jvet-experts.org/doc_end_user/current_document.php?id=13476)  vivo  Z. Lv  JVET-AF0242 |
| 3.4a | TM-based subblock motion refinement | Bytedance  L. Zhao  [JVET-AF0163](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0163-v1.zip) | OPPO  Z. Xie  JVET-AF0270 |
| 3.4b | Interweaved affine prediction | Bytedance  L. Zhao  [JVET-AF0163](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0163-v1.zip) | OPPO  Z. Xie  JVET-AF0270 |
| 3.4c | RMVF candidate derivation with multiple CUs | Bytedance  L. Zhao  [JVET-AF0163](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0163-v1.zip) | OPPO  Z. Xie  JVET-AF0270 |
| 3.4d | Test 3.4a + Test 3.4b + Test 3.4c | Bytedance  L. Zhao  [JVET-AF0163](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0163-v1.zip) | OPPO  Z. Xie  JVET-AF0270 |
| 3.5a | DMVR with robust MV derivation | Ericsson  K. Andersson  [JVET-AF0057](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0057-v1.zip) |  |
| 3.5b | DMVREncSelect from VTM (encoder only) | Ericsson  K. Andersson  [JVET-AF0057](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0057-v1.zip) |  |
| 3.5a\* | Test 3.5a with QP 40, 43, 47, and 50 | Ericsson  K. Andersson  [JVET-AF0057](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0057-v1.zip) |  |
| 3.5b\* | Test 3.5b with QP 40, 43, 47, and 50 | Ericsson  K. Andersson  [JVET-AF0057](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0057-v1.zip) |  |
| 3.6a | Affine subblock BDOF refinement | Qualcomm  Z. Zhang  [JVET-AF0159](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0159-v1.zip) | Kwai  X. Xiu |
| 3.6b | AMVP-merge mode for affine | Qualcomm  Z. Zhang  [JVET-AF0159](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0159-v1.zip) | Kwai  X. Xiu |
| 3.6c | Test 3.6a + Test 3.6b | Qualcomm  Z. Zhang  [JVET-AF0159](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0159-v1.zip) | Kwai  X. Xiu |
| **4 Reference picture resampling** | | | |
| 4.1a | Enabling template-based reordering for scaled reference pictures | Kwai  X. Xiu  [JVET-AF0190](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0190-v1.zip) | Qualcomm  Z. Zhang |
| 4.1b | Test 4.1a + Enabling LIC for scaled reference pictures | Kwai  X. Xiu  [JVET-AF0190](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0190-v1.zip) | Qualcomm  Z. Zhang |
| 4.2a | Filtering applied after motion compensation, the post-processing upsampling is not changed | RWTH Aachen Univ.  T. Claßen  [JVET-AF0173](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0173-v1.zip) |  |
| 4.2b | Test 4.2a + perform the filtering after reconstruction | RWTH Aachen Univ.  T. Claßen  [JVET-AF0173](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0173-v1.zip) |  |
| **5 In-loop filtering** | | | |
| 5.1a | Dynamic TU scale factor for BIF with LUTs interpolation | Ericsson  V. Shchukin  [JVET-AF0112](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0112-v1.zip) | Bytedance  W. Yin  [JVET-AF0203](https://jvet-experts.org/doc_end_user/current_document.php?id=13464) |
| 5.1b | Dynamic TU scale factor for BIF | Ericsson  V. Shchukin  [JVET-AF0112](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0112-v1.zip) | Bytedance  W. Yin  [JVET-AF0203](https://jvet-experts.org/doc_end_user/current_document.php?id=13464) |
| 5.2a | Luma-residual tap in chroma-ALF | Bytedance  W. Yin  [JVET-AF0197](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0197-v1.zip) | Ericsson  V. Shchukin  [JVET-AF0225](https://jvet-experts.org/doc_end_user/current_document.php?id=13488) |
| 5.2b | Test 5.2a + Luma-residual tap in CCALF | Bytedance  W. Yin  [JVET-AF0197](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0197-v1.zip) | Ericsson  V. Shchukin  [JVET-AF0225](https://jvet-experts.org/doc_end_user/current_document.php?id=13488) |
| **6 Entropy coding** | | | |
| 6.1a | Spatial CABAC tuning | Nokia  J. Lainema  [JVET-AF0109](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0109-v1.zip) | Qualcomm  P. Nikitin, I. Jumakulyyev  [JVET-AF0188](https://jvet-experts.org/doc_end_user/current_document.php?id=13446) |
| 6.1b | Spatial CABAC tuning with reduced memory/latency | Nokia  J. Lainema  [JVET-AF0109](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0109-v1.zip) | Qualcomm  P. Nikitin, I. Jumakulyyev  [JVET-AF0188](https://jvet-experts.org/doc_end_user/current_document.php?id=13446) |
| 6.2 | Retrain I-slices context model parameters | Alibaba  R.-L. Liao  [JVET-AF0133](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0133-v1.zip) | InterDigital K. Naser  JVET-AF0239 |
| 6.3 | Test 6.2 + Test 6.1 | Nokia  J. Lainema  Alibaba  R.-L. Liao  [JVET-AF0110](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0110-v1.zip) | Kwai  X. Xiu  JVET-AF0232 |

***Test 1 - Partitioning***

**Test 1.1:** Non-square quadtree partitioning **(**[**JVET-AF0108**](https://jvet-experts.org/doc_end_user/current_document.php?id=13355)**)**

In this test, the existed restriction that any non-square block cannot be split with quadtree is removed for 2NxN and Nx2N blocks, where N is greater or equal to 8. The minimum block size depends on picture size and configuration, and the maximum block sizes can be decided according to the maximum block sizes for binary or ternary tree partitioning. Under the common test conditions, the largest block sizes for the non-square quadtree partitioning are 128×64 or 64×128 for Class A/B and 64×32 or 32×64 for other classes.

Fast encoding algorithm is applied in the test, which is related to constraints of partitioning depths considering that non-square quadtree depth is included in MTT depths, and non-square quadtree partitioning is allowed only once in MTT depths under the common test configurations. The second fast encoding algorithm is an early termination method based on CBF of the previous coding block split with BT or TT in rate-distortion optimization process. The third fast encoding algorithm is an early termination method based on RD-cost comparison between NO\_SPLIT and BT/TT blocks in RDO process.

Test 1.1a: Non-square quadtree with fast encoding algorithms

Test 1.1b: ECM-10.0 with maximum MTT depth plus 1



Tradeoff with encoding time is far from attractive.

It was reported by cross-checkers that similar benefit can be achieved by an encoder-only change (related contribution JVET-AF0280).

Minimum block size is 8x4/4x8 for luma.

Further study recommended – no action at this meeting.

***Test 2 - Intra prediction***

**Test 2.1: DIMD merge (**[**JVET-AF0120**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0120-v1.zip)**)**

In this method, a merged Histogram of Gradient (HoG) is introduced which includes HoG of neighboring blocks coded with DIMD or DIMD merge modes. When a single DIMD or DIMD Merge neighbouring block is available, then its histogram of gradients is used to form the merged HoG for the current block. If more than one DIMD or DIMD Merge neighbouring blocks are available, the corresponding histograms are combined by means of amplitude averaging to derive the merged HoG. Up to maximum 13 neighbor CUs around the current block are considered.

Finally, the merged HoG is used to compute intra-prediction modes and weights, as in conventional DIMD. The directional modes and their weights corresponding to the five highest amplitudes in the merged HoG are selected, and the corresponding predictors are blended as in conventional DIMD.

Test 2.1a: DIMD Merge.

Test 2.1b: DIMD Merge with reduced storage. Neighbour HoGs are combined by averaging the five highest amplitudes from their respective histograms. Only the five highest amplitudes are stored for a DIMD CU.

Worst case memory analysis is performed, considering that a picture is fully partitioned into 4x4 luma blocks and all blocks are encoded using DIMD.

|  |  |  |  |
| --- | --- | --- | --- |
| Class | Number of CUs | Test 2.1a memory  [Bytes] | Test 2.1b memory  [Bytes] |
| A1/A2 | 960 | 515632 | 38480 |
| B/F | 480 | 258352 | 19280 |
| C | 208 | 112560 | 8400 |
| D | 104 | 56816 | 4240 |
| E | 320 | 172592 | 12880 |

It is asserted that in CTC All Intra, the actual memory requirements measured correspond to a 1.1% increase for Test 2.1a, and a 0.2% increase for Test 2.1b.

In CTC Random Access, the actual memory requirements measured correspond to a 0.8% increase for Test 2.1a, and a 0.2% increase for Test 2.1b

**Test 2.2: DIMD with filtered template (**[**JVET-AF0131**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0131-v1.zip)**)**

In the test, a DIMD mode with filtered template is introduced, where the template is filtered using 3x3 filter before deriving HoG.

The 3x3 filter operator is used to filter the template as follows.

In Test 2.2a, the following 3x2 and 2x3 gradient operators are used for the gradient histogram derivation of the left and top templates respectively.

and

and

In Test 2.2b, the following 3x3 Sobel gradient operators are used, which is the same as in ECM, while using unfiltered samples for gradient computation at positions where filtered samples are unavailable.

and

**Test 2.3: Combination of DIMD related tests (**[**JVET-AF0126**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0126-v1.zip)**)**

Test 2.3a: Test 2.1a + Test 2.2a

Test 2.3b: Test 2.1a + Test 2.2b

Test 2.3c: Test 2.1b + Test 2.2a

Test 2.3d: Test 2.1b + Test 2.2b

**Test 2.4: IntraCIIP as additional mode of IntraTMP (**[**JVET-AF0130**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0130-v1.zip)**)**

In this new intra prediction mode, a regular intra prediction derived by TIMD is blended with IntraTMP prediction using the existing CIIP blending of ECM-10.0 as shown in the next figure.

CIIP Blending

Regular Intra Prediction

IntraTMP prediction

TIMD process

IntraTMP process

Intra mode

Block vectors

A flag is signalled to indicate this mode as a submode of IntraTMP, and IntraCIIP is only allowed when IntraTMP fusion, IntraTMP filtering and IntraTMP subpel are not applied.

**Test 2.5: TIMD with IntraTMP/IBC (**[**JVET-AF0136**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0136-v1.zip)**)**

In the test, block vectors from neighboring blocks, which are coded by IntraTMP or IBC, are used to compute predictor candidates for TIMD cost evaluation. In TIMD mode selection process, based on template SATD cost, additional predictors obtained using block vectors from both spatial and non-adjacent neighbors are utilized. The modified TIMD process can therefore result in combining regular and IntraTMP/IBC predictions or combining two different IntraTMP/IBC predictions.

Test 2.5a: TIMD with IntraTMP/IBC.

Test 2.5b: Test 2.5a + Test 2.4

**Test 2.6: IntraTMP block vector storing (**[**JVET-AF0079**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0079-v1.zip)**)**

In ECM, IntraTMP BVs are stored for the purpose of coding future IBC blocks in integer pel resolution, even though some of the IntraTMP BVs may have a quarter-pel resolution. In addition, IntraTMP BV is not stored for HMVP. Contrary to IBC that if the current block is coded in IBC mode, the IBC BV is always stored in a quarter-pel resolution and IBC BV is also stored for HMVP.

In this test, IntraTMP BV is stored in a quarter-pel resolution, and quarter-pel IntraTMP BV is also stored for HMVP for coding future blocks.

Test 2.6a: Fractional-pel IntraTMP BVs are stored in block vector buffer.

Test 2.6b: IntraTMP BVs are stored for HMVP.

Test 2.6c: Test 2.6a + Test 2.6b.

**Test 2.7: Extrapolation filter-based intra prediction mode (**[**JVET-AF0080**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0080-v1.zip)**)**

The extrapolation filter-based intra prediction mode consists of three steps:

1. The extrapolation filter coefficients are derived from a neighboring reconstructed area of the current block or filter shape and coefficients are inherited from a previous block coded with this mode. For the latter, a mode flag is signaled, and candidate list is constructed using spatial and non-adjacent blocks coded with this mode, initial list of 12 candidates is reduced to 6 after reordering based on SAD cost measured on the L-shape template of size 1. An index is further signalled to identify a candidate from the list.
2. The extrapolation process generates predicted signals from the top-left to bottom-right corner within the current block.
3. An intra prediction angle is derived by analyzing the gradient of the predicted block, and then the corresponding intra mode is used to select MTS, NSPT, and LFNST kernels for transform.

The mode is restricted to blocks with sizes not greater than 32x32 and luma component only.

Three filer shapes are used in this method, a choice of reconstructed area and filter shape shown in the next figure is signalled.

A black and white grid

Description automatically generated

A black screen with white squares

Description automatically generated

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

The merge mode is also introduced, where the filter shape and the filter coefficients are inherited from the previous decoded blocks that are coded with the tested extrapolation filter-based intra prediction mode or the merge mode based on this mode. In the merge mode, the positions and inclusion order of the spatial adjacent, temporal, non-adjacent, shifted temporal, and history candidates are the same as those defined in ECM-10.0 for the CCP merge prediction candidates. In addition, the merge candidates are reordered by comparing the SAD cost on an L-shape template with column width and row height equal to 1. In the SAD calculation, predictions of the template area by the mode filters are generated only from reconstructed (neighbouring and template) samples, allowing the filters to be applied in parallel rather than sequentially.

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

A diagram of a graph

Description automatically generated

The predicted samples are calculated as follows,

where is the predicted value at (x, y) in the current block, is the coefficient of the selected filter, the index of the coefficients is from 0 to 14, is a reconstructed or a predicted value used for the current position’s prediction, and are the position offsets to the current position along x and y directions, respectively.

**Test 2.8: DBV signalling modification (**[**JVET-AF0081**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0081-v1.zip)**)**

In this test, the signalling of chroma DM mode and DBV mode is changed. When one of the five corresponding luma positions is coded with either IntraTMP or IBC and DBV mode flag is not true, DM flag is skipped, the binarization of chroma prediction mode for this case is shown in the next table.

|  |  |  |
| --- | --- | --- |
| **intra\_chroma\_pred\_mode** | **bin string** | **chroma intra mode** |
| 0 | 11100 | list[0] |
| 1 | 11101 | list[1] |
| 2 | 11110 | list[2] |
| 3 | 11111 | list[3] |
| 4 | 110 | DIMD chroma |
| 5 | 10 | DM |
| 65 | 0 | DBV |

**Test 2.9: Enable DBV in single tree (**[**JVET-AF0066**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0066-v1.zip)**)**

In this test, the constraint on DBV flag not being used for single tree case is removed.

Test 2.9a: Enable DBV in single tree under CTC.

Test 2.9b: Enable DBV in single tree when DualITree is set to 0. Single tree is enabled for the anchor and the test.

**Test 2.10: Combination test of Test 2.8 and Test 2.9 (**[**JVET-AF0207**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0207-v1.zip)**)**

Test 2.10: Test 2.8 + Test 2.9a

**Test 2.11: Combination test of Tests 2.1, 2.2, 2.3, 2.4, and 2.5 (**[**JVET-AF0229**](https://jvet-experts.org/doc_end_user/current_document.php?id=13492)**)**

Test 2.11a: Test 2.1b + Test 2.5b

Test 2.11b: Test 2.2b + Test 2.5b

Test 2.11c: Test 2.3c + Test 2.5b

Test 2.11d: Test 2.3d + Test 2.5b

Results (AI/RA)



2.1-2.5 and combination 2.11 (2.11d is the best, combining 4 different elements) is giving only close to 0.2% luma gain, but increases processing needs both at encoder and decoder (also reflected in run times) – no attractive tradeoff. 2.6x is targeting unification of the BV storage between IntraTMP and IBC (IBC already using fractional pel). Though such a design polishing is not overly important in the exploration, it also gives a very small coding gain for screen content. Support for 2.6c was expressed by several experts. Decision: Adopt JVET-AF0079 test 2.6c.

2.7 has higher gain than original proposal. The proponents explain that an EIP merge mode was added. Concern is raised that this is new element proposal that was not originally planned in the EE. Furthermore, the previous proposal (test 2.8 from JVET-AE0024) had an encoder run time increase of 2.5% with 0.16% gain, and it had been requested to reduce the run time. In the new 2.7, the run time has further increased to 3% by adding the new element. Though the gain was also increased to 0.24%, this was not the original intent given by the EE definition. Further study in the next EE, to reduce the encoder run time of both the previous proposal and the extension with the new merge mode. It was also commented that the RA performance looks more attractive with the added element.

The signalling modification from 2.8 does not have significant benefit (even less in the combination 2.10 (2.8/ 2.9), enabling DBV for single tree). Test 2.9, enabling DBV with single tree (in CTC only relevant for RA) is asserted to be straightforward, likely simpler than for dual tree where it is already enabled).

Decision: Adopt JVET-AF0066, test 2.9a.

***Test 3 - Inter prediction***

**Test 3.1/3.3: CCP merge mode for chroma inter coding (**[**JVET-AF0073**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0073-v1.zip)**)**

In CCP merge mode, cross-component model parameters of the current chroma block can be inherited from a selected candidate in the CCP merge list comprising spatial adjacent, spatial non-adjacent, history-based, and default candidates. For P- and B- slices, temporal, shifted temporal, and InterCCCM candidates are further added.

In the tests, CCP merge is extended to inter coded blocks, and this mode is indicated by a flag. If applied, a CCP model is implicitly selected from a CCP merge list. The final prediction of the current chroma inter block is formed by combining the motion-compensation predicted signals and the cross-component predicted signals derived using the selected CCP model. The weights for combining predictions are fixed as (wCCP, winter) = (3/4, 1/4).

CCP merge list construction

In addition to the existing candidates of CCP merge mode, the CCP merge list includes a second type of shifted temporal candidates and the CCP models in the list can be inherited from intra and inter blocks. After constructing the CCP merge list, the candidate with the lowest template cost is implicitly selected.

Shifted temporal candidates

While positions of the first type shifted temporal candidates are derived based on neighboring motion vectors, the positions of the second type are derived based on the current motion vector. As depicted in the next figure, the position of the collocated block and the positions of C0i and C1i, 1 ≤ i ≤ 10, are shifted by a motion vector from the current block, where C0i and C1i are the positions of the temporal candidates.

A screenshot of a game

Description automatically generated

Inheritance of CCP models from inter blocks

CCP models can also be inherited from inter blocks, in addition to models from intra blocks coded in CCLM, MMLM, CCCM, GLM, chroma fusion, and CCP merge modes. For each inter block, a CCP model is stored after the block is coded in inter CCP merge mode or InterCCCM, or retrieved from a reference position located by the motion vector of the inter block following the rules of intra prediction mode propagation.

Derived candidates

A CCP model can also be derived based on the neighboring reconstructed samples of the current block without accessing the luma reconstructed samples of the current block. When any of the derived CCCM/CCLM model parameters of Cb and Cr is non-zero, such derived candidate is inserted at the beginning of the CCP merge list. At most 4 derived candidates, including single-model CCCM, multi-model CCCM, single-model CCLM, and multi-model CCLM, are inserted to the CCP merge list for an inter CCP merge mode.

Test 3.1b: The CCP merge list consists of all candidates in CCP merge mode for intra (i.e. spatial adjacent, temporal, spatial non-adjacent, history-based, shifted temporal, and default candidates) and no additional candidates are included.

Test 3.1a: Test 3.1b with the second type of shifted temporal candidates.

Test 3.3b: Test 3.1b (without temporal and shifted temporal candidates) with at most 4 derived candidates.

Test 3.1d: Test 3.1a + Test 3.3b with at most 1 derived candidate added in the front of the CCP merge list for low-delay pictures.

**Test 3.2: LIC flag derivation for merge candidates with template costs (**[**JVET-AF0128**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0128-v1.zip)**)**

In ECM, the LIC flag is inherited for a merge candidate. In the test, the LIC flag is derived for a merge candidate based on template costs by comparing two template costs: a SAD-based template cost, denoted as C0, and a Mean Removal SAD (MRSAD)-based template cost, denoted as C1. The LIC flag is set to be false, if C0 <= C1 and is set to be true, if C0 > C1.

To favor the inherited LIC flag, C0 is multiplied by *α* if the inherited LIC flag is false while C1 is multiplied by *α* if the inherited LIC flag is true, where *α* < 1.

**Test 3.4: Enhanced subblock-based motion compensation (**[**JVET-AF0163**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0163-v1.zip)**)**

Three methods are tested for subblock-based motion compensation enhancement.

TM-based subblock motion refinement

CPMVs of uni-predicted affine merge candidates and the motion shift of SbTMVP candidates are refined using TM. For a uni-predicted affine merge candidate, a same MV offset is assigned to all the CPMVs, and the TM cost of the affine candidate is calculated accordingly. The optimal CPMV offset with the minimum TM cost is used to refine the corresponding affine candidate. For a SbTMVP candidate, the initial motion shift is refined with TM, and then the refined motion shift will be utilized to derive subblock temporal motion information.

Interweaved affine prediction

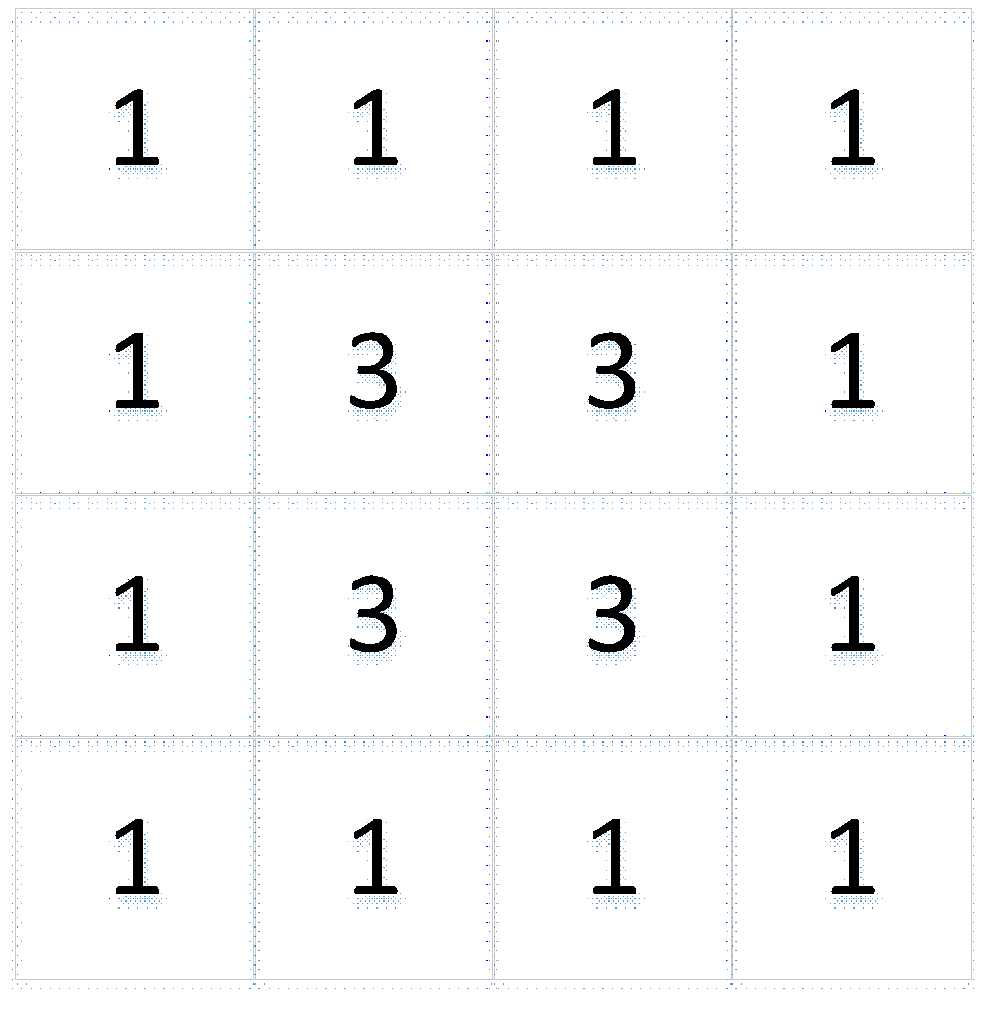
With the interweaved prediction, a coding block is divided into subblocks with two different dividing patterns. The first dividing pattern (i.e., pattern 0) is the same as that in ECM, while the second dividing pattern also divides the coding block into 4×4 subblocks but with a 2×2 offset as pattern 1 shown in the next figure.



Then the two auxiliary predictions are generated by affine motion compensation with the two dividing patterns. The final prediction is calculated as a weighted sum of the two auxiliary predictions.



As shown in next figure, weights are positions dependent, an auxiliary prediction sample located at the center of a subblock is associated with a weighting value 3, while an auxiliary prediction sample located at the boundary of a subblock is associated with a weighting value 1.



The subblocks associated with two dividing patterns are respectively refined by PROF unless the width or height is smaller than 4. Besides, subblock boundary deblocking/OBMC for affine mode is disabled.

RMVF candidate derivation with multiple CUs

Additional RMVF affine candidates are derived by taking the motion vector field of multiple CUs as regression input. In the test, two previously affine-coded CUs are simultaneously used to generate a RMVF candidate. The additional RMVF candidates are reordered together with other RMVF candidates through ARMC.

Test 3.4a: TM-based subblock motion refinement.

Test 3.4b: Interweaved affine prediction.

Test 3.4c: RMVF candidate derivation with multiple CUs.

Test 3.4d: Test 3.4a + Test 3.4b + Test 3.4c.

**Test 3.5: DMVR with robust MV derivation (**[**JVET-AF0057**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0057-v1.zip)**)**

It is asserted that multi-pass DMVR can sometimes produce dislocated subblocks due to use of unreliable motion vectors which leads to subjective artifacts. Two approaches (normative and encoder only) are tested.

Normative approach

Select motion vector that minimize distortion based on both subblock boundary distortion and subblock matching distortion when there is a risk to have unreliable motion vectors based on checks on spatial activity, subblock motion vector differences and boundary differences. The method consists of following steps:

1. Check spatial activity for the reference subblocks centered at the block motion vector. If it is determined that the spatial activity is lower than a threshold go to step 2.

2. Each candidate motion vector for the subblock is compared with a neighboring subblock’s motion vectors across a subblock boundary (if there exists). If the absolute difference between the candidate motion vector and the neighboring subblock’s motion vector for one component is greater than a threshold go to step 3.

3. A boundary check which is based on neighboring samples is made for the top respectively the left subblock boundary. If the boundary check indicates that there not is a true edge along the top subblock boundary or the left subblock boundary continue to step 4.

4. Determine subblock boundary distortion across left and/or top subblock boundary. For any subblock boundary distortion which is greater than a threshold, add the subblock boundary distortion to the block matching distortion and go to step 5.

5. Select motion vectors that minimize the total distortion.

Encoder only

In the encoder only solution the RD cost is set to a maximum value when at least one subblock of a coding unit has unreliable motion vector based on steps 1 to 4 as in the normative method following steps without adding the boundary distortion to the selection criteria. Furthermore, the RD cost is not punished for the highest temporal layer and the method is only used for QP equal to or above 33 to reduce BDR impact and still maintain subjective quality.

Test 3.5a: DMVR with robust MV derivation

Test 3.5b: DMVREncSelect from VTM (encoder only)

Test 3.5a\*: DMVR with robust MV derivation with QP 40, 43, 47, and 50

Test 3.5b\*: DMVREncSelect from VTM (encoder only) with QP 40, 43, 47, and 50

**Test 3.6: Affine subblock BDOF refinement (**[**JVET-AF0159**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0159-v1.zip)**)**

In ECM, BDOF refinement is not applied to subblocks in affine and SbTMVP modes, and AMVP-Merge mode is also not used with subblocks.

In the test, when BDOF condition is satisfied, BDOF subblock MV refinement and sample adjustment is applied to an affine or SbTMVP coded block with subblock MC.

An affine coded block, e.g. affine regular merge mode, affine BM merge mode, affine AMVP mode, derives MVs for each 4×4 subblock from the affine model. The BDOF process starts with the 4×4 subblocks grouping with identical MVs. The first iteration of BDOF MV refinement is processed in 8x8 subblock grid as in ECM-10.0. When the grouped subblock size is less than 256, the second iteration of BDOF MV refinement is processed in 4×4 subblock grid, and otherwise in 8×8 subblock grid. When the grouped subblock size is 4xN or Nx4, the first iteration of BDOF MV refinement is bypassed.

The BDOF enabling condition is the same as ECM-10.0, e.g., two reference pictures have equal POC distance to the current picture, and equal weight prediction.

AMVP-Merge mode applied to affine blocks is also tested. Like the AMVP-Merge mode in ECM-10.0, it signals reference picture list index, reference index and MVP index to indicate the AMVP predictor. A flag is signalled to indicate AMVP-Merge mode followed by affine mode flag if block size is equal or greater than 8×8 and the current picture is not a low-delay picture.

Test 3.6a: Affine subblock BDOF refinement.

Test 3.6b: AMVP-merge mode for affine.

Test 3.6c: Test 3.6a + Test 3.6b.

Results (RA/LB)

Tests 3.1/3.3: Benefit to be expected for chroma. The methods are somewhat different in selection of candidates. Test 3.1d appears most attractive, though mainly for LB.

It was commented that the gain in chroma was significantly decreased, probably due to the recent adoption of CCRM. However, also the encoder had been significantly higher (several percents), and it had been requested to be reduced, which was achieved.

It was asked what the term “low delay picture” means. It was answered that all pictures in the RPL have a POC that is before the current picture. In the case, no derived candidate is put into the list (which explains that 3.1d has no additional gain over 3.1a in RA case).

Support by non-proponents (including some of the cross-checkers).

Decision: Adopt JVET-AF0073 test 3.1d.

Test 3.2 is asserted to provide reasonable tradeoff, straightforward implementation.

Decision: Adopt JVET-AF0128 test 3.2

Tests 3.4: Most benefit seems to come from 3.4a, whereas 3.4b/c do not provide much standalone. Also in combination, the additional benefit is small. Some support was expressed by independent experts.

Decision: Adopt JVET-AF0163 Test 3.4a

Test 3.5: The purpose is avoiding visual artifacts caused by DMVR which had been observed in the context of VTM verification tests (in VTM, that had to be resolved by an encoder modification). The higher QP were tested, because more artifacts would appear in that case. It was commented by proponents that deblocking cannot resolve the problem of edge discontinuity.

It was asked if parallel processing of subblocks was still possible with the normative solution? To some extent yes in terms of the MV search, but the decision has to be done sequentially based on neighbour SB samples.

At this point of exploration, it is not urgent to go for a normative solution on this aspects. In case of a standardization, the normative solution is probably the best way to go (but eben better solutions might be possible than the one investigated in test 3.5a).

Decision (SW): Adopt JVET-AF0057 test 3.5b. Not enabled in CTC, but when visual testing with ECM is performed, it should be enabled.

Test 3.6a has a good tradeoff and is straightforward (applying BDOF for affine SBs). Compared to that, the additional benefit of test 3.6b is rather small. The increase in encoding runtime in the combination is likely caused by the fact that BDOF has then also to be executed in the affine merge mode check.

It was commented that no padding is performed in ECM, therefore combination of BDOF with affine should be straightforward.

Decision: Adopt JVET-AF0159 Test 3.6a.

***Test 4 - RPR***

This section was discussed 1340 to 1410 on Saturday 14 Oct, 2023, chaired by Y. Ye.

In RPR tests, two PSNR calculations 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 4.1: Enabling template-based reordering and LIC for scaled reference pictures (**[**JVET-AF0190**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0190-v1.zip)**)**

In the tests, TM based tools are enabled for RPR.

In Test 4.1a, the template-based inter reordering tools, including ARMC, MMVD/affine MMVD reordering, template-based BCW derivation, reference picture list reordering and MVD sign and magnitude-suffix prediction, are enabled for scaled reference pictures in the RPR.

In Test 4.1b, on top of Test 4.1a LIC is additionally enabled when any of reference pictures is in different resolution comparing to the current picture.

Test 4.1 performance (RA/LB)



It was asked why the impact on encoding/decoding time is so small due to the enabling of TM and/or LIC tools. One expert commented that this is because only 5% of the pictures are impacted by the proposal.

Multiple experts including the cross checker commented that the proposed changes are straightforward, and code change is relatively small, whereas the performance vs. complexity tradeoff is very favorable.

Comparing 4.1a and 4.1b, the latter adds LIC tool on top of TM tool in the case of RPR. Multiple experts commented that 4.1b provided more favorable gain vs complexity.

Decision: Adopt JVET-AF0190 test 4.1b.

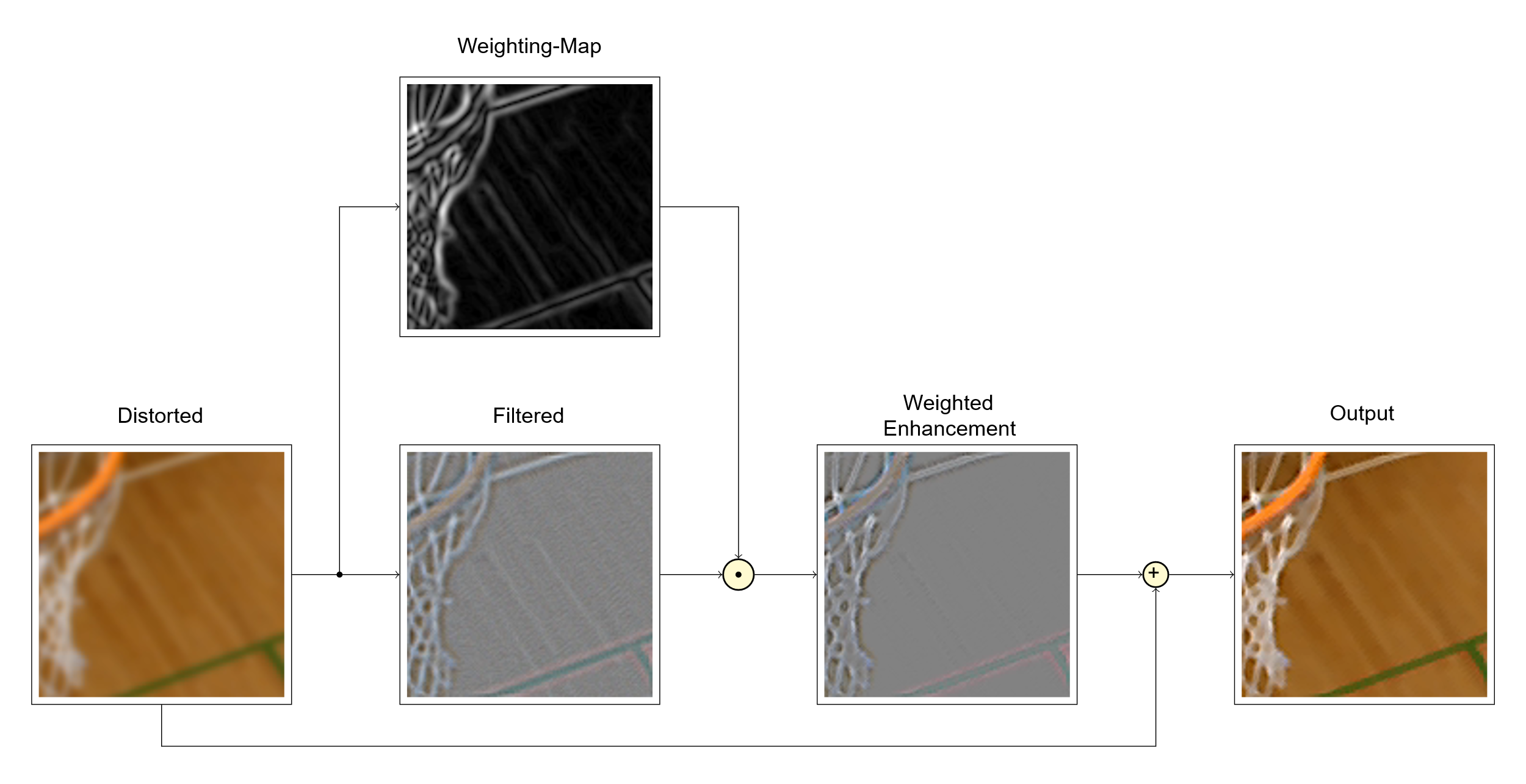
**Test 4.2: Weighted edge enhancement filtering (**[**JVET-AF0173**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0173-v1.zip)**)**

Additional processing step, adaptive locally weighted filter, is added after the upsampling.

The inputs are the decoded, upscaled picture or block and the side information which is encoded in an adaption parameter set in the bitstream. First, the filter parameters are decoded. This includes information regarding the weighting map function, filter shape, luma and chroma flags, and filter coefficients.

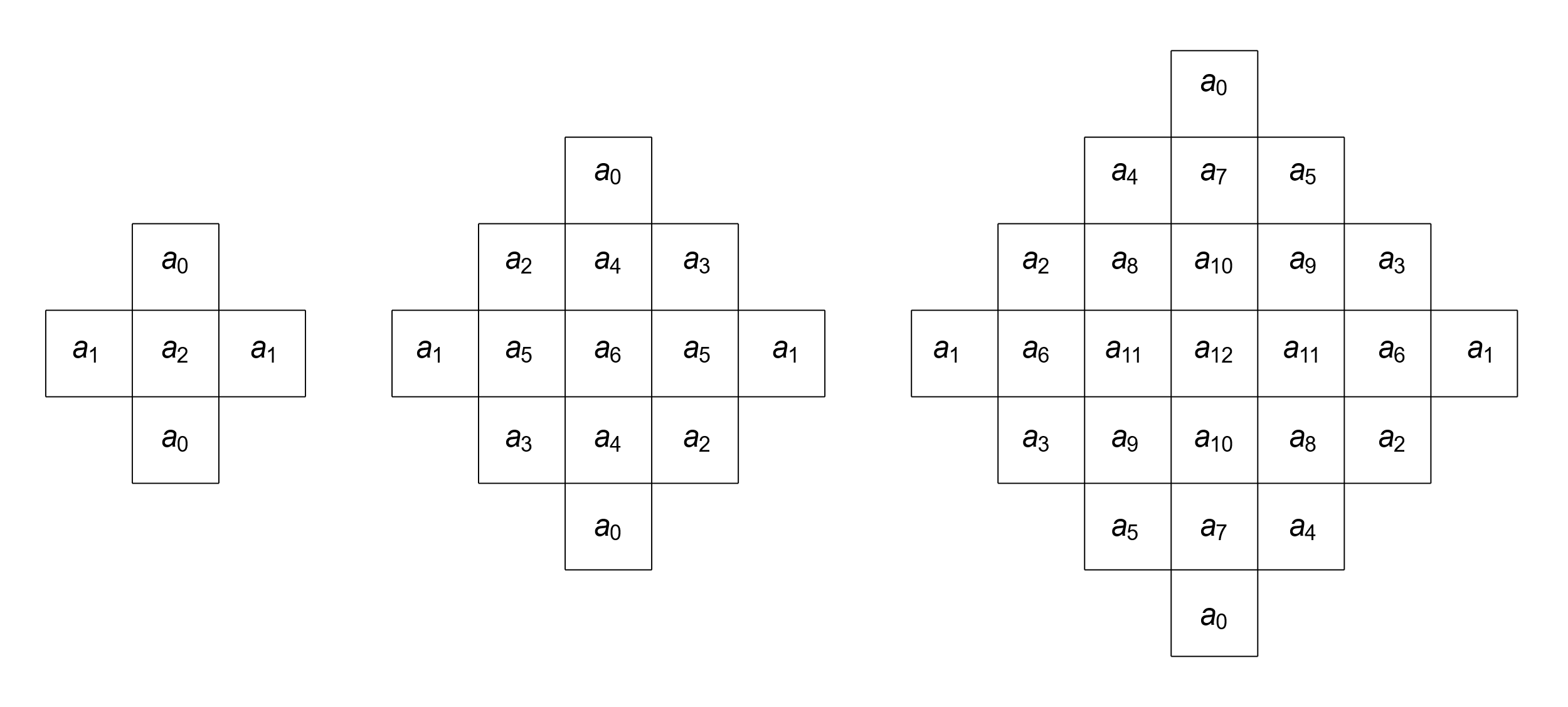
Next, the weighting map function is applied to obtain the local weighting map, which is computed such that the strength of the filter is increased at the location of edges and decreased, or set to zero, in parts of the picture with small local gradient.

Then, the filter is applied to the upscaled picture or block to generate a filtered map. The filtered map is multiplied by the local weighting and added to the upsampled picture as illustrated in next figure.



Filter coefficients are derived by minimizing the difference between the upsampled *s* and original *g* (non-donwsampled) signals.

The filters have an adaptive shape following the available shapes shown in the next figure and filter symmetries of ALF. The shape is signaled for each picture in the adaption parameter set.



In Test 4.2a, the adaptive locally weighted filtering is applied to inter prediction only but not the upscaled output for PSNR calculation.

In Test 4.2b, the filtering is also applied after reconstruction which affects PSNR calculation.

Results are not complete, available PSNR2 results for LDB are as follows: RPR 1.5x, test 4.2a

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Class C | 0.00% | 0.00% | 0.01% | 131% | 117% |
| Class E | -0.02% | 0.05% | -0.02% | 126% | 114% |

RPR2x, test 4.2a

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Class C | -0.02% | -0.05% | -0.02% | 129% | 117% |
| Class E | 0.00% | 0.00% | 0.00% | 129% | 116% |

RPR1.5x, test 4.2b

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Class C | -3.12% | -0.42% | -0.43% | 132% | 151% |
| Class E | -6.21% | 0.29% | 0.89% | 129% | 143% |

RPR2x, test 4.2b

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Class C | -4.03% | -0.69% | -0.70% | 135% | 148% |
| Class E | -12.58% | -1.11% | 0.65% | 134% | 152% |

It was commented that currently the normative aspect in 4.2a is not showing much gain, but the non-normative post processing filter in 4.2b on top of 4.2a is showing more significant coding gain. The proponent suggests that this is mainly because the RDO process is not yet very effective in 4.2a, and if the post processing filter in 4.2b would be brought “in-loop”, it would provide more useful reference for RPR prediction, and could further improve the performance from the normative aspect.

Proponent commented that 4.2a has more potential for gains, which they intend to show in a future proposal. Such a contribution would be welcomed.

Futher study encouraged.

***Test 5 - In-loop filtering***

**Test 5.1: Dynamic scaling of bilateral filter (**[**JVET-AF0112**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0112-v1.zip)**)**

In ECM, the bilateral filter offset applied to a sample is calculated using 5x5 diamond filter shape, which is sum of 12 offsets:

S0,2

S-1,1

S0,1

S1,1

S-2,0

S-1,0

S0,0

S1,0

S2,0

S-1,-1

S0,-1

S1,-1

S0,-2

where is based on a 26x16 LUT of 8-bit integers denoted by (for 26 QPs from 17 to 42 and 16 intervals of sample difference). For the innermost positions, i.e., , the LUT is used as-is, i.e., . For the remaining positions, the LUT output it right shifted: .

The TU-based scale factor is defined from minimum block size as follows.

***TU scale factor in ECM-10.0***

|  |  |  |  |
| --- | --- | --- | --- |
| Prediction mode |  |  |  |
| Inter prediction |  |  |  |
| Intra prediction |  |  |  |

In the test, the LUT and calculation changes are summarized as follows:

1. The TU scale factor depends on the TU shape size and the mean absolute difference (MAD) of the TU,
2. The BIF LUTs are interpolated.

Three scale factors () are used to pre-compute three LUTs for three different neighbor distances , i.e.:

Test 5.1a additionally uses the averaging linear interpolation

where and are successive entries of . For chroma, the right shift is decreased from 3 to 2.

The right shift in the formula for BIF offset calculation is increased from 5 to 8 and

where is based on the TU’s shape sizes and is based on the mean absolute difference (MAD) of the TU. Both and are calculated using LUTs. In total, four 64-byte tables and four 16-byte tables are introduced (for luma/chroma component, for intra/inter prediction).

*Complexity assessment of the tests against ECM-10.0*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| BIF version | Bit depth  (maximum) | Summations  (per sample) | Multiplications  (per sample) | LUT lookups  (per sample) | LUTs memory  (bytes) |
| ECM 10.0 | 12 | 18 | 0 | 6 | 832 |
| Test 5.1a | 15 | 25 | 1 | 12 | 2816 |
| Test 5.1b | 15 | 19 | 1 | 6 | 2816 |

Test 5.1a: Dynamic TU scale factor for BIF with LUTs interpolation.

Test 5.1b: Dynamic TU scale factor for BIF.

**Test 5.2: Luma residual taps in chroma ALF and CCALF (**[**JVET-AF0197**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0197-v1.zip)**)**

In ECM, chroma ALF uses a diamond 9x9 filter shape with 20 taps in total. CCALF filter uses a cross-liked filter shape with 24 taps in total, and all these taps take the spatial luma reconstruction as input.

In the test, all the spatial based taps in chroma ALF and CCALF are kept unchanged. For chroma ALF, only one luma residual tap is added, which takes down-sampled luma residual as input.

图示

描述已自动生成

For CCALF, five luma residual taps in a cross 3x3 shape are added. The extended taps take the co-located and neighbouring luma residual values as input.

**图示

描述已自动生成**

Test 5.2a: Luma residual tap in chroma ALF.

Test 5.2b: Test 5.2a + luma residual tap in CCALF.



Test 5.1: It was asked how “MAD\_TU” is computed: It is the mean absolute difference between the mean value of TU and the samples in it.

Run time and partial results confirmed by cross-checker.

Test 5.1a is slightly more complex than 5.1b (in terms of processing steps), but also has more compression benefit.

Decision: Adopt JVET-AF0112 Test 5.1a.

Test 5.2a/b affects only chroma quality. However, it was critizized that separate results for CCALF modification are not provided (test 5.2b only tests combination with 5.1a). Gain of the combination is significantly higher for LB.

Several experts expressed support for 5.2b, emphasizing the benefit in compression which comes at practically no increase in run time. One concern was raised that some optimization at the encoder is frame size dependent (but it was also said that a similar approach also exists in CCSAO).

Considering that the standalone gain of test 5.2a is relatively small compared to the combination 5.2b, it can be concluded that most benefit in 5.2b (in particular for LB) comes due to the CCALF modification.

Decision: Adopt from JVET-AF0197 the part of luma residual tap in CCALF test 5.2b.

***Test 3 - Entropy coding***

**Test 6.1: Spatial CABAC tuning (**[**JVET-AF0109**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0109-v1.zip)**)**

In ECM, CTUs are scanned in raster scan order, while CUs inside CTUs are scanned in nested tree order. When starting to process the current CTU (CTUC) the processing “jumps” from the bottom-right CU (yellow CU 8) to the top-left CU of the current CTU (blue CU 0) as shown in the next figure. At that stage, the context parameters of the CABAC engine can be expected to be finetuned for the bottom-right area of the previous CTU instead of the ideal case where the context parameters would be finetuned for the actual next CU to be processed.



In the tested method, syntax elements bins of the bottom CUs of the above CTU (the green CUs in the figure) are used to tune the CABAC contexts when starting to process a new CTU.

In the implementation, the certain number of bins are buffered for each CTU of a CTU line, where each entry contains a bin and 10-bit context ID associated with that bin. The buffer size is 512 in Test 6.1a and 256 in Test 6.1b.

Following table summarizes the memory requirement per CTU for the bin buffers.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Test | Bits / element | Number of elements | Total bits / CTU | Total bytes / CTU |
| Test 6.1a | 11 | 512 | 5632 | 704 |
| Test 6.1b | 11 | 256 | 2816 | 352 |

For each context, no more than 8 bins are stored, so 4-bit counter is used. The number of contexts in ECM-10.0 is 852 making the needed size of these counters 4 \* 852 bits = 3408 bits = 426 bytes.

The total worst-case memory requirement then depends on the number of CTUs in a CTU row and summarized as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| Class | CTUs / row | Test 6.1a memory [Bytes] | Test 6.1b memory [Bytes] |
| A1 | 15 | 10986 | 5706 |
| A2 | 15 | 10986 | 5706 |
| B | 15 | 10986 | 5706 |
| C | 7 | 5354 | 2890 |
| D | 4 | 3242 | 1834 |
| E | 10 | 7466 | 3946 |
| F | 15 | 10986 | 5706 |
| TGM | 15 | 10986 | 5706 |

Test 6.1a: Spatial CABAC tuning with 512 buffer.

Test 6.1b: Spatial CABAC tuning with 256 buffer.

**Test 6.2: Updating I-slice context model parameters (**[**JVET-AF0133**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0133-v1.zip)**)**

The context model parameters have not been retrained since ECM-5.0. In this test, the context model parameters for I-slice are retrained. The training scripts are modified from the one used in VVC’s CE on CABAC initialization. The training bitstream are generated using ECM-10.0 and CTC sequences (including class A1, class A2, class B, class C, class D, class E and class F) with QP 17 to QP 42 in AI configuration.

**Test 6.3: Combination of Test 6.1 and Test 6.2 (**[**JVET-AF0110**](https://jvet-experts.org/doc_end_user/documents/32_Hannover/wg11/JVET-AF0110-v1.zip)**)**

Retraining of the contexts Test 6.2 is added first and then the spatial CABAC tuning from Test 6.1 is applied.

Test 6.3a: Test 6.2 + Test 6.1a.

Test 6.3b: Test 6.2 + Test 6.1b.



It was suggested that a regular retraining of CABAC initialization within each meeting cycle would be beneficial. Currently, this was only possible for I slices (as per 6.2). A script is provided, which however is not easy to understand according to crosscheckers.

It was asked why the rate offset parameter was not trained? Because those are shared between intra and inter, such that it would introduce some inconsistency as long as inter initialization is not retrained.

It is agreed that the approach from test 6.2 is a reasonable step forward to perform the re-training, even though futher improvements appear possible and can certainly be investigated in upcoming experiments.

Decision: Adopt the CABAC initialization parameters from JVET-AF0133 Test 6.2. Also the script should be included in the ECM package, such that it can be used by other experts.

The benefit of the 6.1 tests is relatively small, and it has some impact on the memory complexity of the CABAC engine. The combination test with 6.2 had been suggested from the previous meeting to understand possible interaction with an optimized CABAC initialization. This seems not to exist for AI, but some inconsistency is observed e.g. for LB in test 6.3b. Considering that new input is available at this meeting for performing the update of CABAC init also for inter cases, further study in next EE would be valuable in that context.

### 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-AF0057](https://jvet-experts.org/doc_end_user/current_document.php?id=13305) EE2-3.5: DMVR with robust MV derivation [K. Andersson, R. Yu (Ericsson)]

[JVET-AF0218](https://jvet-experts.org/doc_end_user/current_document.php?id=13481) Cross-check of JVET-AF0057 (EE2-3.5: DMVR with robust MV derivation) H. Huang (Qualcomm)

[JVET-AF0066](https://jvet-experts.org/doc_end_user/current_document.php?id=13314) EE2: Test 2.9 Enable DBV in single tree [H. Huang, H. Wang, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-AF0076](https://jvet-experts.org/doc_end_user/current_document.php?id=13324) Crosscheck of JVET-AF0066 (EE2: Test 2.9 Enable DBV in single tree) [L. Xu (OPPO)] [late]

[JVET-AF0073](https://jvet-experts.org/doc_end_user/current_document.php?id=13321) EE2-3.1 and EE2-3.3: 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), Z. Deng, K. Zhang, L. Zhang (Bytedance)]

[JVET-AF0213](https://jvet-experts.org/doc_end_user/current_document.php?id=13476) Crosscheck of JVET-AF0073 (EE2-3.1 and EE2-3.3: Cross-component prediction merge mode for chroma inter coding) test 3.1d F. Wang (OPPO) [late]

[JVET-AF0233](https://jvet-experts.org/doc_end_user/current_document.php?id=13496) Crosscheck of JVET-AF0073 (EE2-3.1b: Cross-component prediction merge mode for chroma inter coding) [Y.-J. Chang (Qualcomm)] [late]

[JVET-AF0240](https://jvet-experts.org/doc_end_user/current_document.php?id=13503) crosscheck of JVET-AF0073: EE2-3.1a: Cross-component prediction merge mode for chroma inter coding [K. Naser (InterDigital)] [late]

[JVET-AF0242](https://jvet-experts.org/doc_end_user/current_document.php?id=13505) Crosscheck of JVET-AF0073 (EE2-3.1 and EE2-3.3: Cross-component prediction merge mode for chroma inter coding) test3.3b and test3.1d [Z. Lv (vivo)] [late]

[JVET-AF0079](https://jvet-experts.org/doc_end_user/current_document.php?id=13327) EE2-2.6: IntraTMP block vector storing [Y. Yu, L. Zhang, L. Xu, H. Yu, J. Gan, F. Wang, Z. Xie, D. Wang (OPPO), P.-H Lin, J. -L Lin, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-AF0227](https://jvet-experts.org/doc_end_user/current_document.php?id=13490) Crosscheck of JVET-AF0079 (EE2: Test 2.6a, Test 2.6b and Test 2.6c - Combination of Test 2.6a and Test 2.6b) [W. Lim, S.-C. Lim (ETRI)] [late]

[JVET-AF0080](https://jvet-experts.org/doc_end_user/current_document.php?id=13328) EE2-2.7: An extrapolation filter-based intra prediction mode [L. Xu, Y. Yu, H. Yu, J. Gan, D. Wang (OPPO)]

[JVET-AF0221](https://jvet-experts.org/doc_end_user/current_document.php?id=13484) Crosscheck of JVET-AF0080 (EE2-2.7: An extrapolation filter-based intra prediction mode) [X. Li (Alibaba)] [late]

[JVET-AF0230](https://jvet-experts.org/doc_end_user/current_document.php?id=13493) Crosscheck of JVET-AF0080 (EE2-2.7: An extrapolation filter-based intra prediction mode) [H.-J. Jhu (Kwai)] [late]

[JVET-AF0081](https://jvet-experts.org/doc_end_user/current_document.php?id=13329) EE2-2.8: DBV improvement [L. Xu, Y. Yu, H. Yu, J. Gan, D. Wang (OPPO)]

[JVET-AF0219](https://jvet-experts.org/doc_end_user/current_document.php?id=13482) Cross-check of JVET-AF0081 (EE2-2.8: DBV improvement) [H. Huang (Qualcomm)] [late]

[JVET-AF0108](https://jvet-experts.org/doc_end_user/current_document.php?id=13355) EE2-1.1: Non-square quadtree partitioning [Y. Ahn, J. Nam, N. Park, J. Lim, S. Kim (LGE)]

[JVET-AF0249](https://jvet-experts.org/doc_end_user/current_document.php?id=13512) Crosscheck of JVET-AF0108 (EE2-1.1: Non-square quadtree partitioning) [R. Utida, F. Le Léannec, T. Dumas (InterDigital)] [late]

[JVET-AF0109](https://jvet-experts.org/doc_end_user/current_document.php?id=13356) EE2-6.1: Spatial CABAC tuning [J. Lainema, A. Aminlou, P. Astola, D. B. Sansli (Nokia)]

[JVET-AF0188](https://jvet-experts.org/doc_end_user/current_document.php?id=13446) Cross-check of JVET-AF0109 (EE2-6.1: Spatial CABAC tuning) [P. Nikitin, I. Jumakulyyev (Qualcomm)] [late]

[JVET-AF0110](https://jvet-experts.org/doc_end_user/current_document.php?id=13357) EE2-6.3: Combination of EE2-6.1 and EE2-6.2 [J. Lainema, A. Aminlou, P. Astola, D. B. Sansli (Nokia), R.-L. Liao, Y. Ye, J. Chen, X. Li (Alibaba)]

[JVET-AF0232](https://jvet-experts.org/doc_end_user/current_document.php?id=13495) Crosscheck of JVET-AF0110 (EE2-6.3: Combination of EE2-6.1 and EE2-6.2) [X. Xiu (Kwai)] [late]

[JVET-AF0112](https://jvet-experts.org/doc_end_user/current_document.php?id=13359) EE2-5.1: Dynamic Scaling of Bilateral Filter (BIF) [V. Shchukin, P. Wennersten, J. Ström (Ericsson)]

[JVET-AF0203](https://jvet-experts.org/doc_end_user/current_document.php?id=13464) Crosscheck of JVET-AF0112 (EE2-5.1: Dynamic Scaling of Bilateral Filter) [W. Yin (Bytedance)] [late]

[JVET-AF0120](https://jvet-experts.org/doc_end_user/current_document.php?id=13367) EE2-2.1 DIMD merge [S. Blasi, I. Zupancic, J. Lainema (Nokia)]

[JVET-AF0124](https://jvet-experts.org/doc_end_user/current_document.php?id=13369) Crosscheck of JVET-AF0120 (EE2-2.1: DIMD merge) [P. Andrivon (Ofinno)]

[JVET-AF0126](https://jvet-experts.org/doc_end_user/current_document.php?id=13371) EE2-2.3: Combination of DIMD related tests (EE2-2.1 + EE2-2.2) [S. Blasi, I. Zupancic, J. Lainema (Nokia)]

[JVET-AF0204](https://jvet-experts.org/doc_end_user/current_document.php?id=13465) Crosscheck of JVET-AF0126 (EE2-2.3-AB: Combination of DIMD Related Tests) [W. Yin (Bytedance)] [late]

[JVET-AF0276](https://jvet-experts.org/doc_end_user/current_document.php?id=13540) crosscheck of JVET-AF0126: EE2-2.3: Combination of DIMD related tests (EE2-2.3c + EE2-2.3d) [K. Naser (InterDigital)] [late]

[JVET-AF0128](https://jvet-experts.org/doc_end_user/current_document.php?id=13386) EE2-3.2: LIC flag derivation for merge candidates with template costs [N. Zhang, K. Zhang, H. Liu, Y. Wang, L. Zhang (Bytedance)]

[JVET-AF0269](https://jvet-experts.org/doc_end_user/current_document.php?id=13533) Crosscheck of JVET-AF0128 (EE2-3.2: LIC flag derivation for merge candidates with template costs) [Z. Xie (OPPO)] [late]

[JVET-AF0130](https://jvet-experts.org/doc_end_user/current_document.php?id=13388) EE2-2.4: IntraCIIP as additional mode of IntraTMP [K. Naser, P. Bordes, F. Galpin, K. Reuzé (InterDigital)]

[JVET-AF0286](https://jvet-experts.org/doc_end_user/current_document.php?id=13550) Crosscheck of JVET-AF0130 (EE2-2.4: IntraCIIP as additional mode of IntraTMP) [B. Ray (Qualcomm)] [late]

[JVET-AF0131](https://jvet-experts.org/doc_end_user/current_document.php?id=13389) EE2-2.2: DIMD with filtered template [C. Zhou, Z. Lv (vivo)]

[JVET-AF0235](https://jvet-experts.org/doc_end_user/current_document.php?id=13498) Crosscheck of JVET-AF0131 (EE2-2.2: DIMD with filtered template [J. Chen (Alibaba)] [late]

[JVET-AF0133](https://jvet-experts.org/doc_end_user/current_document.php?id=13391) EE2-6.2: updating I-slice context model parameters [R.-L. Liao, Y. Ye, J. Chen, X. Li (Alibaba)]

[JVET-AF0239](https://jvet-experts.org/doc_end_user/current_document.php?id=13502) crosscheck of JVET-AF0133: EE2-6.2: updating I-slice context model parameters [K. Naser (InterDigital)] [late] [miss]

[JVET-AF0136](https://jvet-experts.org/doc_end_user/current_document.php?id=13394) EE2-2.5: TIMD with IntraTMP and IBC [K. Naser, F. Le Léannec, P. Bordes, F. Galpin, A. Robert (InterDigital)]

[JVET-AF0125](https://jvet-experts.org/doc_end_user/current_document.php?id=13370) Crosscheck of JVET-AF0136 (EE2-2.5a: TIMD with IntraTMP/IBC merge) [P. Andrivon (Ofinno)]

[JVET-AF0215](https://jvet-experts.org/doc_end_user/current_document.php?id=13478) Crosscheck of JVET-AF0136 (EE2-2.5: TIMD with IntraTMP and IBC) test 2.5b F. Wang (OPPO) [late]

[JVET-AF0243](https://jvet-experts.org/doc_end_user/current_document.php?id=13506) Crosscheck of JVET-AF0136 (EE2-2.5: TIMD with IntraTMP and IBC) test 2.5b [S. Blasi (Nokia)] [late]

[JVET-AF0159](https://jvet-experts.org/doc_end_user/current_document.php?id=13417) EE2-3.6: Affine subblock BDOF refinement [Z. Zhang, H. Huang, J.-L Lin, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-AF0231](https://jvet-experts.org/doc_end_user/current_document.php?id=13494) Crosscheck of JVET-AF0159 (EE2-3.6: Affine subblock BDOF refinement) [X. Xiu (Kwai)] [late]

[JVET-AF0163](https://jvet-experts.org/doc_end_user/current_document.php?id=13421) EE2-3.4: Enhanced subblock-based motion compensation [L. Zhao, K. Zhang, L. Zhang (Bytedance)]

[JVET-AF0270](https://jvet-experts.org/doc_end_user/current_document.php?id=13534) Crosscheck of JVET-AF0163 (EE2-3.4: Enhanced subblock-based motion compensation) [Z. Xie (OPPO)] [late]

[JVET-AF0173](https://jvet-experts.org/doc_end_user/current_document.php?id=13431) EE2-4.2: Weighted edge enhancement filtering [T. Claßen, M. Wien (RWTH)]

[JVET-AF0265](https://jvet-experts.org/doc_end_user/current_document.php?id=13529) Crosscheck of JVET-AF0173 (EE2-2.4: Weighted edge enhancement filtering) [J. Samuelsson-Allendes (Sharp)] [late]

[JVET-AF0190](https://jvet-experts.org/doc_end_user/current_document.php?id=13451) EE2-Test4.1: Enabling template-based inter tools for the RPR [X. Xiu, H.-J. Jhu, C.-W. Kuo, C. Ma, N. Yan, W. Chen, X. Wang (Kwai)]

[JVET-AF0261](https://jvet-experts.org/doc_end_user/current_document.php?id=13525) Crosscheck of JVET-AF0190 (EE2-Test4.1: Enabling template-based inter tools for the RPR) [Z. Zhang (Qualcomm)] [late]

[JVET-AF0197](https://jvet-experts.org/doc_end_user/current_document.php?id=13458) EE2-5.2: Luma Residual Taps in Chroma-ALF and CCALF [W. Yin, K. Zhang, Z. Deng, L. Zhang (Bytedance)]

[JVET-AF0225](https://jvet-experts.org/doc_end_user/current_document.php?id=13488) Crosscheck of JVET-AF0197 (EE2-5.2: Luma Residual Taps in Chroma-ALF and CCALF) [V. Shchukin (Ericsson)] [late]

[JVET-AF0207](https://jvet-experts.org/doc_end_user/current_document.php?id=13468) EE2: Test 2.10 Combination of Test 2.8 and Test 2.9 [H. Huang, H. Wang, V. Seregin, M. Karczewicz (Qualcomm), L. Xu, Y. Yu, H. Yu, J. Gan, D. Wang (OPPO)]

[JVET-AF0220](https://jvet-experts.org/doc_end_user/current_document.php?id=13483) Cross-check of JVET-AF0207 (EE2: Test 2.10 Combination of Test 2.8 and Test 2.9) [C. Zhou (vivo)] [late]

[JVET-AF0229](https://jvet-experts.org/doc_end_user/current_document.php?id=13492) EE2-2.3: Combination of DIMD/TIMD related tests (EE2-2.3 + EE2-2.5) [K. Naser, F. Le Léannec, P. Bordes, A. Robert (InterDigital), S. Blasi, I. Zupancic, J. Lainema (Nokia), C. Zhou, Z. Lv (Vivo)] [late]

[JVET-AF0244](https://jvet-experts.org/doc_end_user/current_document.php?id=13507) Crosscheck of JVET-AF0229 (EE2-2.11: Combination of DIMD/TIMD related tests) [R. G. Youvalari, M. Abdoli (Xiaomi)] [late]

### EE2 related contributions (6)

Contributions in this area were discussed at 1415–1615 on Saturday 14 Oct. 2023 and 1830-1845 on Sunday 15 October 2023 (chaired by Y. Ye).

[JVET-AF0053](https://jvet-experts.org/doc_end_user/current_document.php?id=13301) EE2-related: Test 2.5a and TIMD fusion improvement [P. Andrivon, M. Blestel (Ofinno), K. Naser (InterDigital)]

Contribution JVET-AE0184, studied in EE2-2.5a (JVET-AF0136), proposes to extend TIMD candidates list with IntraTMP and IBC intra prediction modes of neighboring blocks. Contribution JVET-AE0137 introduces a non-angular intra predictor (among DC or Planar) in the TIMD fusion process, additionally to the two selected TIMD modes. Besides, the location-dependent sample-based fusion process from DIMD is adjusted for TIMD. This contribution merges both EE2-2.5a and JVET-AE0137 and proposes to add the IntraTMP or IBC mode of neighboring blocks with the lowest template matching cost to the non-angular intra predictor candidates list. On top of ECM-10.0, the reported PSNR-Y, Cb, Cr, BD-rate and {EncT, DecT} results are as follows:

* AI: -0.08%, -0.05%, -0.04% {100.3%, 99.7%}
* RA: -0.03%%, 0.00%, -0.10% {98.8%, 100.1%}

Cross checker confirmed that partial results match, and supported to study in EE. This contribution doubles the gain reported in EE2-2.5a (JVET-AF0136).

It was commented that JVET-AF0132 is a closely related contribution. See notes under JVET-AF0132.

[JVET-AF0245](https://jvet-experts.org/doc_end_user/current_document.php?id=13508) Crosscheck of JVET-AF0053 (EE2-related: Test 2.5a and TIMD fusion improvement) [S. Blasi (Nokia)] [late]

[JVET-AF0132](https://jvet-experts.org/doc_end_user/current_document.php?id=13390) Non-EE2: Optimization of TIMD blending mode [C. Zhou, Z. Lv (vivo)]

This contribution proposes an optimization method for reference lines selection in TIMD blending mode. For the second mode, the determination of which reference line to be used is not only based on the second mode, but also on the first mode.

On top of the ECM10.0, simulation results are reported as below:

AI: -0.03% (Y), -0.02% (U), -0.10% (V), 99.8% (EncT), 100.3% (DecT)

Cross checker confirmed that partial results (performance and runtime) match.

Multiple experts considered this to be a simple modification to TIMD blending and supported investigation in EE,

Proponent of the related contribution JVET-AF0053 commented that the proposed changes in this contribution is orthogonal to the blending element in JVET-AF0053, and could be additive on top of the gain from JVET-AF0053.

JVET-AF0053 provided above-average gains for some class A and B sequences.

Investigate in EE JVET-AF0053, JVET-AF0132, as well as the combination of the two proposals in the following way:

* From JVET-AF0053 : Add a non-angular intra predictor (among DC or Planar) in the TIMD fusion process using the TIMD candidate list in the next ECM version
* From JVET-AF0053 : In addition, use location-dependent sample-based fusion process from DIMD in TIMD
* From JVET-AF0132 : determination of reference line in TIMD fusion
* Combination of the above

[JVET-AF0279](https://jvet-experts.org/doc_end_user/current_document.php?id=13543) Crosscheck of JVET-AF0132 (Non-EE2: Optimization of TIMD blending mode) [H. Wang (Qualcomm)] [late]

[JVET-AF0106](https://jvet-experts.org/doc_end_user/current_document.php?id=13353) EE2-related: Non-adjacent spatial candidates for DIMD merge [J. Huo, J. Fan, Z. Zhang, Y. Ma, F. Yang (Xidian Univ.), M. Li (OPPO)]

In EE2 test 2.1 described in document JVET-AE2024, DIMD merge is investigated. This contribution proposes to utilize non-adjacent spatial neighbors to further improve the coding efficiency of DIMD merge. Specifically, several new candidates, derived from non-adjacent spatial blocks, are added to the existing candidate list of DIMD merge.

The proposed method + EE2-2.1a leads to the following gain over ECM-10.0:

AI: {-0.12%, 0.04%, -0.05%, xxx%, xxx%}

The proposed method + EE2-2.1b leads to the following gain over ECM-10.0:

AI: {-0.11%, 0.02%, 0.00%, xxx%, xxx%}

The proposed method + EE2-2.1a leads to the following gain over test2.1a:

AI: {-0.05%, -0.01%, -0.01%, xxx%, xxx%}

The proposed method + EE2-2.1b leads to the following gain over test2.1b:

AI: {-0.04%, -0.02%, -0.02%, xxx%, xxx%}

This proposal includes non-adjacent spatial neighbors into the DIMD merge list. The proponent commented that this will enable more DIMD blocks whose adjacent spatial neighbors are not using DIMD. It was commented that this could cause an increase in runtime.

Cross checker reported that encoding time increase is around 4%, which doubles the runtime increase in EE2-2.1x.

It was commented that performance vs. runtime tradeoff in EE2-2.1x was considered not desirable, and performance vs. runtime tradeoff in this contribution seems to be even less.

It was commented that 0.11%-0.12% gain in AI seems desirable.

Investigate in EE, reduction in encoding time is necessary to bring more desirable performance vs. runtime tradeoff.

[JVET-AF0223](https://jvet-experts.org/doc_end_user/current_document.php?id=13486) Crosscheck of JVET-AF0106 (EE2-related: Non-adjacent spatial candidates for DIMD merge) [X. Li (Alibaba)] [late]

[JVET-AF0151](https://jvet-experts.org/doc_end_user/current_document.php?id=13409) EE2 related: CABAC parameters retraining [F. Galpin, F. Lo Bianco, C. Salmon-Legagneur, K. Naser (InterDigital)]

This contribution proposes to update context initialization parameters for all slice types. All CABAC parameters (initial probabilities parameters, window sizes, adaptive weights and rate offsets) are retrained and updated. As previous scripts are not suitable anymore, a set of new scripts taking any account new parameters and constraints is proposed. It is reported that on top of ECM-10.0, the overall coding performance impact for {Y, U, V} is {-0.16 %, -0.03%, -0.10% } {-0.15 %, 0.19%, 0.11% } {-0.11 %, 0.91%, -0.06% } in AI, RA and LDB configurations respectively.

It was commented that the U component in LDB and LDP show a non-negligible performance loss.

The cross checker used the retrained contexts from the proponent to run the simulation, but didn’t cross check the training due to lack of scripts. It was suggested to investigate this in EE and include the training scripts in the EE.

Investigate in EE, include the training script in the EE.

[JVET-AF0252](https://jvet-experts.org/doc_end_user/current_document.php?id=13515) Crosscheck of JVET-AF0151 (EE2 related: CABAC parameters retraining) [M. Abdoli, R. G. Youvalari (Xiaomi)] [late]

[JVET-AF0202](https://jvet-experts.org/doc_end_user/current_document.php?id=13463) EE2 related: CABAC context initialization retraining and slice type based window offsets [V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes CABAC context initialization retraining for ECM-10.0 and to split window offsets initialization per slice type as for other CABAC parameters.

The simulation results are not complete at the time of presentation, Class C results are as follows:

CABAC retraining:

|  |  |  |  |
| --- | --- | --- | --- |
| AI | -0.20% | -0.06% | 0.04% |
| RA | -0.14% | -0.14% | -0.25% |
| LB | -0.06% | -0.55% | -0.80% |

CABAC retraining with **slice type based rate offsets**:

|  |  |  |  |
| --- | --- | --- | --- |
| AI | -0.23% | 0.00% | 0.04% |
| RA | -0.16% | -0.15% | -0.44% |
| LB | -0.09% | -0.33% | -0.30% |

This contribution includes the non-normative retraining aspect, as well as a normative aspect that make the offset dependent on slice types. The second aspect provides additional gain over retraining-only.

Training is done somewhat differently from JVET-AF0151. Both proponents suggest to work on a unified training script.

Investigate in EE, retraining only with scripts, and additionally the normative aspect of slice type-dependent offset. For the normative aspect, additional memory cost should be reported.

[JVET-AF0280](https://jvet-experts.org/doc_end_user/current_document.php?id=13544) EE2-related: complete results for EE2-1.1b in JVET-AF0249 [T. Dumas, F. Le Léannec, R. Utida (InterDigital)] [late]

During the 31st JVET meeting, Geneva, 11-19 July 2023, the tests of ECM with various maximum MTT depth increments were requested, c.f. EE2-1.1b in JVET-AE2024. This will provide a fair comparison for EE2-1.1a evaluating ECM with the non-square QT, c.f. JVET-AF0108. JVET-AF0249 reports the cross-check of JVET-AF0108. Besides, JVET-AF0249 gives the final mean BD-rate versus encoding runtime curve resulting from the runs of ECM with different maximum MTT depth increments. This contribution reports all the results used to construct this curve.

It is reported that, in RA configuration, over all classes, at equivalent encoding runtimes, ECM including the non-square QT yields -0.04% of mean BD-rate gains in luma with respect to ECM with the maximum MTT depth increased by 1 on I-slices and the maximum MTT depth increased by 1 for the first two TIDs.

This contribution is for information. It was suggested that in the future, contributions with normative changes to partitioning should always bring runtime vs. gain tradeoff curves as shown in this contribution and compare to non-normative encoder-only changes.

It was commented that this is reasonable request and future partitioning contributions should bring such information to JVET.

[JVET-AF0288](https://jvet-experts.org/doc_end_user/current_document.php?id=13552) Crosscheck of JVET-AF0280 (EE2-related: complete results for EE2-1.1b in JVET-AF0249) [W. Ahmad (Ericsson)] [late]

[JVET-AF0295](https://jvet-experts.org/doc_end_user/current_document.php?id=13559) Crosscheck of JVET-AF0280 (EE2-related: complete results for EE2-1.1b in JVET-AF0249) [A. Filippov, K. Suverov, V. Rufitskiy (Ofinno)] [late]

[JVET-AF0291](https://jvet-experts.org/doc_end_user/current_document.php?id=13555) EE2 related: Encoder time reduction of Test-2.11 [K. Naser, F. Le Léannec, P. Bordes, A. Robert (InterDigital), S. Blasi, I. Zupancic, J. Lainema (Nokia), C. Zhou, Z. Lv (Vivo)] [late]

This contribution proposes an encoder-side optimisation of tests in EE2-2.11 tests. It is asserted that this corresponds to a minimal change to the EE2-2.11x branch. The simulation results on AI for Test-2.11d with the encoder time reduction in this contribution are:

{ -0.12%, 0.02%, -0.06%, 100.9% (EncT) 98.6% (DecT) 100.3% (VmPeak)}

EE2 Test-2.11d was not adopted for two reasons: 1) the encoder runtime increase for AI was 4%, and 2) Test-2.11d is a combination of four tools.

The combined gain of 4 tools are now only 0.12% for AI, which means the individual gain for each tool is probably only ~0.03%.

Further study is recommended.

### ECM modifications and software improvements beyond EE2 (41)

Contributions in this area were discussed at 1615–1910 on Saturday 14 Oct. 2023, 0840-1830 on Sunday 15 Oct, 2023 (chaired by Y. Ye), and on Monday 16 Oct. 1715-1755 (chaired by JRO)

#### Intra and CIIP (18)

[JVET-AF0059](https://jvet-experts.org/doc_end_user/current_document.php?id=13307) AHG12: Fix to interpolation filter for intra prediction [K. Andersson (Ericsson)]

It was noted that ECM-10.0 can give stripe like visual artifacts and it was discovered that the reason for that was in interpolation in intra prediction due to non-unity gain of one of the 64-phase interpolation filters. It is therefore suggested to correct the filter with non-unity gain using a mirror version of corresponding mirror phase filter.

Objective performance (BDR) versus ECM-10.0 is reported as:

AI: Y: -0,01%, U 0,01%, V -0,05%

It is asserted that fixing the gain of the interpolation filter with non-unity gain also can improve subjective quality.

In ECM-10.0, the intra prediction interpolation at phase 41/64 has filter coefficients { 3, -21, 106, 180, -25, 3 } that sum up to 246 rather than 256. The proposal suggests to use the mirroring coefficients of phase 23/64 for phase 41/64, with coefficients { 3, -20, 106, 189, -25, 3}.

It was commented that this is the only phase that has this unintended behaviour.

It was also shown that fixing the coefficients bring subjective quality improvement.

Decision: adopt JVET-AF0059.

[JVET-AF0072](https://jvet-experts.org/doc_end_user/current_document.php?id=13320) AHG12: FIBC flag inherits IntraTMP-FLM flag [M.-H. Jia, Y. Gao, Y. -X. Bai, S.-W. Xie, P. Wu, C. Huang (ZTE)]

This contribution proposes that FIBC flag in the IBC merge mode can inherit the IntraTMP-FLM flag. In the IBC merge mode, the FIBC flag is inferred from the merge candidate. When the merge candidate is comes from the IntraTMP coded block, the FIBC flag can be inferred through IntraTMP-FLM flag.

The proposed method shows coding gains as follows:

Test A : Proposed method (the FIBC flag inherits the IntraTMP-FLM flag) on top of ECM-10.0

AI:

Class F： -0.02% / 0.04% / 0.02%, 100.4% / 99.8%

ClassTGM：-0.02% / -0.03% /-0.03%, 100.2% / 100.1%

Test B : Proposed method on top of EE2-2.6c (storing quarter-pel resolution IntraTMP BVs + IntraTMP BV storing for HMVP)

AI:

Class F： -0.04% / -0.02% / 0.02%, xx% / xx%

ClassTGM： 0.01% / 0.04% / -0.01%, xx% / xx%

It was commented that the gains are extremely small. One expert commented that in the previous round of EE, proposals with tiny gains were included in EE. Several experts commented that this might not be the most productive way to conduct EE tests.

Further study to bring more gains.

[JVET-AF0074](https://jvet-experts.org/doc_end_user/current_document.php?id=13322) Crosscheck of JVET-AF0072 (AHG12: FIBC flag inherits IntraTMP-FLM flag) [J. Fu, [S. Ma (PKU)](mailto:swma@pku.edu.cn)]

[JVET-AF0084](https://jvet-experts.org/doc_end_user/current_document.php?id=13332) Non-EE2: Update on 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-10.0 software, and the performance is reported as follows:

AI: { -0.47% Y, -0.51% U, -0.36% V, 101.0% EncT, 100.6% DecT } for class F

{ -0.15% Y, -0.13% U, -0.11% V, 100.4% EncT, 102.6% DecT} for class TGM

RA: { -0.30% Y, -0.29% U, -0.23% V, 100.5% EncT, 100.0% DecT } for class F

{ -0.07% Y, -0.13% U, -0.10% V, 100.4% EncT, 101.8% DecT } for class TGM

It was asked if this method can be applied to IBC mode in natural content, the proponent commented that this had been tried on natural content and no gain was obtained. However, in the EE setting, results for natural content can be provided as well.

This proposal would require the model parameters from spatial adjacent and non-adjacent neighbors and history candidates to be additionally stored. Decoding time increase is likely due to more effort at the decoder to derive the IBC block vectors.

Investigate in EE, also consider potentially better tradeoff between performance and decoding time in terms of number of inherited models.

[JVET-AF0271](https://jvet-experts.org/doc_end_user/current_document.php?id=13535) Crosscheck of JVET-AF0084 (Non-EE2: Update on IBC-LIC Model Merge mode) [Y. Wang (Bytedance)] [late]

[JVET-AF0104](https://jvet-experts.org/doc_end_user/current_document.php?id=13351) Non-EE2: Direct Decoder-side Intra Mode Derivation for IntraTMP blocks [J. Huo, N. Qiu, Y. Ma, F. Yang (Xidian Univ.), M. Li (OPPO)]

In ECM10.0, DIMD is used to derive an intra prediction mode for Intra TMP blocks. It is conducted to the prediction signal after filtering, sub-pixel interpolation, or fusion. This contribution is proposed that using the initial predictions to derive *intraTmpDimdMode*. On top of ECM-10.0, simulation results are reported as follows:

AI(Y/U/V): -0.00%/-0.01%/-0.06%, EncT 100.1%, DecT 100.2%

This proposal is for improved parallelism for DIMD, it has no performance impact. However, at this stage of the exploration, refining a small aspect is not our main goal.

No action at this time.

[JVET-AF0222](https://jvet-experts.org/doc_end_user/current_document.php?id=13485) Crosscheck of JVET-AF0104 (Non-EE2: Direct Decoder-side Intra Mode Derivation for IntraTMP blocks) [X. Li (Alibaba)] [late]

[JVET-AF0115](https://jvet-experts.org/doc_end_user/current_document.php?id=13362) Non-EE2: IBC with a further upward-extended reference area for screen content [Y. Kidani, H. Kato, K. Kawamura (KDDI)]

This contribution proposes an intra block copy (IBC) with a further upward-extended reference area for screen content to enhance the coding performance of IBC. In the current ECM design, the reference area for IBC is extended to two CTU rows above the current CTU, but there is room to achieve additional coding gain for screen content by extending the reference area beyond the referenceable two CTU rows. Therefore, this contribution proposes two further upward extension methods for the IBC reference area, qualitatively comparing their advantages and disadvantages in terms of complexity. Specifically, method 1 extends the reference area to the picture's upper side boundary, while method 2 does four CTU rows above the current CTU.

The experimental results reportedly show the proposed method brings some coding gains over ECM-10.0. The results of method 2 are shown as follows:

AI: Class F {-0.14%, -0.22%, -0.10 %, 100.1%, 99.9%}, Class TGM {-1.21%, -1.21%, -1.20 %, 99.4%, 99.3%};

RA: Class F {-0.06%, -0.01%, -0.04%, 100.1%, 100.0%}, Class TGM {-0.66%, -0.61%, -0.64%, 99.9%, 100.9%}.

It was asked why increased IBC search area doesn’t incur an increase in encoder runtime. The proponent said that encoder search strategy in ECM-10 had not been changed, and the fact that encoding time is not increase could be due to the fact that only integer BVs are used for SCC content, and that hash-based search is used. It was commented that a previous adoption that increased IBC search range also didn’t result in encoder runtime increase for screen content.

It was commented that the IntraTMP search range and IBC search range in ECM-10 are different, and this proposal only suggests to extend the search range for IBC.

It was commented that currently the search range for natural content and screen content is the same, whereas this proposal could create different search ranges depending on the type of content.

The proponent said that no gain was obtained for natural content when the search range was enlarged.

Investigate in EE both methods in JVET-AF0115, results for natural content should be made available as well.

[JVET-AF0116](https://jvet-experts.org/doc_end_user/current_document.php?id=13363) Non-EE2: slope adjustment for IBC LIC [C. Ma, X. Xiu, W. Chen, H.-H.Jhu, C.-W. Kuo, N. Yan, X. Wang (Kwai)]

This contribution proposes to improve IBC LIC with slope adjustment, where an offset parameter is utilized to adjust the slope parameter of IBC LIC. The proposed method is only applied to IBC AMVP CUs. The simulation results on top of ECM-10.0 are summarized as follows:

Screen Content:

AI:

Class F : {-0.21%, -0.14%, -0.26%, 100.9%, 99.6%, 100.0%}, Class TGM : {-0.12%, -0.13 %, -0.19%, 100.5%, 100.3%, 100.0%};

RA:

Class F : {-0.19%, 0.06%, -0.05%, 100.0%, 99.2%, 100.0%}, Class TGM : {-0.13%, -0.15%, -0.14%, 99.5%, 99.8%, 100.0%};

An offset adjustment is determined by encoder through RDO to adjust the slope of IBC LIC. Encoder speedup algorithms are applied to reduce the impact to encoding time due to additional RDO.

The offset is signaled in a similar way as CCLM slope adjustment (range of value is the same). It was requested that the proponent upload a revised contribution that provides more details about the algorithm.

Cross checker confirmed that simulation results are matched, and support study in EE.

The proponent commented that no gain was observed for natural content.

Investigate in EE, results for natural content should be provided as well.

[JVET-AF0289](https://jvet-experts.org/doc_end_user/current_document.php?id=13553) Crosscheck of JVET-AF0116 (Non-EE2: slope adjustment for IBC LIC) [Y. Wang (Bytedance)] [late]

[JVET-AF0117](https://jvet-experts.org/doc_end_user/current_document.php?id=13364) Non-EE2: extensions of IBC GPM [C. Ma, X. Xiu, W. Chen, H.-H.Jhu, C.-W. Kuo, N. Yan, X. Wang (Kwai)]

This contribution proposes to extend IBC GPM scheme in ECM10.0 from two aspects: First, IBC GPM with block vector difference (IBC GPM BVD) is introduced. Second, the split mode of IBC GPM is reordered with template matching methods. The simulation results on top of ECM-10.0 are summarized as follows:

1. IBC GPM BVD:

Screen Content:

AI:

Class F : {-0.20%, -0.12%, 0.01%, 103.4%, 99.1%, 100.0%}, Class TGM : {-0.26%, -0.24 %, -0.23%, 103.3%, 99.8%, 100.0%};

RA:

Class F : {-0.10%, 0.04%, 0.02%, 102.3%, 99.8%, 99.9%}, Class TGM : {-0.21%, -0.14%, -0.17%, 101.4%, 99.4%, 100.0%};

Natural Content :

AI : {0.00%, 0.00%, 0.00%, 100.0%, 100.0%, 100.0%}, RA : {0.00%, 0.00 %, 0.00%, 100.0%, 100.0%, 100.0%};

1. IBC GPM BVD + Split mode reordering :

Screen Content:

AI:

Class F : {-0.21%, -0.22%, -0.05%, 105.6%, 99.9%, 101.4%}, Class TGM : {-0.39%, -0.34 %, -0.36%, 106.4%, 99.5%, 101.1%};

RA:

Class F : {-0.16%, 0.22 %, -0.04%, 103.4%, 99.8%, 101.2%}, Class TGM : {-0.26%, -0.16%, -0.26%, 102.3%, 99.1%, 100.7%};

Natural Content :

AI : {0.00%, 0.00%, 0.00%, 100.0%, 100.0%, 100.0%}, RA : {0.00%, 0.00 %, 0.00%, 100.0%, 100.0%, 100.0%};

It was commented that the split mode reordering seems to provide performance vs. runtime tradeoff that is less than IBC GPM BVD. The proponent said current encoder could be further optimized.

It was asked why there is no decoding time increase for split mode. This could be because only the first set is reordered, which is not always activated, therefore the impact on decoding time is minimal. However, encoder time increase due to split mode reordering is about 2%, which is likely due to additional RDO (currently a coarse RDO followed by a fine RDO) process in the current implementation being unoptimized.

Cross checker confirmed that results are matched, and suggests that the proposal is a straightforward extension. Upon examination of the code, the encoding time increase seems to be related to unoptimized code, and can be potentially reduced.

Several experts expressed interest in studying this proposal in EE.

Currently in CTC for natural content, IBC GPM is not enabled, therefore this proposal has no impact on natural content performance. However, it might be interesting to test the proposal with IBC GPM turned on for natural content as the anchor, and see if there would be additional gain brought by this proposal. It could also be interesting to check the gain from IBC GPM on natural content in the latest ECM software. Proponents are requested to bring this as additional data to the next meeting.

Investigate in EE, with separate EE tests on the two elements (IBC GPM BVD, and split mode reordering), and reduction in encoder runtime is requested.

[JVET-AF0290](https://jvet-experts.org/doc_end_user/current_document.php?id=13554) Crosscheck of JVET-AF0117 (Non-EE2: extensions of IBC GPM) [R.-L. Liao (Alibaba)] [late]

[JVET-AF0129](https://jvet-experts.org/doc_end_user/current_document.php?id=13387) Non-EE2: Auto relocated block vector prediction [N. Zhang, K. Zhang, L. Zhang (Bytedance)]

In this contribution, auto relocated block vector prediction (AR-BVP) is introduced into IBC merge/AMVP candidate list construction. AR-BVP can be derived as a guiding BV plus a BV of a reference block, located by the guiding BV. The AR-BVP candidates are inserted after the HBVP candidates.

On top of the ECM10.0, simulation results of the proposed method are reported as below:

AI:

Overall: {-0.07%, -0.06%, -0.12%, 101.1%, 99.8%, 100.0%};

Class F: {-0.33%, -0.23%, -0.10%, 101.0%, 101.2%, 100.0%};

Class TGM: {-0.47%, -0.44%, -0.48%, 101.3%, 100.6%, 100.0%};

RA:

Overall: {-0.03%, 0.01%, -0.02%, 100.4%, 99.4%, 100.2%};

Class F: {-0.20%, -0.10%, -0.12%, 101.1%, 99.4%, 100.0%};

Class TGM: { -0.36%, -0.25%, -0.22%, 100.8%, 100.5%, 100.0%}.

Encoding time is increased by about 1%, which is due to the need to check additional BVs.

It was asked how many additional BVs are added. The total number of BVs in the merge list is not changed.

Several experts commented that this proposal presents an interesting idea, and provides some gains in natural content in addition to screen content. They supported further study in EE.

Investigate in EE, also study the effect on the length of the path that AR BV traces (i.e. the value of “n” in the proposal).

[JVET-AF0137](https://jvet-experts.org/doc_end_user/current_document.php?id=13395) AHG12: IntraTMP with Merge Candidates [K. Naser, F. Le Léannec, A. Robert, F. Galpin (InterDigital)]

In the current IntraTMP design, the template matching is performed on a predefined search area to obtain a candidate list of predictors specified by their block vectors. Compared to IBC mode, the readily available block vectors from neighboring merge candidates are not explored in IntraTMP mode despite being more correlated with the properties of the current block. This contribution proposes extending the IntraTMP candidate list with all the merge candidates block vectors. The following results are obtained for natural content:

AI: -0.05% 0.01% -0.02% EncT 100.6% DecT 101.3%

For screen content, AI configuration, the following results are provided:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Class F | -0.16% | -0.11% | 0.09% | 99.0% | 99.6% | 99.9% |
| Class TGM | -0.20% | -0.16% | -0.14% | 96.5% | 93.4% | 99.7% |

The provided results show a reduction in encoding/decoding time for Class TGM, but according to the proponent, such runtime may not be accurate. Close to 100% runtime is expected.

One expert expressed support for investigating in EE.

Investigate in EE.

[JVET-AF0161](https://jvet-experts.org/doc_end_user/current_document.php?id=13419) AHG12: Harmonized IBC/ITMP search area with adaptive sampling [D. Ruiz Coll, B. Chen (Ofinno), K. Naser, P. Bordes, F. Le Léannec, F. Galpin (InterDigital)]

This contribution proposes the harmonization of the IBC and IntraTMP search areas. The IBC and IntraTMP tools use the reconstructed samples of the current picture to select a candidate prediction block for the current block. The IBC search area comprises a constant number of CTUs, while the IntraTMP search area is based on the current block dimension, which is mainly smaller than the IBC search area. IntraTMP searches for the most similar L-shaped, Top-Only, or Left-Only template that matches the template of the current block in the search area. In IntraTMP, a constant sampling factor of 3 is applied in the horizontal and vertical search directions to alleviate the template-matching complexity. However, the extension of the IntraTMP search area demands more searches. This contribution proposes two methods to determine an adaptive sampling factor when the IntraTMP search region is extended to the IBC search region.

On top of ECM-10.0, simulation results are reported as follows:

**Test-1**: Sampling rate based on the searching distance.

AI: Class F: {-0.43%, -0.35%, -0.28%, 100.9%, 105.0%},

AI: Class TGM: {-0.70%, -0.72 %, -0.73%, 101.5%, 108.3%};

RA: Class F: {-0.16%, -0.00%, -0.04%, 100.5%, 100.5%},

RA Class TGM: {-0.43%, -0.54 %, -0.52%, 100.2%, 102.6%};

**Test-2**: Sampling rate based on the IntraTMP regions dimension.

AI: Class F: {-0.29%, -0.26%, -0.14%, 101.4%, 104.1%},

AI: Class TGM: {-0.43%, -0.43 %, -0.48%, 100.4%, 101.1%};

**Test-3**: Combination of Test-1 and Test-2.

AI: Class F\*: {-0.32%, -0.32%, -0.24%, 100.8%, 101.5%},

AI: Class TGM: {-0.32%, -0.33%, -0.36%, 101.1%, 100.0%};

Contribution JVET-AF0115 also extends IBC search range.

It was asked if this was tested on natural content. An earlier interation of the proposal was tested and showed unfavorable performance vs. runtime tradeoff for natural content. For this current proposal, it could be tested again on natural content to see the performance vs. runtime tradeoff.

This proposal changes the search range of IntraTMP, and uses the entire area of IBC search range. It further applies two elements to the IntraTMP search area: 1) sampling rate based on search distance, and 2) sampling rate based on dimensions.

Since JVET-AF0115 extends IBC search range, and the concept of this proposal is to use the IBC search range, then thre is natural interaction between the two proposals.

It was commented that combining the two elements in this proposal would actually result in worse performance (Test 3 vs. Test 1 and Test 2). However, the proponents said that Test 3 was mainly targeted at reducing decoding time.

Investigate in EE, test the two elements separately and in combination, also test the interaction between this proposal and JVET-AF0115 in the form of additional combination tests. Results on natural content are also requested.

[JVET-AF0166](https://jvet-experts.org/doc_end_user/current_document.php?id=13424) Non-EE2: Intra TMP fusion probing [J.-L. Lin, P.-H. Lin, Y.-J. Chang, Z. Zhang, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes an Intra TMP fusion probing method to select a fusion candidate with minimum probing cost from a fusion candidate list. The probing cost is derived as the SAD between the reference and current block’s probing lines. The simulation results on top of ECM-10.0 are summarized as follows:

AI(Y/U/V): -0.09 %/ -0.04 %/ -0.11 %, EncT 102.1 %, DecT 101.5 %

It was asked why there is a 2% runtime increase. According to the proponent, there are two factors that could cause the runtime increase: the calculation of the probing cost, and an additional RDO that is conditionally triggered.

Encoding time should be further reduced to provide better tradeoff.

Several experts expressed interest in studying this in EE.

Investigate in EE, better performance vs. runtime tradeoff should be investigated.

[JVET-AF0170](https://jvet-experts.org/doc_end_user/current_document.php?id=13428) Non-EE2: CIIP with subblock-based motion compensation [L. Zhao, K. Zhang, L. Zhang (Bytedance)]

In this contribution, CIIP with subblock-based motion compensation is presented, where the inter prediction signal of CIIP may be generated based on affine or sbTMVP motion compensation. It is reported that on top of ECM-10.0, simulation results of the proposed method are as below:

RA: -0.07 % (Y), -0.06 % (U), -0.07 % (V), 100.4% (EncT), 100.4% (DecT)

LB: -0.08 % (Y), -0.21 % (U), -0.28 % (V), 101.2% (EncT), 100.4% (DecT)

In the current ECM, the inter part of CIIP doesn’t use subblock motion candidate, this contribution extends the inter part to also use subblock motion (affine and SbTMVP).

It was commented that though gain is small, the performance vs. runtime tradeoff seems to be reasonable. Support was expressed to study in EE by several experts.

It was commented that there is a 0.33% luma loss for class F in LDB. It was suggested to add class TGM to the EE test, such that it could be better understood if this proposal is suitable for screen content or not.

Investigate in EE, screen content results are requested to be provided.

[JVET-AF0171](https://jvet-experts.org/doc_end_user/current_document.php?id=13429) Non-EE2: Sample distance-based weighting for CIIP [J. Zhao, S. Kim (LGE)]

In this contribution, sample distance-based weighting for CIIP is proposed for CIIP angular modes. Weights of intra and inter prediction are calculated based on sample distance to the boundary, intra mode angle, as well as block size.

Experimental results show that there is BD rate gain of -0.03%, 0.04%, 0.03% in RA with little change to encoding and decoding time.

The proponent commented that currently gain is quite small, and requested to further study the aspect of improving the weighting design for CIIP.

Further study is encouraged.

[JVET-AF0184](https://jvet-experts.org/doc_end_user/current_document.php?id=13442) Non-EE2: Search range optimization for Intra TMP [G. Verba, Z. Zhang, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes to extend Intra TMP search area by enlarging the existed rectangular search area, adding horizontal and vertical oriented areas relative to the current block to the left and above of the rectangular search area, and adding areas around BV candidates. The experimental results on top of ECM-10.0 for SCC classes are reported as follows:

AI:

Overall: { -0.06%, -0.12%, -0.10%, 105.0%, 111.4%},

Class F: {-0.48%, -0.55%, -0.48%, 102.1%, 109.4%},

Class TGM: {-0.53%, -0.52%, -0.43%, 102.8%, 115.1%}

It was commented that the decoder runtime increase is quite substantial, and given this is intraTMP that relies on decoder search, encoder-only algorithm will not be able to solve the problem of decoding time increase.

It was commented that looking for a better performance vs. runtime tradeoff for this contribution would require changing the normative aspects during EE, which is not in alignment with the purpose of an EE test.

Proponent claims that an individual element in the contribution provides better performance vs. runtime tradeoff, but this information is not yet uploaded. Proponent is requested to provide a revised contribution with detailed information about this aspect as well as its performance.

This contribution was further discussed 1815 to XXXX Sunday 15 October 2023, after JVET-AF0184-v4 was made available.

In v4, the following results are additionally provided for when only adding areas around IBC merge BV candidate. The experimental results on top of ECM-10.0 are as follows:

AI:

Overall: {-0.06%, -0.03%, -0.05%, 101.4%, 104.8%},

Class F: {-0.21%, -0.13%, -0.15%, 98.9%, 103.0%},

Class TGM: {-0.35%, -0.31%, -0.26%, 99.8%, 109.2%}

Compared to the earlier results, the new results had less increase in decoder runtime. However, the decoder runtime is still increased by 5%, 3% and 9% for natural content, class F, and class TGM, respectively. The available performance data do not seem to justify the increase in decoder runtime.

Better performance vs. runtime tradeoff should be provided, before conducting EE.

Further study is recommended.

[JVET-AF0199](https://jvet-experts.org/doc_end_user/current_document.php?id=13460) Non-EE2: Bilateral Filtering for Intra Prediction [W. Yin, K. Zhang, Y. Wang, L. Zhang (Bytedance)]

In this contribution, Bilateral Filtering (BF) for intra-prediction is proposed. In the proposed method #1, the intra reference samples are filtered by BF before they are used to generate the intra-prediction. In the proposed method #2, the intra prediction samples are filtered by BF before they are used in later processes.

On top of ECM-10.0, simulation results of proposed method #1 are reported as below:

AI: -0.04%, -0.02%, -0.08%, 100.3%, 100.2%.

On top of ECM-10.0, simulation results of proposed method #2 are reported as below:

AI:\* -0.06%, -0.04%, -0.10%, 100.2%, 100.2%.

RA results are not complete by the time of presentation.

This proposal applies an additional filtering process using bilateral filter on the angular prediction modes. Essentially, the intra prediction block is formed as in ECM-10, a BF filter is applied, and then the filtered block is used as the prediction block. This is applied as an additional process without an on/off flag. This is the method #2 in the proposal. For method #1, a BF filter is applied to the reference samples before 1:2:1 filtering. Method 2 uses 2D bilateral filter, and method #1 uses 1D bilateral filter.

In the current implementation, this is not applied to angular prediction blocks with width/height larger than 128. All other angular prediction blocks use the additional BF filter.

The cross checker reported that the results are matched, and commented that the proposal seems to be simple.

It was asked if the proponent tried to replace the 1:2:1 filter on reference samples with BF, the proponent said that it was tried but no performance benefits were found.

It was commented that the method #2 of this proposal puts the BF at the same stage of intra prediction as the previously proposed diffusion filter for inter prediction during the development of VVC. It was commented that at the time, block size restriction was placed to limit the application of diffusion filter on small blocks (e.g. blocks smaller than 8x8). However, at that time, the block size restriction was placed for practical implementation consideration, rather than for performance gain.

Gains from both methods are small, with the gains from method #2 being somewhat larger and perhaps a bit more interesting.

Several experts expressed interest in testing method #2.

It was asked if method #1 and method #2 could be combined. Currently no.

Method #2 puts BF after PDPC (without any modification to PDPC).

Investigate in EE, only the method #2.

[JVET-AF0](https://jvet-experts.org/doc_end_user/current_document.php?id=13566)302 Crosscheck of JVET-AF0199 (Non-EE2: Bilateral Filtering for Intra Prediction) [X. Li (Alibaba)] [late]

[JVET-AF0226](https://jvet-experts.org/doc_end_user/current_document.php?id=13489) AHG12: SGPM with IntraTMP and IBC [K. Naser, F. Le Léannec, P. Bordes, Y. Chen (InterDigital)] [late]

SGPM is an intra mode with prediction signal generated by blending 2 intra prediction modes with specific split line. This contribution proposes further including block-vector based prediction obtained from IntraTMP or IBC modes of the neighboring blocks in the candidate list of SGPM. The results are obtained:

AI: -0.04% -0.05% -0.07% EncT 100.6% DecT 100.5%

In SC classes (classF and TGM)

AI: -0.54% -0.44% -0.46% EncT 99.65% DecT 102.35%

IntraTMP and IBC modes are made available to SGPM, which seems to be a straightforward extension.

It was asked why decoding time fluctuates from class to class. The proponent said some small increase in decoding time is expected, and more accurate decoding time could be provided during EE.

Several experts expressed interest in studying this in EE.

Investigate in EE.

[JVET-AF0275](https://jvet-experts.org/doc_end_user/current_document.php?id=13539) AHG12: Intra TMP extension to DIMD, LIC and SGPM [F. Le Léannec, K. Naser, G. Rath, T. Dumas, P. Bordes, Y. Chen (InterDigital)] [late]

This contribution proposes 3 extensions to the intra TMP prediction mode in ECM.

First extension concerns DIMD. It is proposed to adaptively select between planar or block-vector based prediction obtained from IntraTMP or IBC mode of neighboring blocks for the blending with an angular intra prediction.

Second extension consists in allowing Local Illumination Compensation in blocks coded in IntraTMP mode. LIC is already allowed for IBC blocks. Thus, extending it to IntraTMP harmonizes IntraTMP and IBC with regards to LIC usage.

Last, block-vector-based prediction obtained from IntraTMP or IBC modes of the neighboring blocks are included in the candidate list of SGPM, as proposed in contribution JVET-AF0226.

The following results are obtained for AI and RA configurations.

AI: -0,09% (Y) -0,01% (U) -0,12% (V) 100,8% (EncT) 100,6% (DecT)

For screen contents in AI configuration:

Class F: -0,40% -0,36% -0,29% 101,7% (EncT) 100,5% (DecT)

Class TGM: -0,78% -0,79% -0,78% 99,9% (EncT) 100,3% (DecT)

The two elements (LIC and DIMD) in this contribution that are not in JVET-AF0226 bring more gain for natural content, but have limited additional benefit for screen content.

One expert expressed support for studying in EE, considering all elements are intraTMP related.

Investigate in EE, conduct separate tests on individual elements (LIC, DIMD, SGPM) as well as combination tests.

#### Inter (13)

[JVET-AF0082](https://jvet-experts.org/doc_end_user/current_document.php?id=13330) Non-EE2: on LFNST/NSPT for inter coding [F. Wang, J. Gan, Y. Yue, H. Yu, D. Wang (OPPO)]

This contribution proposes to utilize LFNST/NSPT for inter coding. For inter coded block, an intra prediction mode is first derived according to the inter prediction block. Then the derived intra prediction mode is used to select an LFNST/NSPT transform set and the transform can be processed with the selected LFNST/NSPT kernel, like in the intra coding process. On top of ECM-10.0, the performance of the proposed inter LFNST/NSPT is reported as follows:

RA {-0.37% Y, 0.14% U, 0.09% V, 114.0% EncT, 101.3% DecT, 100.0% VmPeak}

It was asked if this would create any parsing dependency, it was confirmed by multiple experts that this proposal doesn’t have parsing dependency issues.

It was asked if new transform kernels for inter coding are trained, the answer is no. Only signaling (of which kernel is used) is different from intra coding.

Some preliminary results of improved encoding time are shown in the uploaded PPT package.

Several experts commented that the gains are attractive, and that there is evidence that encoding time can be reduced.

It was asked if there is a block size restriction, the answer is that current implementation has no such limitation.

It was suggested to enable intra LFNST in LDB and use that as the anchor for EE test in LDB config. This was agreed.

Is this applied to luma only? Yes.

Investigate in EE, with reduction in encoding time desirable.

[JVET-AF0273](https://jvet-experts.org/doc_end_user/current_document.php?id=13537) Crosscheck of JVET-AF0082 (Non-EE2: On LFNST/NSPT for inter coding) [C. Zhou (vivo)] [late]

[JVET-AF0113](https://jvet-experts.org/doc_end_user/current_document.php?id=13360) Non-EE2: Modifications of affine merge candidates [K. Kim, D. Kim, J.-H. Son, [J.-S. Kwak (WILUS)](mailto:jinsam.kwak@wilusgroup.com)]

This contribution proposes to use the SbTMVP for the derivation of both constructed affine merge candidates and history-affine-parameter-based candidates if TMVP is not available. Compared with ECM-10.0, the proposed methods achieve -0.01% and -0.07% gain in RA and LDB configurations, respectively. Also, additional memory usage is 99.9% and 100.3% gain in RA and LDB configurations, respectively.

RA: {-0.01%, -0.02%, 0.00%, 99.8%, 100.3%, 99.9%}

LDB: {-0.07%, 0.00%, -0.16%, 100.0%, 100.6%, 100.3%}

It was commented that the gains from the proposal are quite small.

Further study to improve performance.

[JVET-AF0251](https://jvet-experts.org/doc_end_user/current_document.php?id=13514) Crosscheck of JVET-AF0113 (Non-EE2: Modifications of affine merge candidates) [R.-L. Liao (Alibaba)] [late]

[JVET-AF0118](https://jvet-experts.org/doc_end_user/current_document.php?id=13365) Non-EE2: Regression-based GPM blending [P. Bordes, F. Galpin, K. Naser, F. Urban, K. Reuzé, F. Le Léannec, E. François (InterDigital)]

This contribution presents an enhanced GPM mode where the splitting shape and the blending function are derived from the template. The proposed extended GPM mode was implemented on top of ECM-10.0, it reportedly provides {Y, U, V} BD-reduction with encoder and decoder runtimes as follows:

Method-A: RA: -0.07%, -0.11%, -0.14%, 100.8% (EncT), 99.9% (DecT), LDB: -0.11%, 0.21%, -0.25%, 101.0% (EncT), 99.9% (DecT).

Method-B: RA: -0.06%, -0.08%, -0.10%, 101.9% (EncT), 99.6% (DecT), LDB: -0.15%, -0.11%, -0.31%, 101.9% (EncT), 100.1% (DecT).

Two variations of the proposed method are tested, with method-B using more combinations of mode pairs than method-A. The proposed method is an additional GPM mode (one flag signaled to indicate whether it is used).

It was commented that there are two parts to the proposal, modified blending function and mofiication of the GPM index coding (with template matching based reordering). It was asked if these two parts have been tested separately. Currently no. It was commented that further understanding the coding gain from individual parts might be interesting.

It was commented that there is some performance loss in individual sequences (e.g. FourPeople) in LDB config. The proponent is requested to further investigate this phenomenon and try to resolve the issue.

Several experts expressed interest in further studying the proposed method.

Investigate in EE, test the two elements (modified blending, and modified index coding) separately, use method-A as the basis for the EE tests. Bring results for natural content as well as screen content.

[JVET-AF0297](https://jvet-experts.org/doc_end_user/current_document.php?id=13561) Crosscheck of JVET-AF0118 (Non-EE2: Regression-based GPM blending) [M. Blestel (Ofinno)] [late]

[JVET-AF0121](https://jvet-experts.org/doc_end_user/current_document.php?id=13368) AHG12: Adjusting out-of-boundary prediction samples [P. Astola, J. Lainema (Nokia)]

This contribution proposes to improve the out-of-boundary (OOB) prediction by adjusting the uni-prediction samples using a scaling parameter and an offset to better match with the bi-prediction samples. The scale and offset are derived using the uni-prediction samples and the averaged bi-prediction samples.

The impact on coding efficiency and runtimes over ECM-10.0 is reportedly {for Y, U, V, EncT, DecT, VmPeak}: RA { -0.06, 0.03%, 0.01%, 100.3% 100.4%, 100.1% }, LB { -0.07%, 0.23%, -0.07%, 99.5%, 100.2%, 100.1%}.

The proposal tries to improve OOB prediction using scaling and offset parameters, which are derived by the encoder and the decoder in the same manner, and the scaling and offset parameters are constant for a given block.

This proposed method would replace the current OOB handling in ECM-10.0.

Investigate in EE.

[JVET-AF0134](https://jvet-experts.org/doc_end_user/current_document.php?id=13392) AHG12: AMVP with SbTMVP mode [R.-L. Liao, J. Chen, Y. Ye, X. Li (Alibaba)]

The AMVP with SbTMVP mode was firstly proposed in JVET-AC0103, and was studied in EE2. The simulation result on top of ECM-8.0 was reported in JVET-AD0152, and the overall coding performance and runtime for {Y, U, V, EncT, DecT} was {-0.06%, -0.03%, -0.04%, 102%, 100%} for RA and {-0.22%, 0.06%, 0.04%, 102%, 100%} for LB. It was commented that the results did not provide attractive benefit in terms of compression, compared to the increased encoder.

In this contribution, the coding efficiency and complexity of the AMVP with SbTMVP mode are improved. Given a CU coded in AMVP with SbTMVP mode, the CU is predicted in a similar way as that of SbTMVP in merge mode except that the motion shift is signaled in the bitstream instead of being derived from neighboring blocks. It is reported that on top of ECM-10.0, the overall coding performance impact for {Y, U, V, EncT, DecT, VmPeak} is {-0.10%, -0.01%, -0.02%, 101.1%, 99.8%, 100.2%} for RA and {-0.36%, -0.21%, -0.29%, 103.2%, 99.7%, 100.1%} for LB.

Was presented Monday Oct. 16 1715 (chaired by JRO).

Investigate in EE.

[JVET-AF0250](https://jvet-experts.org/doc_end_user/current_document.php?id=13513) Crosscheck of JVET-AF0134 (AHG12: AMVP with SbTMVP mode) [F. Pu (Dolby)] [late]

[JVET-AF0135](https://jvet-experts.org/doc_end_user/current_document.php?id=13393) AHG12: Additional SbTMVP candidates [R.-L. Liao, J. Chen, Y. Ye, X. Li (Alibaba)]

The contribution proposes to add additional SbTMVP candidates whose reference picture indices are set equal to 0 to subblock merge list. At most 4 additional SbTMVP candidates are added to the subblock merge list and are sorted together with existing SbTMVP candidates and other affine candidates. It is reported that on top of ECM-10.0, the overall coding performance and runtime for {Y, U, V, EncT, DecT, VmPeak} is {-0.05%, -0.03%, -0.01%, 100.0%, 101.5%, 100.4%} for RA and {-0.09%, -0.26%, -0.08%, 100.5%, 102.7%, 100.2%} for LB.

Was presented Mon. 16 Oct. 1735 (chaired by JRO)

It is noted that this is a relative simple change

Tradeoff enc time vs. compression not attractive. Further study recommended for improvements.

[JVET-AF0238](https://jvet-experts.org/doc_end_user/current_document.php?id=13501) Crosscheck of JVET-AF0135 (AHG12: Additional SbTMVP candidates) [L. Zhao (Bytedance)] [late]

[JVET-AF0160](https://jvet-experts.org/doc_end_user/current_document.php?id=13418) Non-EE2: On DMVR Extensions [M. Salehifar, Y. He, K. Zhang, L. Zhang (Bytedance)]

This proposal introduces improvements for DMVR. Specifically, DMVR is extended to non-equal POC distance cases and 16×16 BDOF DMVR is added. Also, a mean removed formula is introduced for BDOF DMVR.

Simulation results on top of the ECM-10.0 are reported as:

RA: -0.17%, -0.20%, -0.24%, 102.1%, 102.9%.

It was asked what is the gain for each of the three elements in the contribution (non-equal POC distance, 16x16 BDOF DMVR, and mean-removed BDOF DMVR). The proponent said each element approximately contributed 1/3 of the gain.

It was commented that the removal of equal distance constaint for when DMVR is applied would have some impact on complexity, as the process would be applied to many more blocks than in ECM-10.

For the non-equal POC distance cases, only the first two stages of DMVR process is applied, i.e. BDOF DMVR is not applied.

It was asked if putting some constraint on the non-equal POC distance case (e.g. applying a threshold of how large the POC distance can be) would be beneficial. Proponent is requested to bring results for a threshold that provides desirable tradeoff.

The proposed mean-removal formula is applied to all DMVR stages.

It was commented that although gain is somewhat interesting, the increase in encoder and especially decoder runtime is not desirable. Therefore the proponent is encouraged to investigate how to reduce runtime increase.

Several experts expressed interest in testing the proposal in EE, and requested to study each element of the proposal separately.

Investigate in EE, with subtests on each of the elements and combination of all 3 elements. For the non-equal POC distance element, two subtests are to be established: the original design in this proposal, and an alternative design with a threshold value that provides more desirable tradeoff.

[JVET-AF0162](https://jvet-experts.org/doc_end_user/current_document.php?id=13420) AHG12: Fixes to template matching [J. Chen, R.-L. Liao, X. Li, Y. Ye (Alibaba)]

This contribution reports and fixes the issues in the design of template matching in ECM-10.0. It is reported that the proposed fixes give {-0.01%, 0.05%, 0.03%} in RA and {-0.05%, 0.09%, -0.08%} in LD overall coding gain with no encoding and decoding time increase.

Currently some cost is computed in TM that is never used. The contribution shows that by using the value of this cost some minor gain is achieved in LB.

Not relevant for action. A ticket might be filed to remove the computation of the value that is not used without change in performance.

[JVET-AF0264](https://jvet-experts.org/doc_end_user/current_document.php?id=13528) Crosscheck of JVET-AF0162 (AHG12: Fixes to template matching) [Y. Wang (Bytedance)] [late]

[JVET-AF0168](https://jvet-experts.org/doc_end_user/current_document.php?id=13426) Non-EE2: Geometric partitioning mode with affine prediction [K. Zhang, Z. Deng, L. Zhang (Bytedance)]

This proposal presents a method of geometric partitioning mode (GPM) with affine prediction. A GPM partition can be predicted by affine motion compensation (AMC) or non-affine motion compensation (non-AMC), indicated by a flag. A GPM partition predicted by AMC can be blended with the other GPM partition predicted by AMC, non-AMC, or intra-prediction. Simulation results based on ECM-10.0 are reported as:

RA:{ -0.19%, -0.18%, -0.20%, 101.2%, 99.8%}.

It was asked when ARMC is applied, the proponent said that the uni-prediction AMC candidates are formed first, and then ARMC is applied to these candidates.

It was asked why only uni-prediction AMC is used, the proponent said conceptually bi-prediction can be applied as well, but may not provide desirable performance vs. complexity tradeoff.

It was commented that the performance vs. runtime tradeoff looks good.

Seveeral experts expressed interest in conducting EE test based on this proposal.

Investigate in EE.

[JVET-AF0191](https://jvet-experts.org/doc_end_user/current_document.php?id=13452) Non-EE2: Enhancements on local illumination compensation [X. Xiu, C. Ma, N. Yan, H.-J. Jhu, C.-W. Kuo, W. Chen, X. Wang (Kwai)]

This contribution proposes two aspects to further enhance the coding efficiency of local illumination compensation (LIC). In the first aspect, one non-local illumination compensation (NLIC) is proposed, where extra linear parameters (i.e., scale and offset) are derived by minimizing the difference between the reconstruction and prediction samples of inter blocks. The derived linear models can be inherited to generate the prediction samples of other inter CUs through merge mode. In the second aspect, two extra LIC modes, which only use the reconstructed samples either above or left to one CU for calculating the LIC parameters, are introduced for the LIC. The two modes are only applied to uni-predicted AMVP CUs. Simulation results show that the proposed schemes reportedly provide BD-rate savings of 0.12%% and 0.27%% for RA and LDB configurations with enc/dec run-time impacts of 102.0%/ 102.7% and 102.1%/ 103.8%, respectively.

It was commented that LIC with multiple templates was tested before in EE, but this proposal includes further restrictions on when LIC with multiple templates can be applied, i.e. this is activated for only uni-prediction AMVP blocks.

It was commented that the first element (non-local IC) mainly impacts the decoder runtime, and the second element (multi-template LIC) mainly impacts the encoder runtime. It was commented that it would be desirable to reduce the encoder runtime from the second element.

It was commented that for the multi-template LIC element, the additional two templates are just the top and left portions of the existing template, no additional samples are used in these two templates.

For the first element (non-local IC), it was commented by the proponent that although NLIC includes adjacent samples as well as non-adjacent samples, most of the gain comes from the non-adjacent samples. The same downsampling derivation as in the ECM-10 is used on the non-adjacent samples.

It was commented that the first element (non-local LIC) provides 0.08% gain in RA for 2.7% increase in decoder runtime. The proponent said that the increase in decoder runtime is mainly due to the reordering step in the first element. Currently this step is not optimized (up to 6 candidates are used). The proponent claimed that further optimization of this step (e.g. reducing the maximum number of candidates) should reduce decoder runtime without too much impact on performance.

Mulitple experts commented that this proposal provides interesting performance gain and should be studied in EE.

Investigate in EE, test the two elements separately, also reduction in encoder and decoder runtime is requested.

[JVET-AF0194](https://jvet-experts.org/doc_end_user/current_document.php?id=13455) Non-EE2: On LIC flag in merge mode [Y. Zhang, V. Seregin, H. Wang, Z. Zhang, C.-C. Chen, H. Huang, M. Karczewicz (Qualcomm)]

This contribution proposes to signal LIC flag for inter-prediction merge modes and to apply BDOF for bi-predicted LIC. On top of ECM-10.0, the simulation results are reported as follows.

RA: -0.10% / 0.09% / 0.02%; 102.4% enc, 100.0% dec

LB: -0.20% / -0.12% / 0.05%; 102.8% enc, 100.9% dec.

The proposal contains two orthogonal aspects, 1) signal LIC flag, and 2) apply BDOF to bi-predicted LIC (without changes to BDOF conditions or the BDOF process itself).

The results shown are the combination test results for both elements. The proponent reported that majority (~70% of gain) of the gain is from the first element, and the rest is from the second element.

EE2-3.2, which determine the LIC flag based on template matching cost, was adopted at this meeting. It was asked if there is still any additional gain for the first element of this proposal on top of EE2-3.2. The proponent reported that partial results from internal testing shows some additional gain (though the gains are somewhat reduced).

It was commented that there could be some interactions between this proposal and JVET-AF0191, and it would be interesting to study them together.

It was asked what could be the cause for the encoder runtime increase. The proponent said that there is no additional RDO to determine the LIC flag. Instead, this proposal applies more SATD calcuations on the merge candidates to determine the value of the LIC flag, and this is the main cause for encoder runtime increase. Software is not optimized in the current implementation, and it is reported by the proponent that further software optimization should bring down encoder runtime.

Investigate in EE, test the two elements separately, and reduction in encoder runtime is requested to be studied.

[JVET-AF0200](https://jvet-experts.org/doc_end_user/current_document.php?id=13461) Non-EE2: Extension of local illumination compensation [Y. Wang, K. Zhang, Y. He, H. Liu, L. Zhang (Bytedance)]

This contribution presents two extensions of local illumination compensation (LIC). In Aspect #1, LIC with multiple templates is proposed. Besides using a template comprising neighbouring samples both top to and left to the current block, two new LIC modes using top-only or left-only template are proposed to derive the parameters of LIC. In Aspect #2, LIC with slope adjustment is proposed, in which an adjustment parameter is used to modify parameters of LIC. On top of ECM-10.0, simulation results of the proposed method are estimated as below:

Aspect #1:

RA\*: {-0.06%, -0.07%, -0.04%; 101.4%, 98.1%};

LDB\*: {-0.09%, 0.11%, -0.01%; 101.6%, 99.4%} (Average of Class C and Class E).

Aspect #2:

RA\*: {-0.05%, -0.08%, -0.14%; 101.2%, 98.0%};

LDB\*: {-0.07%, -0.18%, 0.19%; 99.9%, 99.0%} (Average of Class C and Class E).

It was asked how the multi-template aspect is related to the multi-template part proposed in JVET-AF0191. Both proponents commented that the normative part is likely the same, and different might only be in the encoder implementation. The proponents of JVET-AF0191 and JVET-AF0200 are requested to work on a unified encoder implementation for the EE test on multi-template LIC.

It was commented by multiple experts that the second element, which adjusts the slope for LIC, looks interesting (JVET-AF0116 proposes to make slope adjustment to IBC).

Several experts expressed interest in studying the second element in EE as well.

Investigate in EE, test multi-template LIC (a unified encoder is to be tested) and slope adjustment for LIC. Also combination tests with other LIC-related EE tests are requested to be conducted.

[JVET-AF0301](https://jvet-experts.org/doc_end_user/current_document.php?id=13565) AHG12: SbTMVP with MMVD [L.-F. Chen, R. Chernyak, X. Zhao, X. Xu, S. Liu (Tencent)] [late]

In this proposal, it is proposed to apply the MMVD on SbTMVP candidate in subblock MMVD merge list. The syntax element, MMVD index, is signaled to indicate the MMVD offset for the SbTMVP merge candidate, and this MMVD offset is used to derive the offset value of the displacement vector (DV) for each SbTMVP-MMVD candidate. The final DV for each SbTMVP-MMVD candidate is calculated from the DV of SbTMVP merge candidate with the selected offset. Different subblock-based motion field could be obtained by using the different DV offset at different MMVD offset position. In this proposal, the maximum candidate list size is 16. The concept of MMVD offset index in this proposal is also utilized in JVET-AF0134. Compared with JVET-AF0134, signaling these SbTMVP-MMVD index at subblock merge mode is the major difference, other design aspects in JVET-AF0134 could be harmonized in this proposal. In this proposal, two different tests are tested and the overall simulation results of these two tests on top of ECM-10.0 reference software are reported as below.

Test 1:

* RA: -0.xx%/-0.xx %/-0.xx % with xxx.x% EncT and xxx.x% DecT
* LB: -0.xx%/-0.xx %/-0.xx % with xxx.x% EncT and xxx.x% DecT

Test 2:

* RA: -0.xx%/-0.xx %/-0.xx % with xxx.x% EncT and xxx.x% DecT
* LB: -0.xx%/-0.xx %/-0.xx % with xxx.x% EncT and xxx.x% DecT

Was presented Monday Oct. 16 1715 (chaired by JRO).

At the time of presentation, largely incomplete results. Similar concept as in JVET-AF0134, but comparably worse tradeoff. Further study recommended for improvements.

#### Cross Component Prediction (3)

[JVET-AF0083](https://jvet-experts.org/doc_end_user/current_document.php?id=13331) Non-EE2: Enhancements on CCP merge [H. Huang, Y. Yu, H. Yu, D. Wang (OPPO), Z. Deng, K. Zhang, L. Zhang (Bytedance)]

This contribution proposes a fusion mode for the CCP merge mode, in which the final prediction is generated by a weighted sum between the CCP-merge prediction and the MM-CCCM prediction or the DIMD prediction. A fusion type flag is signalled to indicate whether the MM-CCCM or DIMD is selected. Additionally, CCP merge mode is harmonized with LB-CCP by including the LB-CCP flag in the CCP merge candidate. On top of ECM-10.0, the performance of the proposed contribution is reported as follows:

AI {-0.03%, -0.65%, -0.74%, 100.6%, 100.8%}.

The proposal proposes two orthogonal aspects, 1) CCP merge fusion, 2) inherit LB-CCP flag.

For the first aspect, two syntax elements are added, one flag to indicate whether to use the proposed CCP merge fusion mode, and another flag to indicate whether to perform fusion with MM-CCCM or with DIMD. The values of these two flags are determined using RDO.

For the second aspect, it was asked how much gain from inheriting the LB-CCP flag could be obtained, the proponent did not yet test it separately.

Cross checher confirms that results are matched, and makes a comment that over 1% average chroma gain is obtained for class A1.

Investigate in EE. Test the two aspects separately.

[JVET-AF0299](https://jvet-experts.org/doc_end_user/current_document.php?id=13563) Crosscheck of JVET-AF0083 (Non-EE2: Enhancements on CCP merge for chroma intra coding) [C. Zhou (vivo)] [late]

[JVET-AF0114](https://jvet-experts.org/doc_end_user/current_document.php?id=13361) AHG12: Local-Boosting on Cross-Component Merge Mode [H. Qin, K. Ding, Z. Xu (TCL)]

It is proposed to enable local-boosting cross-component prediction (LBCCP) in cross-component merge mode. At cross-component merge mode candidate reordering stage, for each MM-CCLM/MM-CCCM candidate, a filter is tried on template cost calculation in order to determine whether LBCCP is used on the corresponding candidate.

The experimental results over ECM-10.0 are summarised as follows:

**AI:** -0.02% (Y), -0.01 % (U), -0.04 % (V), 100.2 % (EncT), 100.7 % (DecT), 100% (VmPeak)

It was commented that the current gains are rather small.

It was commented that this contribution is related to the second aspect in JVET-AF0083 (inheriting LB-CCP flag). The second aspect in JVET-AF0083 is a straighforward inheritence, and this proposal uses template matching cost to determine to use LB-CCP.

Further study is recommended, esp. in relationship with the EE test on inheriting LB-CCP flag.

[JVET-AF0176](https://jvet-experts.org/doc_end_user/current_document.php?id=13434) Non-EE2: Decoder Derived Cross-Component Prediction [Y.-J. Chang, P.-H. Lin, V. Seregin, J.-L. Lin, M. Karczewicz (Qualcomm)]

This contribution proposes a decoder derived cross-component prediction (DDCCP) to derive the best CCP mode by comparing the template costs between the evaluated CCPs. A decoder-derived CCP fusion mode is also introduced to derive CCP modes for fusion selected based on the template costs. The simulation results for 3 different test setting on top of ECM-10.0 are summarized as follows:

Test 1: CCP list constructed by {CCCM, CCCMwMDF, GLCCCM} with the constraint of (CU\_Width \* CU\_Height > 64)

AI: -0.08% (Y), -0.34% (Cb), -0.42% (Cr), 101.6% (EncT), 103.4% (DecT)

Test 2: CCP list constructed by {CCCM, CCCMwMDF, GLCCCM} with the constraint of (CU\_Width \* CU\_Height > 8)

AI: -0.10% (Y), -0.34% (Cb), -0.34% (Cr), 102.3% (EncT), 105.5% (DecT)

Test 3: CCP list constructed by {CCLM, CCCM, CCCMwMDF, GLCCCM, LBCCP} with the constraint of (CU\_Width \* CU\_Height > 8)

AI: -0.14% (Y), -0.16% (Cb), -0.16% (Cr)

It was noted that the abstract in JVET-AF0176r1 is incorrect. It was requested for the proponent to revise the abstact and upload a revision.

The proposed method added two additional CCP modes, the DDCCP and DDCCP with fusion.

Compared to Test 1, which applies the two proposed CCP modes on blocks with more than 64 luma samples, Test 2 applies the same method on more blocks (blocks with more than 8 luma samples are affected). Comparing the performance of Test 1 and Test 2, there is an additional luma gain of 0.02% in AI, and additional 2% decoder runtime increase.

Compared to Test 2, which uses {CCCM, CCCMwMDF, GLCCCM} for list construction, Test 3 uses two more modes {CCLM, CCCM, CCCMwMDF, GLCCCM, LBCCP} for list contruction. The block size restriction in Test 2 and Test 3 are the same. Comparing the performance of Test 2 and Test 3, there is additional 0.04% luma gain in AI, but chroma gain is reduced by 0.18%. Decoder runtime is not available at the time of presentation for Test 3.

The attached PowerPoint file included some RA results that are not available in the contribution itself. The RA results show chroma loss in class B.

It was commented that decoder runtime increase of 3-5% is not negligible. The proponent commented that the current implementation on the decoder side contained redundant operations, which should be removed during EE. This could result in a lower decoder runtime increase.

It was asked how many predictions are used in the DDCCP fusion process, and the answer is 2 predictions.

It was requested to test the two proposed DDCCP modes separately.

In terms of block size restriction and modes to be included in the CCP list, the proponent is requested to bring one configuration that provides the best performance vs. runtime tradeoff.

Investigate in EE, two subsets with the first one using the DDCCP mode only, and the second subtest using both DDCCP and DDCCP fusion modes. These two subtests should be based on the same configuration of block size restriction and modes in CCP list.

#### In-Loop Filters (4)

[JVET-AF0169](https://jvet-experts.org/doc_end_user/current_document.php?id=13427) Non-EE2: Adaptive clipping with signaled lower and upper bounds [K. Cui, Z. Zhang, H. Huang, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes a method of applying adaptive clipping with signaled lower and upper bounds. For each picture, the minimum and maximum values of the Luma channel, which are derived from the original picture or MCTF pre-filtered picture, are signaled and used to perform clipping at the reconstruction stage. The proposed method reports the following BD-rate performance on top of ECM-10.0:

AI: {-0.06%, 0.06%, -0.02%}, EncT 100.4%, DecT 100.3%, VmPeak 100.0%.

RA: {-0.08%, 0.01%, -0.11%}, EncT 100.4%, DecT 100.6%, VmPeak 100.0%.

The proposed method sends min/max values per picture. It covers two cases:

* For MCTF filtered pictures, the min/max values are obtained from the filter picture
* For original pictures (i.e. MCTF is not applied), the min/max values are obtained from the original picture

Additionally, if LMCS is enabled, the signaled values are mapped to the LMCS-mapped domain using the LMCS LUT, and applied in the LMCS mapped domain.

It was commented that for the above two cases, the min/max values are always obtained from the encoder input.

Two clipping operations are applied, before loop filtering, and once again after loop filtering. Both operations use the same min/max values.

It was asked if the min/max values could always be obtained from the original picture (even if MCTF is applied). This hasn’t been tested, but could be tested. The proponent commented that the performance gain could be due to sticking with the actual dynamic range of the input picture, rather than a dynamic range defined by bit depth. From this point of view, it would be interesting to also test getting the min/max values always from the original picture.

Currently, the proposed clippings are applied only to luma. Chroma is not changed.

It was commented that there is performance loss for some sequences in RA config. The proponent said this could be related to signaling overhead of sending the min/max values.

This is also applicable to LDB config, although the LDB results are not yet available at time of presentation.

Multiple experts find this proposal interesting.

Investigate in EE, with two variations, i..e. obtaining min/max values as proposed in JVET-AF0169, and obtaining min/max values always from the original image.

[JVET-AF0178](https://jvet-experts.org/doc_end_user/current_document.php?id=13436) Non-EE2: Applying fixed filters to output of the Gaussian filter [N. Hu, M. Karczewicz, V. Seregin (Qualcomm)]

In ECM-10.0, two Laplacian classifier-based fixed filters in adaptive loop filter (ALF) are applied to input to deblocking filters (DBF) and input to ALF. A third fixed filter, which is known as a Gaussian filter, is applied to input to DBF. The outputs of the three fixed filters are then used to derive ALF output. In this contribution, the first two fixed filters are proposed to apply to output of the Gaussian filter. On top of ECM-10.0 with common test condition, simulation results of the proposed method are reported as follows.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | BD-rate Y | BD-rate U | BD-rate V | EncT | DecT | VmPeak |
| All intra | -0.07% | 0.00% | 0.00% | 100.7% | 104.0% | 100.1% |
| Random access | -0.11% | 0.02% | -0.04% | 100.2% | 102.8% | 100.1% |
| Low delay B | -0.12% | 0.13% | -0.26% | 100.2% | 103.0% | 100.1% |

It was proposed to add Gaussian filtered output samples to fixed filters with Laplacian classifiers. This would require addionally performing Gaussian filter in order to obtain the output samples.

Before the proposed change, around 100 multiplications per ALF invocation, this proposal would add another 40 multiplications per ALF invocation. Additionally, this would also increase the number of coefficients for fixed filters by around 40%.

It was commented that this doesn’t seem to be good return (i.e. gain) on investiment (i.e. additional computation and memory to store coefficients).

No change to Gaussian filters is made. The fixed filters are retrained, that is, in order to take additional Gaussian-filtered inputs.

It was commented that the amount of actual changes (completely new fixed filters, additional computation, and additional memory required) is much more than the concept (of adding Gaussian-filtered output as input to fixed filters) itself.

It was commented that this also causes 3-4% increase in decoder runtime.

Retraining of fixed filters included CTC sequences. It was commented that this might create overfitting.

Futher study is recommended.

[JVET-AF0179](https://jvet-experts.org/doc_end_user/current_document.php?id=13437) Non-EE2: Fixed filter for Chroma ALF [N. Hu, M. Karczewicz, H. Wang, V. Seregin (Qualcomm)]

In ECM-10.0, for luma component, two Laplacian and variance classifier-based fixed filters in adaptive loop filter (ALF) are applied to input to deblocking filters (DBF) and input to ALF. The outputs of the two fixed filters are then used to derive ALF output. In this contribution, a Laplacian and variance classifier-based fixed filter is applied to chroma components. On top of ECM-10.0 with common test condition, simulation results of the proposed method are reported as follows (RA and LDB results are estimated by copying one data point from anchor).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | BD-rate Y | BD-rate U | BD-rate V | EncT | DecT | VmPeak |
| All intra | -0.03% | -1.62% | -2.03% | 100.8% | 102.3% | 100.3% |
| Random access | -0.03% | -1.44% | -1.87% | 101.0% | 101.8% | 100.1% |
| Low delay B | -0.08% | -1.52% | -1.55% | 100.8% | 102.1% | 100.2% |

In ECM-10.0, there are no fixed filters for chroma. This contribution proposes to add 8 sets of fixed chroma filters. Cb and Cr share the 8 sets of filters, classification dictates which fixed filter is used for Cb and which filter is used for Cr, which means different filters could be applied to Cb and Cr of the same block.

Sample-based classification is used. One classifier is used.

The additional memory required to store the chroma fixed filters is less than 0.5MB.

Around 2% decoder runtime increase.

In ECM-10, 20 multiplications per chroma sample; in this proposal, 55 additional multiplications per chroma sample are needed.

This contribution provides insights into the benefits to applying fixed filters to the chroma components. The performance on chroma is obvious, and could be interesting to have.

Several experts expressed interest in the concept of fixed filters for chroma, and supported studying this in EE.

Training included CTC and non-CTC sequences.

Investigate in EE. It is requested to bring fixed filters that are training on non-CTC sequences.

[JVET-AF0198](https://jvet-experts.org/doc_end_user/current_document.php?id=13459) Non-EE2: Coefficient Precision Adjustment for ALF [W. Yin, K. Zhang, L. Zhang (Bytedance), N. Hu, M. Karczewicz, V. Seregin (Qualcomm)]

In this contribution, the coefficient precision of luma ALF is increased from 8-bit to 9-bit.

On top of ECM-10.0, simulation results of the proposed methods are reported as below:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **AI** | | | | | **RA** | | | | | **LB** | | | | |
|  | Y | U | V | EncT | DecT | Y | U | V | EncT | DecT | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Class B | -0.03% | -0.01% | -0.01% | 100.7% | 99.4% | -0.07% | 0.02% | -0.10% | 99.4% | 99.3% |  |  |  |  |  |
| Class C | -0.07% | -0.01% | -0.01% | 101.2% | 101.2% | -0.10% | -0.07% | -0.19% | 99.6% | 100.2% | -0.06% | -0.22% | -0.09% | 100.3% | 99.6% |
| Class E | -0.04% | -0.02% | -0.02% | 99.7% | 99.4% |  |  |  |  |  | -0.04% | 0.12% | 0.41% | 99.0% | 100.7% |
| **Overall** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Class D | -0.06% | 0.02% | 0.02% | 100.4% | 99.7% | -0.13% | 0.08% | -0.15% | 99.4% | 100.0% | -0.06% | 0.38% | -0.56% | 99.7% | 100.3% |
| Class F |  |  |  |  |  | -0.04% | 0.04% | -0.05% | 99.5% | 100.2% |  |  |  |  |  |

This contribution proposes to increase coefficient precision from 8-bit to 9-bit for non-fixed filters, and for luma only.

Fixed filters are not changed.

It was asked if chroma could benefit from this change. The proponent said internal results showed no gain for chroma.

At time of presentation, only partial results are available. Cross checher reported that it is unlikely to finish all simulation before the end of this meeting.

While the current results show some small gain at no additional cost, there is no urgency to adopt the proposal without seeing full results.

Investigate in EE.

[JVET-AF0298](https://jvet-experts.org/doc_end_user/current_document.php?id=13562) Crosscheck of JVET-AF0198 (Non-EE2: Coefficient Precision Adjustment for ALF) [C. Ma (Kwai)] [late]

#### Entropy coding and transform coefficient coding (2)

[JVET-AF0164](https://jvet-experts.org/doc_end_user/current_document.php?id=13422) AHG12: CABAC inter/intra model switch [F. Lo Bianco, F. Galpin, C. Salmon-Legagneur (InterDigital)]

This contribution proposes to

1. Use intra CABAC models to encode residual information of CU in inter slices
2. Use inter CABAC models to encode residual information of some CU in intra slices

This contribution provides the individual test results for the following configurations:

Test 1: Use intra CABAC models to encode residual information of intra-predicted CU in inter slices

RA:

|  |  |  |  |
| --- | --- | --- | --- |
| Class B | 0.03% | -1.58% | -1.13% |
| Class C | 0.06% | -1.17% | -1.21% |

LDB:

|  |  |  |  |
| --- | --- | --- | --- |
| Class C | -0.08% | -0.56% | -0.40% |
| Class E | -0.09% | 0.72% | -0.07% |

Test 2: As in Test 1, and also use inter CABAC models to encode residual information of IBC or IntraTMP-predicted CU in intra slices

AI:

|  |  |  |  |
| --- | --- | --- | --- |
| Class B | -0.06% | 0.08% | 0.02% |
| Class C | -0.04% | 0.13% | 0.09% |
| Class E | -0.10% | 0.05% | -0.18% |

RA:

|  |  |  |  |
| --- | --- | --- | --- |
| Class B | -0.04% | -1.56% | -1.13% |
| Class C | 0.04% | -1.09% | -1.39% |

LDB:

|  |  |  |  |
| --- | --- | --- | --- |
| Class B | -0.03% | -0.49% | -1.00% |
| Class C | -0.11% | -0.39% | -0.58% |
| Class E | 0.00% | 1.31% | -0.95% |

Test 3: As in Test 2, but QtCbf flag is excluded from the syntax elements for which a model switch is allowed

AI:

|  |  |  |  |
| --- | --- | --- | --- |
| Class B | -0.08% | 0.07% | -0.15% |
| Class C | -0.05% | 0.10% | 0.12% |
| Class E | -0.11% | 0.00% | -0.02% |

RA:

|  |  |  |  |
| --- | --- | --- | --- |
| Class B | -0.12% | 0.00% | -0.05% |
| Class C | -0.09% | -0.08% | -0.16% |

LDB:

|  |  |  |  |
| --- | --- | --- | --- |
| Class C | -0.05% | -0.28% | -0.30% |
| Class E | -0.07% | 1.19% | 0.18% |

It was noted that the presentation was not uploaded. The proponent was requested to upload a revised package with presentation as an attachment.

The encoder runtime showed significant increase, but this was likely due to inaccurate timing from the grid. The proponent commented that no encoding time increase should be expected.

This would double the number of contexts for coefficient coding.

Compared to Test 2, Test 3 excludes QtCbf flag from the proposed method. The main motivation is to maintain a better balance between luma and chroma gain.

Investigate in EE, subtests to include 3 variations:

* Use intra CABAC models to encode residual information of intra-predicted CU in inter slices
* As above, and also use inter CABAC models to encode residual information of IBC or IntraTMP-predicted CU in intra slices
* As above, but QtCbf flag is excluded from the syntax elements for which a model switch is allowed

Also conduct combination tests with the EE tests on CABAC context retraining.

[JVET-AF0185](https://jvet-experts.org/doc_end_user/current_document.php?id=13443) Non-EE2: Transform Coefficient Coding [P. Nikitin, M. Karczewicz, M. Coban, V. Seregin (Qualcomm)]

This contribution proposes a change to the number of regular coded bins per coefficient in transform residual coefficient coding. It is reported that on top of ECM-10.0, the overall coding performance impact for {Y, U, V} in AI, RA configurations.

AI {-0.09% Y, -0.04% U, -0.13% V, 100.7% EncT, 100.0% DecT, 99.9% VmPeak}

RA {-0.07% Y, -0.02% U, -0.04% V, 99.4% EncT, 98.5% DecT, 99.9% VmPeak}

The proposal only changes transform coefficient coding, it does not impact the blocks in transform skip mode.

It was asked why 8 gtX flag. The proponent did internal testing and this seems to provide the best results.

It was asked why the location of the parity flag was moved, and whether there are results for not moving the parity flag. The proponent hasn’t teste dit.

It was suggested that if EE test is performed, it should be tested on lower QP values.

It was commented that increasing the number of gtX flag might decrease the throughput of transform coefficient coding.

Among the 8 gtX flags, the first 3 flags have their own contexts, and the remaining 5 flags share the same context.

It was requested for the proponent to upload a revised document that provides further details on the design of context model for each flag.

For gt1 and gt3 flags, which are existing flags in ECM-10, the context initialization values have not been changed in this proposal.

Investigate in EE, results on all configurations, CTC QP, as well as QP 17 should be provided. Also test in combination with the method in JVET-AF0164.

#### RPR (1)

[JVET-AF0058](https://jvet-experts.org/doc_end_user/current_document.php?id=13306) AHG12: GOP-based RPR encoder control for ECM [K. Andersson, J. Ström, R. Yu, W. Ahmad, P. Wennersten (Ericsson)]

GOP-based RPR encoder control has been available for VTM since VTM-19.0. It selects coding in a reduced resolution when impact of re-scaling performance between source resolution and reduced resolution is expected to be limited. ECM does not have such functionality. This contribution proposes to add the GOP-based RPR encoder control from VTM to ECM to enable further experimentation of RPR functionality in ECM.

Objective performance (BDR) for RA 4K resolutions and QP 40, 43, 47 and 50 is reported as:

ECM-10.0 with GOP-based RPR versus ECM-10.0: Y: -3,25%, U 5,95%, V 5,15%

VTM-22.0 with GOP-based RPR versus VTM-22.0: Y: -8,92%, U -3,38%, V -5,87%

Objective performance (BDR) for RA all resolutions and QP 40, 43, 47 and 50 is reported as:

ECM-10.0 with GOP-based RPR versus ECM-10.0: Y: -1,48%, U 2,37%, V 1,54%

VTM-22.0 with GOP-based RPR versus VTM-22.0: Y: -4,08%, U -2,82%, V -3,92%

It is noted that the benefit of using the GOP-based RPR encoder control in ECM is less compared to its usage in VTM but that it still shows promising coding gains at challenging bitrates.

It is also asserted that subjective quality can be improved by the proposal.

The QP values experimented here are not in RPR CTC.

GOP-based RPR in VTM is provided as an optional encoder control but not enabled in RPR CTC in the context of VTM.

If performance gain is the goal, then CTC should be used to conduct experiments.

Further study is recommended.

### CTC for EE2/ECM and general ECM improvements (7)

Contributions in this area were discussed at 1400–1610 on Tuesday 17 Oct. 2023 (chaired by Y. Ye).

[JVET-AF0090](https://jvet-experts.org/doc_end_user/current_document.php?id=13337) AHG6: Encoder memory profiling on ECM software [Y. Yasugi, T. Ikai (Sharp), R. Chernyak, Shan Liu (Tencent)]

In this contribution, the effect of the encoder memory reduction of the ECM software was analyzed by profiling encoders using Valgrind’s Massif profiler. It is reported that the encoder memory consumption was reduced by 36% at the version with commit-ID 9e651d0 (after ECM-10.0) compared to ECM-9.1.

It was asked if a comparison of VTM memory consumption vs. ECM memory comsumption was performed. This is currently not available, but it was agreed that such information would be very valuable. And therefore it is highly enoucraged for people to bring such information to the next JVET meeting.

After ECM-10.0 was tagged, a lot of work had been done to further reduce ECM encoder’s memory consumption. Such reduction is reflected in commit-ID 9e651d0.

JVET-AF0156 reports further detail on the work that has been done to reduce ECM’s meomory comsumption.

[JVET-AF0101](https://jvet-experts.org/doc_end_user/current_document.php?id=13348) AHG12: ECM software cleanup for decoder side intra predictions acceleration [Z. Fan, Y. Yasugi, T. Ikai (Sharp)]

This proposal proposes a software change to accelerate the ECM by addressing the multiple initialization of prediction image in three intra prediction methods: TMRL in JVET-AB0157, TIMD in JVET-W0123 , and sortedMPM in JVET-AD0085. The modification is software only and no changes in simulation results. The proposed modification reportedly shows 2% and 6% encoding and decoding time reduction in AI and 1% encoding and decoding time reduction in RA respectively.

Cross checker confirms the losslessness nature of this contribution, and reported 2% and 4% encoding and decoding in AI config, respectively. RA results are not yet completed at the time of presentation.

It was commented that such software-only changes could also be submitted as merge requests, rather than having a contribution and a full discussion at the JVET meeting.

It was commented that it is strongly encouraged for software coordinators and AHG chairs to acknowledge such merge reqests (i.e. those that are targeted toward encoding/decoding/memory consumption reduction without causing any algorithm-level changes) in the AHG report. This is agreed.

Decision (SW): adopt JVET-AF0101

[JVET-AF0212](https://jvet-experts.org/doc_end_user/current_document.php?id=13475) Crosscheck of JVET-AF0101 (AHG12: ECM software cleanup for decoder side intra predictions acceleration) F. Wang (OPPO) [late]

[JVET-AF0156](https://jvet-experts.org/doc_end_user/current_document.php?id=13414) AHG6: ECM software optimizations [F. Urban, T. Poirier, F. Le Léannec (InterDigital)]

Recent developments of ECM have increased its memory consumption to a critically large amount, especially when running multiple instances on a computer. In previous meetings, several contributions proposed memory profiling and software modifications to tackle this issue. To monitor the impact on memory consumption of contributions, memory consumption reporting is now part of performance reporting.

During the development of ECM 10.0, memory reduction was achieved, with about 5 GB memory saving for the worst-case scenario (class A RA).

This contribution presents a few more optimizations that have been pushed to the ECM software after the ECM 10.0 was released. Compared to ECM 10.0, the worst-case memory consumption has been further reduced by:

* (AI, RA, LDB): 1.0 GB, 4.0 GB, 3.3 GB

and the average encoding time has been reduced to

* (AI, RA, LDB): 94.2%, 95%, 95%.

On the decoder side, memory is also reduced by up to

* (AI, RA, LDB): 0.15 GB, 1 GB, 0.15 GB

and the average decoding time has been reduced to

* (AI, RA, LDB): 83.9%, 82.7%, 83.6%.

Comprehensive memory reduction efforts have been made since ECM-10.0 was tagged, the following table provides a few examples and the respective amount of reduction.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Optimization | RA peak memory saving (MB and %) | | | | | AI peak memory saving (MB and %) | | | | |
| Class A | Class B | Class C | Class D | Class F | Class A | Class B | Class C | Class D | Class F |
| MAX\_CTU | -61 | -2857 | -2850 | -2856 | -3049 | -4 | -2856 | -2855 | -2856 | -2897 |
| **0%** | **-30%** | **-54%** | **-65%** | **-43%** | **0%** | **-40%** | **-60%** | **-68%** | **-56%** |
| CS | -1636 | -395 | -97 | -20 | -220 | -130 | -35 | -9 | -2 | -12 |
| **-9%** | **-4%** | **-2%** | **-0%** | **-3%** | **-1%** | **0%** | **0%** | **0%** | **0%** |
| PIC\_Buf | -1523 | -397 | -84 | -21 | -118 |  |  |  |  |  |
| **-8%** | **-4%** | **-1%** | **0%** | **-1%** |  |  |  |  |  |
| sign\_pred | -695 | -579 | -551 | -544 | -592 | -596 | -560 | -545 | -544 | -546 |
| **-4%** | **-6%** | **-10%** | **-12%** | **-8%** | **-7%** | **-8%** | **-12%** | **-13%** | **-10%** |
| CU | -246 | -192 | -187 | -175 | -198 | -246 | -200 | -182 | -174 | -179 |
| **-1%** | **-2%** | **-4%** | **-4%** | **-3%** | **-3%** | **-3%** | **-4%** | **-4%** | **-3%** |
| PU | -74 | -8 | -7 | -7 | -14 | -9 | -10 | -7 | -7 | -6 |
| **0%** | **0%** | **0%** | **0%** | **0%** | **0%** | **0%** | **0%** | **0%** | **0%** |
| MCTF | -93 | -9 | -1 | 0 | -27 |  |  |  |  |  |
| **-1%** | **0%** | **0%** | **0%** | **-0%** |  |  |  |  |  |
| DBF | -84 | -8 | 0 | -1 | -28 | -3 | -3 | 0 | -1 | -1 |
| **0%** | **0%** | **0%** | **0%** | **0%** | **0%** | **0%** | **0%** | **0%** | **0%** |

Though the efforts were mainly targeted at encoder memory reduction, the decoder had also benefited. Further, encoding and decoding times were also reduced, as noted above.

AHG6 plans to tag ECM-10.1 which contains these changes in the near future.

Some of the memory reduction efforts detailed here are also applicable to VTM. The software coordinators are working together to bring those changes that are relevant to VTM to reduce VTM’s memory consumption as well.

Similar efforts could also be done to the NNVC software (which is based on VTM).

It is noted that these efforts are very much appreciated. Further work is also encouraged.

[JVET-AF0287](https://jvet-experts.org/doc_end_user/current_document.php?id=13551) Crosscheck of JVET-AF0156 (AHG6: ECM software optimizations) [Y. Yasugi (Sharp)] [late]

[JVET-AF0165](https://jvet-experts.org/doc_end_user/current_document.php?id=13423) Ahg7/Ahg12: On VTM-11ecm anchor low-delay test conditions [F. Le Léannec, S. Puri, E. François, K. Naser (InterDigital)]

This contribution proposes to correct the VTM11-ecm test conditions, to align them with the ECM test conditions used in Ahg12.

In 31st JVET meeting, contribution JVET-AE0163 was adopted to significantly reduce the worst-case encoding time in low-delay configurations. This was achieved by reducing maximum MTT depth from 3 to 2 for inter slices with temporal depth greater or equal to 2 in class B.

However, this CTC modification was not ported to the test conditions of VTM-11ecm software, which is used as the anchor to evaluate ECM performances.

This contribution proposes to fix this CTC misalignment by using same encoder parameter modification for VTM-11ecm as for ECM.

This issue (of misalignment of CTC between ECM and VTM-11ecm) was raised by people involved in AHG7 testing.

For LB configurations, because of the desire to reduce the worst case runtime, we have previously adopted a number of “special” configuration parameters in our CTC. It was asked if we always align the “special” configuration parameters in VTM and ECM at the same time. The proponent reported that previously in a similar case, JVET-AA0098 attempted to reduce worst-case coding tim. The CTC modification in JVET-AA0098 was adopted for both ECM and VTM-11ecm anchor. This was confirmed by other experts.

The proposed change only requires a config file change, no changes to JVET-AF2017 is needed.

Decision (CTC): adopt JVET-AF0165.

[JVET-AF0177](https://jvet-experts.org/doc_end_user/current_document.php?id=13435) AHG6: Memory reduction for ECM encoder [N. Hu, V. Seregin, M. Karczewicz (Qualcomm), Y. Yasugi, T. Ikai (Sharp)]

To reduce the memory usage of ECM encoder, this contribution proposes to change the data type of adaptive loop filter covariance from double into float. On top of ECM-10.0 with common test condition, simulation results of the proposed method are reported as follows. For class A and B sequences, encoder memory usage is reduced by around 1 GiB.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | BD-rate Y | BD-rate U | BD-rate V | EncT | DecT | VmPeak |
| All intra | 0.00% | 0.01% | -0.01% | 98.7% | 100.0% | 89.7% |
| Random access | 0.00% | 0.05% | -0.01% | 98.8% | 100.4% | 94.1% |
| Low delay B | -0.02% | 0.29% | -0.01% | 97.8% | 99.7% | 92.5% |

Cross checker reported that partial results are matched in terms of rate distortion performance and memory consumption.

A similar concept was previously proposed in JVET-Z0150, but no action was taken at the time since the memory consiumption reduction was not very obvious at the time.

It was commented that a recent ALF-related software optimization on VTM caused little difference for SDR but caused a bigger (though still slight) difference in performance for HDR. A software fix was applied in that case. It was suggested for the proponent to also try that fix within the context of this contribution as a side effort.

Decision (SW): adopt JVET-AF0177

[JVET-AF0303](https://jvet-experts.org/doc_end_user/current_document.php?id=13567) Crosscheck of JVET-AF0177 (AHG6: Memory reduction for ECM encoder) [C.-W. Kuo (Kwai)] [late]

[JVET-AF0201](https://jvet-experts.org/doc_end_user/current_document.php?id=13462) AHG6: On ECM SW memory consumption [R. Chernyak, M. Xu, S. Liu (Tencent), Y. Yasugi, T. Ikai (Sharp)]

This document provides a memory consumption comparison between ECM9, ECM10 and one of the recent versions of ECM master branch. It is observed that significant improvement of the ECM memory consumption was performed after ECM9 release and worst case of encoder time memory consumption was reduced for more than 50%. This document also proposes to extend the current ECM encoder memory measurement mechanism to the decoder side.

This contribution pointed out an increase in SAO-related encoder memory consumption between ECM-9.1 and ECM-10.0.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **m\_cCuEncoder** | **m\_cEncAlf** | **m\_cEnc\_Sao** | **Other** | **Total** |
| **ECM9** | 0.7 | 21 | 0.5 | 0.2 | 22.4 |
| **ECM10** | 0.7 | 6.9 | 2.3 | 0.2 | 10.1 |
| **ECM10-9e651** | 0.2 | 6.9 | 2.3 | 0.1 | 9.5 |

It was asked what might have caused the big difference in SAO-related memory consumption. The proponent has not looked into this issue.

Another expert commented that this could be because more edge classifiers were adopted into CCSAO that caused the memory increase. The proponent of more edge classifiers (JVET-AE0151) reported a much smaller memory consumption increase based on internal test, and the increase observed by the JVET-AE0151 proponent was more in line with the numbers reported independently in JVET-AF0090.

It was commented that encoder memory consumption is a much bigger issue than decoder memory consumption. However, several experts commented that it would be a good practice to also track decoder memory consumption.

It was agreed to update the CTC document JVET-AF2017 to include a xls template that contains decoder memory comsuption information.

[JVET-AF0237](https://jvet-experts.org/doc_end_user/current_document.php?id=13500) AHG6: ECM encoder memory reduction [C.-W. Kuo, X. Xiu, W. Chen, H.-J. Jhu, N. Yan, C. Ma, X. Wang (Kwai)] [late]

In the current ECM, SAO and ALF statistics are stored in the encoder memory. A memory space is reserved for both SAO and ALF objects at the beginning of encoding. To reduce the ECM encoder memory, this contribution proposes to only allocate the SAO/ALF statistics memory when SAO/ALF encoding. Simulation results show no BD-rate change and around 1GB worst memory reduction for RA.

It was asked how this proposal interacts with JVET-AF0177. This proposal basically reduces the SAO+ALF memory consumption to max(SAO, ALF) consumption. Therefore, it is expected that although the two changes are not entirely additive, there should be some benefits to having both of them,

Cross checker noted that upon examination of the software changes, they could confirm the change is lossless in terms of rate distortion performance.

Decision (SW): adopt JVET-AF0237.

[JVET-AF0293](https://jvet-experts.org/doc_end_user/current_document.php?id=13557) Crosscheck of JVET-AF0237 (AHG6: ECM encoder memory reduction) [N. Hu (Qualcomm)] [late]

# High-level syntax (HLS) and related proposals (44)

## AHG9: SEI messages on NNPF aspects other than grouping (1)

Contributions in this area were discussed at 1330–1410 on Saturday 14 Oct. 2023 (chaired by JRO).

[JVET-AF0045](https://jvet-experts.org/doc_end_user/current_document.php?id=13289) AHG2/AHG9: On the use of NNPF SEI messages in AVC [M. M. Hannuksela (Nokia)]

JVET-AE1016 includes the draft text for changes of the AVC, including the new subclause D.3.3 specifying the use of the neural-network post-filter characteristics (NNPFC) and neural-network post-filter activation (NNPFA) SEI messages in AVC. This input document proposes the following changes to subclause D.3.3:

1. It is proposed to specify that the non-nested NNPFC and NNPFA SEI messages apply to decoded pictures resulting from the decoding process specified in Clauses 2 to 9 and to remove references to nuh\_layer\_id, which does not exist in AVC.
2. It is proposed to remove the constraint and the NOTE that refer to output\_flag, since output\_flag is present only in the NAL unit header SVC extension and is therefore meaningless unless subclause D.3.3 would be extended to apply to layer representations of SVC.
3. Since AVC does not define picture unit, it is proposed to replace the term picture unit with the term access unit in constraints related to the occurrence and content of NNPFC SEI messages within a picture unit.
4. It is proposed to fix the following syntax element and variable names:
   1. fp\_arrangement\_type 🡪 frame\_packing\_arrangement\_type
   2. SliceQpY 🡪 SliceQPY

Item 1 (nuh\_layer\_id not defined in AVC): Another option suggested is defining a variable layer\_id to generate the corresponding syntax element that is not existing in AVC.

Item 2 (output\_flag not defined in AVC): Could be resolved similarly by variable. The aspect that prefix NAL unit is not mandatory in AVC has been studied in that context.

Item 3 (picture unit not defined in AVC): In non-nested case (single layer), replacing it by “access unit” would be OK. Was resolved in DoCR.

Item 4: editorial – agreed.

It was discussed whether it would be appropriate to restrict usage of NNPF to single layer (not SVC/MVC). In that case, it would need to be explicitly stated. Text would be simpler. However, it is commented that so far, usage of SEI messages had not been restricted to certain profiles.

The question whether there was also a problem in context of AVC’s definition of cropping was raised, but accrding to other experts’ opinion that should not be a problem.

Offline consideration (M. Hannuksela, Y.-K. Wang, Y. Sanchez) to discuss and provide text to resolve the issues.

All issues were resolved during drafting and reviewing of DoCR on CD ballot. All items were agreed as proposed.

## AHG9: General direction of grouping of post-processing filters (PPFs), including NNPFs (3)

Contributions in this area were discussed at 1410–1925 on Saturday 14 Oct. 2023 (chaired by JRO).

[JVET-AF0050](https://jvet-experts.org/doc_end_user/current_document.php?id=13298) AHG9: On NNPF groups [M. M. Hannuksela, F. Cricri (Nokia)]

Item 3 of this contribution belongs to this category.

Item 3: Except for the changes proposed in items 1 and 2 in this contributiion, other aspects of the NNPF group design in the VSEI TuC are considered mature in this contribution and are proposed to be included in the VSEI output document of the JVET-AF meeting, which may either be a working draft or a new version of the VSEI TuC document.

During the presentation of the document, it was suggested to get a conclusion on whether a specific SEI message is needed for the case of combining different NNPF devices, or if the processing order SEI can also be used here. It is also commented that evidence about real need for specific cases would be needed.

The contribution gives arguments why a separate SEI is needed for NNPF

1. Different types of grouping (alternatives, alternating, parallel) – it was commented that some of this may be resolved by other proposals
2. Optional processing steps – it is agreed that this is identical with the “essential” definition of SEIPO, if it is not essential, it can be condidered optional.
3. Signalling complexity of NNPF group – it was commented that a decoder could also retrieve the complexity information from the different NNPF SEI messages, and decide which operations are possible with its capability (e.g. skipping optional ones, or choosing the simpler one of alternating ones)

It was discussed which complexity would be signalled? Would it be the maximum or the minimum that would be possible? What would a decoder conclude, if has several options that would be possible by its capabilities

It was suggested that another option could be to signal complexity of the different contained NNPF messages in the prefix.

[JVET-AF0061](https://jvet-experts.org/doc_end_user/current_document.php?id=13309) AHG9: On grouping of post-processing filters [Y.-K. Wang, W. Jia, J. Xu, L. Zhang (Bytedance)]

At the 30th JVET meeting in April 2023, the BoG on NNPF and SEI processing order SEI messages agreed and the JVET endorsed that it is desirable to support all the following NNPF related capabilities in v4 of VSEI and VVC:

1. Multiple activated NNPFs for a picture in a cascading manner.
2. Multiple activated NNPFs (with the same or different purposes) for a picture in an alternative manner (e.g. choose to apply none, filter a only, or filter b only).
3. Multiple activated NNPFs (with the same or different purposes) for a picture in a cascading plus alternative manner (e.g. choose to apply none, filter a only, filter b only, or both filter a and b in a cascading manner in the indicated order).
4. Multiple activated NNPFs for a picture while the receiver may process the bitstream multiple times to generate multiple different results (e.g. for each time, to choose to apply none, filter a only, or filter b only).
5. Multiple activated NNPF cascades in an alternative manner (e.g. choose to apply none, both filter a and b in a cascading manner in the indicated order, or both filter c and d in a cascading manner in the indicated order).

JVET-AE2032 (Technologies under consideration for future extensions of VSEI (draft 1)) includes a NNPF grouping design resulted from proposals motivated by the above agreement and endorsement. The design involves addition of two new SEI messages, the neural-network post-filter group characteristics (NNPFGC) SEI message and the neural-network post-filter group activation (NNPFGA) SEI message, among other changes.

This contribution asserts that the NNPF grouping design is complicated and hard to understand. The design is also limited to only support NN-based post-processing filters (PPFs).

To address the agreed/endorsed NNPF related capabilities in v4 of VSEI and VVC, and extend the scope to both NN-based and non-NN-based PPFs, this contribution proposes the following changes:

1. Add a ue(v)-coded syntax element po\_id to the SEI processing order (SPO) SEI message, replace reserved\_alignment\_6bits in the SPO SEI message with byte alignment checking followed by po\_alignment\_zero\_bit, and specify the following PPF grouping mechanism (either as part of the semantics of the SPO SEI message or as part of the general post-processing filtering process):
   1. For each picture, there can be multiple PPFs activated and belonging to one or more PPF groups. PPF groups are alternative to each other, i.e., at most one group can be chosen to be applied.
   2. For backward compatibitiy with v3 of VVC/VSEI, a special PPF cascading case is defined, as the case when such two PPFs are both activated for a picture: the two PPFs are both NNPFs, one of the two NNPFs has nnpfc\_purpose equal to 4 and the other has multiple input pictures, and neither of the two NNPFs is associated with an SPO SEI message. In this case, the two NNPFs are implicitly considered as belonging to one PPF group, and the NNPF with nnpfc\_purpose equal to 4 is applied first.
   3. Except for the special PPF cascading case, each PPF group containing multiple PPFs is associated with an SPO SEI message with a particular value of po\_id.
   4. Except for the special PPF cascading case, any PPF not associated with an SPO SEI message is in its own PPF group.
   5. One or more PPFs in the chosen PPF group can be applied. When multiple PPFs (in the chosen PPF group) are applied, they are applied in the cascading manner where they are applied in the order indicated by the SPO SEI message associated with the chosen PPF group, and for each applied PPF that is not the last applied PPF, the output is used as the input of the next applied PPF.
2. Update the filtering process to apply to both NNPFs and non-NN-based PPFs and to enable the use of multiple activated PPFs from one PPF group in the cascading manner. Accordingly, both the NNPFC semantics in VSEI and the interface text in VVC are updated, such that when the PPF is not the first PPF that is applied for the current picture, the input to the PPF is the output of the previously applied PPF.

About 5) in the first item list, it was asked if a v3 legacy decoder might operate in non-desirable manner. It was commented that wrapping concepts could resolve that, i.e. make the NNPF SEI within a group invisible if its standalone usage is undesirable.

From the concepts of JVET-AF0050, “alternating” and “parallel” are missing according to opinion of its proponents. It is explained that one application of “parallel” would be to generate two outputs, e.g. one for human vision, one for machine vision. The view of “alternating” would be more from encoder perspective, compared to “alternative” where a decoder chooses which one to use.

It was commented that similar could be achieved by an encoder activating both of the “parallel” or “alternating” NNPFs, and let the decoder decide which to use.

It is understood that the concept of JVET-AF0061 is, that if it is possible to activate (and choose at the decoder to run) more than one group, such intended behaviour could be accommodated.

It was commented that a case where number of inputs and outputs in the cascade is not matching (e.g., a temporal upsampler expecting two separately post-filtered frames) might not be supported.

Follow-up presentation on Sun. 15 Oct. on remaining aspects/proposed changes 1)/2) in the second item list:

1) refers to text changes in the SPO SEI.

2) refers to text changes in VVC (particularly specifying input/output interfaces of sequential postprocessing elements, including aspects of cropping and size changes), and changes in VSEI, particularly NNPFC

The aspects of po id are to be clarified in the text development for the agreed overall concept, for changes in the SPO SEI text. It was agreed that fixed length would be preferable

The other suggested changes are largely agreed to be beneficial, but many changes are proposed, and more in-depth investigation is necessary to sort out which changes are necessary in the SPO SEI itself, in VVC/VSEI body text, and in VSEI NNPF. Therefore, it was agreed to move the include the suggested text changes (except for po\_id) in the TuC document for further study, and add notes in the SPO SEI text where specifications might be missing (e.g. for input/output interfaces of sequential filters), and that a possible text can be found in the TuC.

It was suggested to move the paragraphs below semantics of po\_id to the section on more general post processing filtering (as per yellow highlight above).

[JVET-AF0174](https://jvet-experts.org/doc_end_user/current_document.php?id=13432) AHG9: On the SEI processing order SEI message and NNPF groups [Y. Sanchez, C. Hellge, T. Schierl (Fraunhofer HHI)]

In this document a proposal is described that suggests modifications to the SEI processing order SEI message to allow referencing payload type and Id for association of an SEI message, as well as to specify a length of the prefix of an SEI indicting the bits that are used to identify the SEI message which a particular order id refers to. The proposal also suggests to allow that an SEI processing order SEI message can reference another SEI processing order SEI message and thereby the processing order specified in the latter is inherited by the first SEI processing order SEI message. Besides, a group activation SEI message is proposed that activates all NNPF filters that are associated with an SEI processing order SEI message.

Similar concept as JVET-AF0061, multiple SPO messages with different content would be allowed, but they need to have different id. Concepts of identifying and activating different SPO more explicitly introduced here.

The additionally proposed NNPF group activation SEI message would activate all NNPF SEI contained in an SPO together. The SPO could still contain other (non-NNPF) SEIs. It commented that one functionality of TuC (activating an update of a specific NNPF in a group individually) is discarded by this concept. It was further commented that when an NNPFA would be sent separately, it might be misinterpreted by a legacy decoder.

Presentation of remaining items from JVET-AF0174 on Sunday Oct. 15.

A syntax element po\_sei\_payload\_id is proposed. In the associated semantics, the content of SEI messages grouped elsewhere in the nested structure can be inherited, such that duplication of information is avoided.

Alternatively, association is possible via the prefix mechanism. It is explained that this is useful in cases where some part of the prefix is changing, e.g. for sending complexity information of NNPF.

It was asked if the number of nested stages would somehow be limited? This is currently not foreseen in the proposal, but would appear reasonable.

It was commented that this appears generally more complicated than implementing nesting by a separate SEI message as proposed in JVET-AF0049. If the main purpose of JVET-AF0174 is saving bits, there would be higher priority of keeping the approach simple. For example, re-sending prefix would not happen often.

It is also proposed to use NNPF group activation with SPO. According to previous discussion, it had been agreed to handle NNPF in SPO not different from other SEI messages. The proponent was asked to study whether this mechanism could be useful in the context of the TuC.

Some more aspects related to this category are contained in JVET-AF0049 item 1, JVET-AF0065, JVET-AF0067, and JVET-AF0189 (mainly aspect of persistence). After presentation of these (see below), following aspects were summarized to be clarified as intermediate conclusion:

* Which amount in flexibility is needed in processing order in terms of sequential, parallel, alternating/alternative?
* Need for separate messages for processing order (only supporting non-NNPF) and NNPF grouping?
* Allowing for several SPO with different processing order?
* Concepts of nesting (with or without a separate SEI)
* Aspects of persistence

Why should NNPF be treated differently than other SEI messages?

Can we expect in the near future (e.g., targeting v4 of VVC/VSEI) a sophisticated combination of different NNPF SEI messages, rather than combining NNPF with “conventional” SEI?

NNPF would work in the current definition of SPO, if the scope/persistence issue is resolved.

It is agreed to proceed with an approach based on modifying SPO such that

* Scope/persistence is resolved (possible solutions in JVET-AF0049/JVET-AF0174/JVET-AF0189)
* Add po\_id to support for alternative cascaded SEI messages (possible solutions in JVET-AF0061/JVET-AF0067/JVET-AF0174)

This is the general design concept to be implemented towards v4 of VVC/VSEI. Some other aspects from the discussion are secondary.

For further study of potential future aspects, the NNPF group design shall be kept in TuC.

Proponents of the contributions listed above were asked to provide a text proposal (see JVET-AF0310).

[JVET-AF0310](https://jvet-experts.org/doc_end_user/current_document.php?id=13574) AHG9: Combined text for scope/persistence changes and addition of po\_id to the SEI processing order SEI message and addition of the processing order nesting SEI message [M. M. Hannuksela (Nokia), G. J. Sullivan (Dolby), Y. Sanchez (Fraunhofer HHI), Y.-K. Wang (Bytedance), L. Chen (MediaTek)] [late]

In response to the review and agreements in the JVET track meeting on 14 Oct 2023, this contribution proposes a combination of aspects from contributions JVET-AF0049, JVET-AF0061, JVET-AF0174, and JVET-AF0189 related to the scope/persistence of the SEI processing order SEI message, and the addition of po\_id from contributions in JVET-AF0061, JVET-AF0067, and JVET-AF0174. The aspects proposed in this contribution can be summarized as follows:

1. Change the persistence scope of the SEI processing order SEI message as follows: Each SEI message identified within the SEI processing order SEI message has the same persistence scope as if the SEI message was carried outside of the SEI processing order SEI message and not identified within an SEI processing order SEI message. The number of SEI messages and the type codes of the SEI messages indicated within the SEI processing order SEI message persist in decoding order from the first access unit in decoding order until the end of the CVS in output order.
2. Remove of the capability of wrapping SEI messages within the SEI processing order SEI message and add of a processing order nesting SEI message that includes one or more SEI messages having the same persistence and are used in an indicated position within the processing order.
3. Add processing order ID, po\_id, as a u(8) syntax element.
4. Replace the requirement

"When there are multiple SEI processing order SEI messages present in a CVS, they shall have the same content"

with

"When more than one SEI processing order SEI message with a particular value of po\_id is present in a CVS, the values of po\_num\_sei\_messages\_minus2 and, for each value of i, the values of po\_sei\_wrapping\_flag[ i ], po\_sei\_prefix\_flag[ i ], po\_sei\_importance\_flag[ i ], po\_sei\_payload\_type[ i ], po\_sei\_processing\_order[ i ] shall be the same as in the other SEI processing order SEI messages in the CVS with the same value of po\_id."

1. To cope with multiple PO SEI messages with different po\_id values, relax the requirement of po\_sei\_processing\_order[ i ] from being equal or incremented by 1 compared to the po\_sei\_processing\_order[ i − 1 ] to being greater than or equal to po\_sei\_processing\_order[ i − 1 ].
2. To cope with multiple PO SEI messages with different po\_id values, include one or more pon\_target\_po\_id[ i ] values in the processing order nesting SEI message to identify associated PO SEI messages.

Was presented Thu 19 Oct. 0940.

Decision: Adopt JVET-AF0310 into JVET-AF2027 (draft 6, prelim. WD in WG5)

The earliest possibility of putting this into a new version of H.266 would be by the April 2024 meeting. Usually, new editions of VVC and VSEI should be synchronized, and it is to be determined if any technology from the TuC would be mature enough by that time. This should be further discussed in the January meeting, where potentially a preliminary WD of a VSEI extension could be established. In terms of timing in ISO, there is no benefit of starting a more official standardization effort on SPO by this meeting.

It was commented that the SPO SEI would also be useful in the AVC and HEVC context. Provided that an action on this would be made for VVC in April, it might be reasonable to include a similar approach in the next editions of AVC and HEVC as well. On the side of ISO, this would require NB comments in the DIS and DAM ballots.

## AHG9: Detailed aspects of NNPF grouping (9)

Contributions in this area were discussed at 1345–1700 on Sunday 15 Oct. 2023 (chaired by JRO).

[JVET-AF0050](https://jvet-experts.org/doc_end_user/current_document.php?id=13298) AHG9: On NNPF groups [M. M. Hannuksela, F. Cricri (Nokia)]

Items 1 and 2 of this contribution belong to this category.

This input document presents a study of the neural-network post-filter (NNPF) group design in JVET-AE2032 "Technologies under consideration for future extensions of VSEI (draft 1)" (VSEI TuC). The following changes are proposed relative to JVET-AE2032:

1. In the JVET-AE meeting, nnpfa\_no\_prev\_clvs\_flag and nnpfa\_no\_foll\_clvs\_flag were adopted to the neural-network post-filter activation (NNPFA) SEI message to control multi-input NNPFs in bitstream splicing. This contribution proposes to introduce similar syntax elements also to the NNPF group activation (NNPFGA) SEI message.
2. It is asserted that bitstreams where nnpfa\_target\_id is equal to an nnpfgc\_id value would be regarded as non-conforming bitstreams according to VSEI v3. Consequently, and also because the NNPFGA SEI message is asserted to provide more flexibility than the NNPFA SEI message, it is proposed to remove the possibility to reference NNPF group identifiers from the NNPFA SEI message.

Except for the changes proposed above, other aspects of the NNPF group design in the VSEI TuC are considered mature in this contribution and are proposed to be included in the VSEI output document of the JVET-AF meeting, which may either be a working draft or a new version of the VSEI TuC document.

Item 1: Straightforward alignment.

Item 2: Activating a group by NNPFA would cause problems with legacy decoders that don’t know groups. This problem is fixed by removing that option.

Agreed for inclusion in TuC

[JVET-AF0051](https://jvet-experts.org/doc_end_user/current_document.php?id=13299) AHG9: Signalling the gain provided by NNPFs and NNPF groups [M. M. Hannuksela, F. Cricri, A. Hallapuro (Nokia)]

Signalling the quality improvement or the gain resulting from applying a neural-network post-filter (NNPF) enables the decoding system to select whether to apply the NNPF given its signalled complexity. Moreover, if there are multiple alternative NNPFs, by signaling the gain and complexity of the NNPFs, the decoding system can choose the NNPF that provides a good trade-off between the gain and complexity given the resources of the decoding system. This document proposes the following items to indicate the gain of NNPFs and NNPF groups:

1. An NNPF gain syntax structure that implements the following design decisions:
2. As a starting point, it is proposed to specify only the following PSNR-based metric in the semantics:
   1. Specify the formula for PSNR identically to the common test conditions.
   2. The improvement of between the average PSNRs before and after post-filtering is indicated as the gain value.
3. Since new metrics may evolve over time and it may be tedious for JVET and its parent organizations to standardize metrics, it is proposed to allow URI-based identification of metrics.
4. Signalling multiple metrics is enabled, since a decoder may understand one metric but not the other, or the same NNPF may provide gains for different purposes that could be characterized by different metrics.
5. It is proposed to indicate whether the indicated gain concerns colour components collectively, and, if not, the components (Y, Cb, or Cr) that the indicated gain concerns.
6. The representation format of the signalled post-filter gain, such as a fixed-point value or a floating-point value of a certain precision, is proposed to depend on the metric itself. The size of the signalled post-filter gain is proposed to be indicated in 16-bit words so that metrics that are not recognized by a decoder can be skipped.
7. Either of the following options for carrying the NNPF gain syntax structure for individual NNPFs:
8. An NNPF gain SEI message.
9. Metadata added to NNPFC message using nnpfc\_reserved\_metadata\_extension bits.
10. Either of the following options for carrying the NNPF gain syntax structure for NNPF groups:
11. Enabling the use of the NNPF gain SEI message for NNPF groups.
12. Metadata added to NNPF group characteristics (NNPFGC) SEI message.

As separate SEI, it could also be used for other post-processing approaches.

It was commented that expression of PSNR should be made generally for all bit depths. It was also commented that a precision of 16 bit may not be needed for the purpose.

Can there be negative gain? Not foreseen in the current version. It was commented that there could be subjective benefit but PSNR loss.

It was commented that the URI referencing might not be useful for interoperability. There were however also different opinions

It was suggested to also allow absolute scale expression of quality (potentially before and after).

It was asked for which pictures the quality improvement is valid? Average for those pictures for which the filter is activated in the CVS. Other granularities might be useful

It was commented that green metadata has a similar SEI that also contains more metrics than PSNR, and has subpicture granularity.

It is generally agreed that this is useful.

It is agreed to include the approach of separate SEI (applicable for any PF) to TuC. Remove the URI, define PSNR in a bit-depth neutral. Other aspacts for further study.

[JVET-AF0067](https://jvet-experts.org/doc_end_user/current_document.php?id=13315) AHG9: On the SEI processing order SEI message (Part 2) [L. Chen, O. Chubach, Y.-W. Huang, S. Lei (MediaTek)]

This contribution proposes to restructure the SEI processing order SEI message in JVET-AE2027 into SEI processing order group characteristic SEI message and SEI processing order group activation SEI message with the following brief description.

1. The SEI processing order group characteristic SEI message carries information indicating the preferred initial processing order type for different types of SEI messages that may be present in a CVS. The use of SEI processing order groups combined with any other types of SEI messages that may be present in a CVS for specific pictures is indicated with SEI processing order group activation SEI messages indicating the preferred final processing order types.
2. The SEI processing order group activation SEI message activates or de-activates the use of the combination of target SEI messages and/or target SEI processing order groups with the preferred final processing order type for processing of a set of pictures.
3. When no SEI processing order groups specified by SEI processing order group characteristic SEI messages rather than SEI messages indicated in the activation message are present for activation, the SEI processing order group activation SEI message may be used without a preceding SPOGC SEI message.

From first presentation Sat. Oct. 14:

Contains definition of purpose (unspec/human/machine), and type (cascade/alternate/parallel). However, in the current concept, parallel could not be used as one branch to human, one for machine.

Wrapping is supported, as well as importance signalling.

There could several of these SEI messages be present and activated.

Follow-up presentation Sunday Oct. 16:

* No aspects that would be relevant for NNPFGC in TuC
* Aspect of purpose: See notes under JVET-AF0065
* Aspect of group id covered in collaborative effort with other contributions.

[JVET-AF0091](https://jvet-experts.org/doc_end_user/current_document.php?id=13338) [AHG9] On design for signalling purpose information for NNPFGC [[Hendry](mailto:dr.hendry@lge.com), J. Nam, S. Kim, J. Lim (LGE)]

It is asserted that the design for signalling purpose information in neural-network post-filter group characteristic (NNPFGC) SEI message may benefit from some modifications for improvement. The proposed modifications are as follows:

1. Add the following definition:
   * Direct member of an NNPFGC: An NNPFC with nnpfc\_id equal to nnpfgc\_member\_id[ i ] for any i in the range of 0 to nnpfgc\_num\_members\_minus2 + 1, inclusive.
   * Indirect member of an NNPFGC: An NNPFC that is a direct or indirect member of another NNPFGC nnpfgcB, in which the value of nnpfgc\_id of nnpfgcB is equal to nnpfgc\_member\_id[ i ] for any i in the range of 0 to nnpfgc\_num\_members\_minus2 + 1, inclusive.
2. Clarify that the value of nnpfgc\_purpose in NNPFGC, when present, includes all purposes of NNPFC or NNPFGC included in the NNPFGC.
3. When nnpfgc\_grouping\_type is equal to 1 (i.e., alternative grouping) or 3 (i.e., parallel grouping), purpose information is signalled for each member of the NNPFGC. This purpose information may be signalled as nnpfgc\_member\_purpose[ i ] and the following applies:
   * If nnpfgc\_member\_id[ i ] is equal to nnpfc\_id of a particular NNPFC, the value of nnpfgc\_member\_purpose[ i ] shall be equal to the value of nnpfc\_purpose of the NNPFC with nnpfc\_id is equal to nnpfgc\_member\_id[ i ].
   * Otherwise (i.e., nnpfgc\_member\_id[ i ] is equal to nnpfgc\_id of a particular NNPFGC), the value of nnpfgc\_member\_purpose[ i ] shall be the union of the values of nnpfgc\_member\_purpose[ i ] in the NNPFGC.

Item 1 is agreed as beneficial. An “indirect member” is an NNPF that is contained in another group that is member of the current group. Conceptually, this is already allowed, therefore adding this definition is somewhat editorial.

It was commented that the different NNPF messages included in the group can have different purposes, and it is relevant to access them (item 3). It was commented that again signalling the purpose of the entire group, as per item 2, somewhat duplicates information. It was claimed that it is more convenient for accessing the purpose overall.

Items 1, 2 and 3 agreed for inclusion in TuC for NNPFGC.

[JVET-AF0092](https://jvet-experts.org/doc_end_user/current_document.php?id=13339) [AHG9] On the output pictures from NNPFGC with parallel grouping type [[Hendry](mailto:dr.hendry@lge.com), J. Nam, S. Kim, J. Lim (LGE)]

It is asserted that the current text for neural-network post-filter group characteristic (NNPFGC) SEI message when the grouping type is parallel grouping is not clear on the following two cases:

* When two or more NNPFs in a parallel group produce interpolated pictures in between two same consecutive input pictures. Without clear direction what to do, different decoder may order these interpolated pictures in different ways in between those two input pictures.
* When two or more NNPFs in a parallel group produce output pictures associated with the same input picture. Without clear direction what to do, different decoder may override the output of other parallel filter(s).

To address the asserted problems above, the following modifications are proposed:

1. When two interpolated pictures, intPicA and intPicB, are outputted by members of an NNPFGC such that intPicA is an interpolated picture outputted by nnpfgc\_member\_id[ m ] and intPicB is an interpolated picture outputted by nnpfgc\_member\_id[ n ] and all the following applies:
   * Both intPicA and intPicB are in between the same two consecutive input pictures
   * The value of nnpfgc\_grouping\_type is equal to 3
   * m is less than n

Then the order of intPicA and intPicB shall be such that intPicA precedes intPicB in output order.

1. When the value of nnpfgc\_grouping\_type of an NNPFGC is equal to 3, it is constrained that there shall be no two or more filters have output pictures that are associated with the same input picture(s).

With the current parallel concept (also specified in NNPFGC), the outputs are independent. Therefore, theproblem does not exist.

[JVET-AF0093](https://jvet-experts.org/doc_end_user/current_document.php?id=13340) [AHG9] On intermediary output picture(s) from activation of an NNPFGC [[Hendry](mailto:dr.hendry@lge.com), J. Nam, S. Kim, J. Lim (LGE)]

It is asserted that the current text for neural-network post-filter group activation (NNPFGA) SEI message lacks of support for the case when there is a need to have intermediate output pictures (i.e., pictures that are output by previous filter(s) in the same NNPFGC but are not needed for final output). This contribution proposes modification with aims at enabling the NNPFGA SEI message to support such case.

The proposed modifications are as follows:

1. At the end of the current syntax table of NNFPGA, signal a new syntax element called nnpfga\_num\_output\_pic\_update that specifies the number of output picture output flag in the output picture list to be updated. The value of nnpfga\_num\_output\_pic\_update shall be less than or equal to the number of pictures in NnpfgaOutputPicList.
2. If nnpfga\_num\_output\_pic\_update is greater than 0, signal a loop of nnpfga\_output\_pic\_update\_flag[ i ] for i ranges from 1 to nnpfga\_num\_output\_pic\_update, inclusive.
   * nnpfga\_output\_pic\_update\_flag[ i ] is 1 specifies that the i-th picture in NnpfgaOutputPicList is included in the final output picture.
   * nnpfga\_output\_pic\_update\_flag[ i ] is 0 specifies that the i-th picture in NnpfgaOutputPicList is not included in the final output picture
3. When the value of nnpfga\_num\_output\_pic\_update is less than the number of pictures in NnpfgaOutputPicList, the value of nnpfga\_output\_pic\_update\_flag[ i ] for i in the range from nnpfga\_num\_output\_pic\_update to the number of pictures in NnpfgaOutputPicList − 1, inclusive, is inferred to be equal to 1.

This can be seen as a generalization of hierarchical structures. This could be used to generate non-dyadic temporal upsampling ratio by using only dyadic filters.

Agreed for inclusion in TuC for NNPFGC

[JVET-AF0094](https://jvet-experts.org/doc_end_user/current_document.php?id=13341) [AHG9] On activation of an NNPFGC that contains another NNPFGC [[Hendry](mailto:dr.hendry@lge.com), J. Nam, S. Kim, J. Lim (LGE)]

It is asserted that the current text for neural-network post-filter group activation (NNPFGA) SEI message has problem such that NNPFGA SEI message cannot activate an NNPFGC that has another NNPFGC as its member. This contribution proposes the following modification to overcome the asserted problem. The proposed modifications are as follows:

1. Remove the constraint that the value of nnpfga\_num\_filters\_minus2 shall be equal to the value of nnpfgc\_num\_members\_minus2 in an NNPFGC SEI message with nnpfgc\_id equal to nnpfga\_target\_id.
2. Specify that the value of nnpfga\_num\_filters\_minus2 shall be equal to or greater than the value of nnpfgc\_num\_members\_minus2 in an NNPFGC SEI message with nnpfgc\_id equal to nnpfga\_target\_id.
3. Specify that nnpfga\_num\_filters\_minus2 plus 2 specifies the number of direct or indirect NNPF members of an NNPFGC SEI message with nnpfgc\_id equal to nnpfga\_target\_id.
4. The list of NNPFs activated by an NNPFGA is the list consisting the direct and indirect NNPF members of the target NNPFGC in the order based on depth-first-search traversal of the NNPFGC’s members.

It was commented that the constraint was introduced for the purpose of hierarchical grouping, and associated activation level by level. No need to change, but some text clarification of the intent could be useful.

[JVET-AF0095](https://jvet-experts.org/doc_end_user/current_document.php?id=13342) [AHG9] On order of output pictures when skipped candidate input pictures are present in NNPFGA [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

It is asserted that the current text for neural-network post-filter group activation (NNPFGA) SEI message has problem when skipped input pictures are used and one or more filters in the target NNPFGC are filters that produce interpolated pictures. When one or more pictures are skipped from being used as input pictures, the skipped pictures may cause problem for decoder to determine where to put the interpolated output pictures from the filter that are supposed to be in between the two input pictures.

This contribution proposes the following modification to overcome the asserted problem. The proposed modifications are as follows:

1. For NNPFGA design, signal a flag for the i-th NNPF to specify whether or not the NNPF has interpolated picture(s) output. This flag is constrained to be equal to 1 when the NNPF includes picture rate upsampling purpose and be equal to 0 when the NNPF does not include picture rate upsampling purpose.
2. When the value of nnpfga\_input\_pic\_skip\_count[ i ][ j ] is greater than 0, signal a new syntax element called nnpfga\_num\_interpolated\_pic[ k ][ j ][ k ] for k ranges from 0 to nnpfga\_input\_pic\_skip\_count[ i ][ j ], inclusive, to specify the number of interpolated pictures generated by the i-th NNPF in the target NNPFG between the k-th and the ( k + 1 )-th pictures that are skipped in the list of candidate input pictures candInputPicList[ i ] after the j-th input picture of the NNPF
3. It is constrained that when nnpfga\_num\_interpolated\_pic[ k ][ j ][ k ] is present, the sum of nnpfga\_num\_interpolated\_pic[ i ][ j ][ k ] for k ranges from 0 to nnpfga\_input\_pic\_skip\_count[ i ][ j ], inclusive, shall be equal to the value of nnpfc\_interpolated\_pics[ j ] of the NNPF.

It was commented that the example is an unusual cornercase with two different filters generating new pictures at different relative positions . Also, timing of output should be managed by systems level, where the system should know at which temporal positions pictures are generated. No need for action.

[JVET-AF0096](https://jvet-experts.org/doc_end_user/current_document.php?id=13343) [AHG9] On miscellaneous aspects of NNPFGC and NNPFGA [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

This contribution proposes several modifications that are asserted to provide improvement to some aspects of the NNPFGC and NNPFGA SEI messages design.

Item 1 relates to complexity: Complexity for alternative to be signalled should be maximum complexity.

It was commented that the complexity information can also be determined from the individual complexities of the filters in the group, and a decoder that is not capable of processing the maximum complexity might still be capable to process an alternative configuration. From that point of view, signalling minimum complexity (or both min and max of the group) might be better. Further study appears necessary which complexity is useful to be signalled for a group for best benefit of a decoder. Such a study on the needs of signalling group complexity would also be useful for SPO, where it is not signalled. There, the consideration of minimum complexity would apply to cascaded configuration where it would only include the essential filters.

Item 2 relates to input pictures information signaling for first NNPF in the group. It was commented that this signalling was introduced by purpose in cases where different input pictures are candidates that could be selected. No action necessary on this.

## AHG9: SEI processing order SEI message aspects other than PPF grouping (5)

Contributions in this area were discussed at 0835–1235 on Sunday 15 Oct. 2023 (chaired by JRO).

Remaining aspects of JVET-AF0174 were also presented on Sunday Oct. 15 (see in sec. xx).

[JVET-AF0049](https://jvet-experts.org/doc_end_user/current_document.php?id=13297) AHG9: On the SEI processing order SEI message [M. M. Hannuksela (Nokia)]

This contribution proposes the following changes related to the SEI processing order SEI message:

1. Removal of the capability of wrapping SEI messages within the SEI processing order SEI message and addition of a processing order nesting SEI message that includes one or more SEI messages having the same persistence and are used in an indicated position within the processing order.
2. Removal of the following constraints, which are asserted to be imposed elsewhere: "It is a requirement of bitstream conformance that, within an SEI processing order SEI message, there shall be at least two pairs of the syntax elements po\_sei\_payload\_type[ i ] and po\_sei\_processing\_order[ i ], and there shall be at least two values of po\_sei\_processing\_order[ i ] that are not equal."
3. Indication of the increment po\_sei\_processing\_order\_incr\_flag[ i ] relative to the previous derived processing order value, instead of an absolute processing order value po\_sei\_processing\_order[ i ].
4. Indication of the number of bits instead of bytes for SEI message prefixes included in the SEI processing order SEI message.

Item 1 (discussed Sat. 14 Oct.) points out a problem that persistence of various SEI messages contained in an SPO might be different.

Could this also be resolved by an encoder constraint?

More flexibility is introduced by the processing order nesting SEI (not replacing SPO) the syntax element pon\_processing\_order[i] could allow gaps, and be not strictly sequential. Some similarity with JVET-AF0174, but the latter does not require a separate SEI.

SEI messages can be contained as payload (but don’t need to), to hide them from legacy decoders (similar to wrapping).

Items 2-4 presented Sun. 15 Oct.

Item 2 (removal of constraints). Agreed. Also referred in JVET-AF0062, and slightly different in JVET-AF0189. Purely editorial.

Item 3 (delta signalling). JVET-AF0062 items 3&4 are similar. It was commented that there might be a problem if there are gaps (as in the concept of the proposed SPO nesting SEI), but also in certain cases with no gaps. Also there may be a problem when two SPO messages refer the same SEI but at different positions. Even if the nesting concept is not used, the conflict with it indicates that the approach of item 3 might restrict possible further extensions for the benefit of saving some bits in the SPO SEI. No action.

Item 4 (ed. note on SEI prefix). It was commented and agreed that it is better resolving this issue consistenly with other prefix SEI cases (as proposed in JVET-AF0189 and JVET-AF0062 item 2).

[JVET-AF0062](https://jvet-experts.org/doc_end_user/current_document.php?id=13310) AHG9: Some syntax and semantics changes for the SEI processing order SEI message [Y.-K. Wang (Bytedance)]

This contribution proposes the following syntax and semantics changes for the SEI processing order SEI message:

1. Move po\_sei\_prefix\_flag[ i ] from immediately before po\_sei\_payload\_type[ i ] to be immediately after po\_sei\_payload\_type[ i ], because the SEI payload type is a higher-level information and this move can also avoid potential semantics dependency on a syntax element later in decoding order. – agreed.
2. To avoid potential syntax parsing problems, the SEI prefix indications, when present, are signalled in units of bits, and an SEI prefix indication for a particular SEI payloadType shall be a bit string that follows the SEI payload syntax of that value of payloadType and contains a number of complete syntax elements starting from the first syntax element in the SEI payload, same as in the SEI prefix indication SEI message.
3. The signalling of the preferred order of processing for the 0-th SEI message type is skipped, and the value is derived to be equal to 0. – no action (see notes under JVET-AF0049)
4. The preferred order of processing for the i-th SEI message type for i greater than 0 is signalled using a flag instead of 8 bits, the flag indicating whether the preferred order of processing for the i-th SEI message type is equal to that of the ( i – 1 )-th SEI message type plus 1 or is the same as that of the ( i – 1 )-th SEI message type. – no action (see notes under JVET-AF0049)
5. Add a NOTE to clarify the following aspects: In the semantics of the SPO SEI message, two different types of SEI messages may have the same SEI payloadType value but are differentiated by some syntax elements in the SEI payload. For example, two NNPFC SEI messages with different nnpfc\_id values are considered as having two different SEI message types. This change is asserted to be purely editorial. agreed
6. Remove the first constraint below as it is redundant to the second constraint below (it would also be OK to keep the first constraint and remove the second one, but the first one is more complicated and has some buggy issues that are asserted to need some effort to fix). This change is asserted to be purely editorial. – agreed by notes elsewhere

*The first constraint:* It is a requirement of bitstream conformance that, within an SEI processing order SEI message, there shall be at least two pairs of the syntax elements po\_sei\_payload\_type[ i ] and po\_sei\_processing\_order[ i ], and there shall be at least two values of po\_sei\_processing\_order[ I ] that are not equal.

*The second constraint:* The value of po\_sei\_processing\_order[ po\_num\_sei\_messages\_minus2 + 1 ] shall not be equal to 0.

1. It is constrained that, an SEI prefix indication for a colour transform information (CTI) SEI message shall include at least all bits of the colour\_transform\_id syntax element of the CTI SEI message.
2. It is constrained that, an SEI prefix indication for a neural-network post-filter characteristics (NNPFC) SEI message shall include at least all bits of the nnpfc\_id syntax element of the NNPFC SEI message.

The proposed text changes, relative to JVET-AE2027, are included in an attachment of this contribution. There are also some changes included in the attachment that are not covered in the summary above, and those changes asserted to be purely editorial and very minor. An example of such changes is replacing of “if” without a corresponding “otherwise” with “when”.

For item 2, conceptually similar to an item in JVET-AF0189. It was pointed out that there may be a problem with byte alignment in the syntax of JVET-AF0062.

For items 7 and 8, further study is recommended about ways to impose constraints for specific SEI messages. E.g. the listing which is planned to be developed as per JVET-AF0189 item 5 could have footnotes for specific SEI messages which constraints would need to be imposed on them.

It was also commented on item 8 that there may be cases where the id is not needed for NNPF, but the purpose would be needed.

[JVET-AF0065](https://jvet-experts.org/doc_end_user/current_document.php?id=13313) AHG9: On the SEI processing order SEI message (Part 1) [L. Chen, O. Chubach, Y.-W. Huang, S. Lei (MediaTek)]

This contribution addresses comments related to the JVET-AE2027 made at the recent JVET-AE meeting. It is proposed to add following aspects in the JVET-AE2027 - SEI processing order SEI message in VVC (draft 5):

1. Scope of persistence
2. Processing order type, such as Cascade, Parallel, and Alternate etc.
3. Processing purpose, such as general visual quality improvement, processing for human viewing, and processing for machine etc.
4. Support of the neural network post filter (NNPF) SEI messages with two options

Shortly presented Sat. 14 Oct. in the context of JVET-AF0067.

Introduces functionality from NNPF grouping SEI into SPO SEI.

There can not be more than one SPO SEI in this concept.

More detailed presentation on Sun. 15

1) it is proposed that persistence of all SEI messages shall be identical. However, as per decision from previous discussion, only the sequence of processing shall be persistent for the entire CVS, whereas SEI messages contained in the chain can change. No need for.

2) as per previous discussion, cascade and alternative will be supported. Other order types for further study.

3) it is kept for further study whether it is useful to signal a purpose of the chain, and if yes, which types of purpose would be useful.

4) No specific handling is necessary for NNPF messages. NNPFC and NNPFC can both be signalled to be part of the chain (e.g. with same processing order number). Regarding wrapping in the SPO, both could be sent within, or only one of them. The case that NNPFC is wrapped and inside and NNPFA outside is not possible per constraint in v3 Perhaps informative note may be useful.

[JVET-AF0070](https://jvet-experts.org/doc_end_user/current_document.php?id=13318) AHG9: On the SEI processing order SEI message [T. Chujoh, Y. Yasugi, T. Ikai (Sharp)]

The SEI processing order SEI message carries information indicating the preferred processing order, as determined by the encoder, for different SEI messages. As a result of examining the specification of the SEI processing order SEI message, three problems have been identified. This contribution discusses the solutions to these problems. The first problem is an editing problem for a variable. The second problem is the difference between the input picture size of two SEI messages and the other SEI messages. The third problem is that the input picture information of concatenated SEI messages needs to be re-specified. Several possible solutions to these problems have been presented.

The following problems exist in the texts of the SEI processing order SEI message.

* **Problem 1**: The definition of the variable SeiProcessingOrderSeiList is not clear. This variable may not be needed.
* **Problem 2**: The input picture size of SEI processing order messages is not constant. According to H.266 (V3), the input picture size of the film grain characteristics SEI message and the decoded picture hash SEI message is the non-cropped picture. The input picture size of the other SEI massages of image processing is cropped picture.
* **Problem 3**: There are differences between the input and output picture information of several SEI massages. For example, NNPFC and NNPFA SEI massages satisfy this condition. When those SE messages are concatenated, the input picture information of the order list of SEI processing order SEI message should be redefined.

Problem 1 already resolved (see notes elsewhere)

About problems 2 and 3: It is agreed that input/output relationships at every interface of the chain shall precisely be defined. JVET-AF0061 proposes a solution on that which does not need additional syntax. (see further notes there, these aspects were reviewed Sun 15 Oct. as well)

About problem 2: Disallowing decoded picture hash is useful and agreed

[JVET-AF0189](https://jvet-experts.org/doc_end_user/current_document.php?id=13449) AHG9: Proposed modifications of the draft SEI processing order SEI message in VVC [G. J. Sullivan, S. McCarthy, P. Yin (Dolby Labs)]

This document contains proposed some modifications of 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. The proposed modifications include the following:

* Using separate loops for the payload type and processing order information and the wrapped and prefixed other SEI messages;
* Modifying the persistence scope and constraints to allow changes of the content of the associated other SEI messages but not changes of their processing order or their way of being handled within the SEI processing order SEI message (this aspect was discussed in context of section 6.2);
* Making the syntax and semantics of the SEI processing order SEI message more consistent with the existing SEI prefix indication and scalable nesting SEI messages by using similar syntax, variable names and semantics;
* Removing the constraint that “The value of po\_sei\_processing\_order[ po\_num\_sei\_messages\_‌minus2 + 1 ] shall not be equal to 0”, since, for example, if an alternative transfer characteristics SEI message and a colour transform information SEI message are both indicated with processing order 0, the input of the colour transform should be interpreted as having the alternative transfer characteristics, and the output of the colour transform could be intended for some other display characteristics associated with the ID value sent in the colour transform information SEI message.
* Modifying the SeiProcessingOrderSeiList that determines which SEI messages are allowed to appear in an SEI processing order SEI message.

Revision marks are used to show the proposed changes relative to the Draft 5 text of the JVET output document JVET-AE2027.

It is also suggested to consider future clarification action of the existing text of VVC by appending “in decoding order” after “in the first access unit of the CVS” in the semantics of the SEI manifest and SEI prefix indication SEI messages.

Persistence can change for individual SEI messages, also prefix can change, but the overall order cannot change.

Several SPO SEI shall define same overall order (including the presence of wrapping and prefix). Enforcing wrapping in all of them may cause need for largely duplicating payload. It was pointed out that nesting (e.g. JVET-AF0049) may be better on that.

For usage with NNPF, further modification would be necessary due to separate activation.

Second bullet point (item 2) discussed Sat. Oct. 14, other bullet points Sun. Oct. 15

Item 1: It was commented that, if a separate nesting message would be defined, the loop for “wrapping” would not be necessary. Otherwise, agreed.

Items 3/4 conceptually agreed (see notes elsewhere). Task for editors.

For item5, the SEI messages that could be contained are restricted. It was commented that the aspect of prefixing is misplaced in the semantics proposed. It was further commented that it would be useful to have this list in a tabular form in VVC, associated with this SEI message but not part of it. At this moment, there is no need for the list to be complete, and it could be modified anyway when new SEI messages are specified in the future. Agreed - task for editors

The proposal also points out a possible need of clarification of semantics for the existing SEI manifest and SEI prefix messages. Agreed to include this aspect in JVET-AF1004.

## AHG9: SEI messages related to generative face video (3)

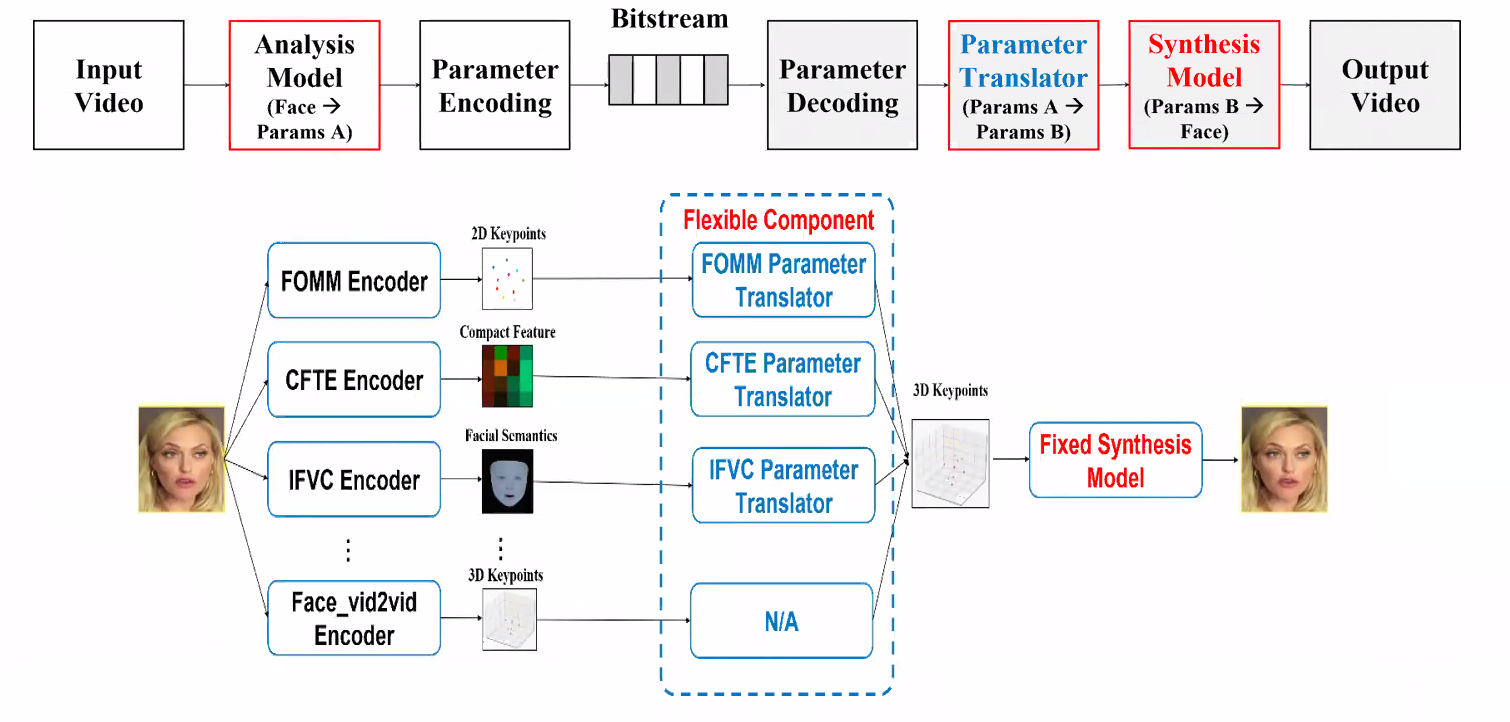
Contributions in this area were discussed at 1450–1610 on Wednesday 18 Oct. 2023 (chaired by JRO).

[JVET-AF0048](https://jvet-experts.org/doc_end_user/current_document.php?id=13296) A Study on Decoder Interoperability of Generative Face Video Compression [B. Chen, S. Yin, J. Chen, Y. Ye (Alibaba), S. Wang (CityU HK)]

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, including 2D keypoints, 2D landmarks, 3D keypoints, temporal motion features, facial semantics and other formats. At the 31st JVET meeting, JVET-AE0080, JVET-AE0083 and JVET-AE0280 were proposed to support more common syntax design of generative face video SEI message and define the interface between the decoder (which decodes the base pictures and parses the SEI message) and the generative neural network.

One question that was raised during the discussion of these contributions was the interoperability between different facial representations and their associated decoder networks. This contribution attempts to shed some lights on how the interoperability problem might be solved. Specifically, we further analyze the decoder design of generative face video SEI message, and discuss how to support the decoder interoperability and parameter translatability in the two following schemes:

1. **Scheme 1**: recognizing that the end-to-end decoder consists of two independent modules, optical flow estimation module and generator module, we retrain the original end-to-end decoder as independent modules, where we fix the generator module and train the optical flow estimation module to adapt to different types of parameters.
2. **Scheme 2**: we establish a generalized face parameter translator so that different facial representations can be converted into the required type that can be accepted by a fixed decoder.



Analysis and synthesis don’t have to be jointly optimized, a translator aligns the information needed.

Quality somewhat suffers relative to the jointly optimized case, but still acceptable.

The SEI message would represent 3D keypoints.

The network parameters of the translator could also be sent (1.3 million parameters).

The fixed synthesis model would not be mandatory.

It was commented that rather than supporting various approaches, it might be more appropriate selecting one of them.

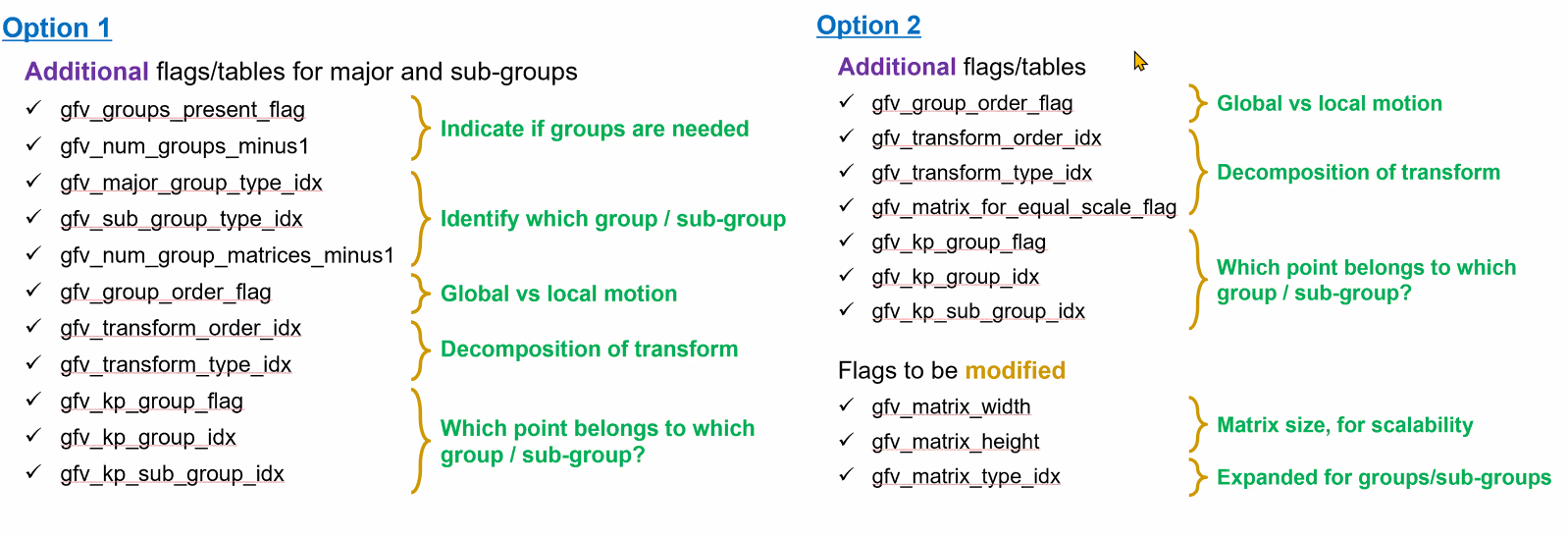
[JVET-AF0146](https://jvet-experts.org/doc_end_user/current_document.php?id=13404) AHG9: On Face Motion Information for Generative Face Video [H.-B. Teo, J.-Y Thong, K. Jayashree, C.-S. Lim, K. Abe (Panasonic)]

At the 29th JVET meeting, JVET-AC0088 presented a generative face video SEI message to allow for compression of talking-face videos at ultra-low bitrate. An expansion of its scope was presented in JVET-AD0051 at the 30th JVET meeting to encompass multiple facial representations. During the 31st JVET meeting, further enhancements to the generative face video SEI message syntax were propositioned in JVET-AE0280 which combined ideas proposed in JVET-AE0080 and JVET-AE0083. Also, the informative proposal JVET-AE0088 demonstrated the effectiveness of 3-D facial landmarks for generating talking-face videos, and the framework’s independence across number of landmarks for generation.

This contribution aims to provide further details presented in JVET-AE0088 and enhanced refinements to the generative face video SEI message from JVET-AE0280.

A concept of “subgrouping” landmark points such as eyes, contour and mouth is introduced to support more complex facial actions while keeping rate low by only sending changes in specific subgroups.

Affine matrix represents 3D global movements



The approach is primarily for efficient representation of keypoints.

[JVET-AF0234](https://jvet-experts.org/doc_end_user/current_document.php?id=13497) 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 updates the common text of generative face video (GFV) SEI message proposed in JVET-AE0280 by introducing the parameter translator proposed in JVET-AF0048 which attempts to solve the interoperability issue between different facial representations and their associated decoder networks. 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 SEI message contains

* Keypoints (2D or 3D)
* Matrices of different types (e.g. affine, cov, mouth,…)
* Optionally a translator (using NNR)

Depending on the algorithm at the encoder, not all of these parameters are present in the SEI.

Depending on the parameters received, and the synthesis algorithm available, a certain translator may be necessary.

Translator and synthesizer are NNs.

If a decoder does not have a suitable translator and synthesizer available for the parameters found in the SEI, it would give up.

Basically by this concept, many different formats would be supported.

It was concluded that establishing a new AHG is the best way of studying generative face, providing software, coordinating experiments, investigating which parameters would be needed, how they can be efficiently coded, interoperability aspects, quality, stability, etc.

The AHG should also be mandated to develop a document summarizing the potential approaches.

## Chairs: Y. Ye, H.-B. Teo, Z. Lyu, S. McCarthyAHG9: Source picture timing information SEI message aspects (6)

Contributions in this area were discussed at 1830–2200 on Monday 16 Oct. 2023 (chaired by JRO).

[JVET-AF0055](https://jvet-experts.org/doc_end_user/current_document.php?id=13303) AHG9: On Source Picture Timing Information (SPTI) SEI message [[J. Samuelsson-Allendes](mailto:samuelssonj@sharplabs.com), S. Deshpande (Sharp)]

This input document relates to the Source Picture Timing Information (SPTI) SEI message included in the Technology under Consideration for future extensions of VSEI (draft 1). The following items are proposed:

1. Syntax and semantics modification is proposed such that an SPTI message cannot contain information about pictures in lower temporal layers.
2. Syntax modification is proposed to condition the presence of the sublayer indicator on that persistence flag equals 1.
3. It is proposed not to signal spti\_sublayer\_source\_picture\_interval\_scale\_factor[ spti\_max\_sublayers\_minus1 ] for the highest temporal sublayer and instead infer its value.
4. It is proposed to signal the time scale (spti\_time\_scale) with a minus 1 coding to syntactically avoid signalling the value 0 for time scale.
5. Addition of text for sections “Use of source picture timing information” for the relevant standards is proposed.
6. Syntax rearrangement is proposed such that the syntax element spti\_source\_picture\_timing\_type is signalled only when spti\_source\_timing\_equals\_output\_timing\_flag is equal to 0.

The items are independent of each other, except for item 2 which builds on top of item 1. The second version of this document adds the 6th item.

For item 1, it was explained that the SEI message would be lost if a higher layer is discarded in which it is associated.

It was commented that an alternative to the solution proposed would be to repeat the SEI in lower layers.

It was also commented that there might also be a problem with splicing.

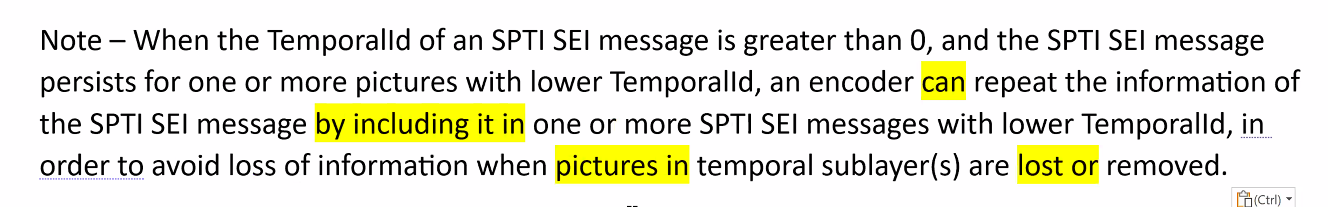
It is recognized that there is a problem, but it was commented that there might be an option of adding a note rather than the suggested solution.

The change proposed in context of item 1 (avoiding duplication of interval scale factor when the SEI message is sent in higher layers) is asserted not to be necessary.

Item 2: It is agreed that conditioning the syntax element “spti\_max\_sublayer\_minus1” on persistence is appropriate, but the rule of inferring it has to be modified as item 1 was not adopted.

Though it is uncommon that the SEI message will be often applied together with temporal scalability (or only for one-time playout where the loss is irrelevant, it is agreed that a note should be added relating to item 1.

The following text was agreed to be included in a follow-up discussion on Thu 19 Oct.:



Item 3: It was commented that in some applications the source picture interval scale factor might not be 1, so inferring it might not be appropriate in all cases. However, the semantics may not be clear on that. Editorial action to clarify that the interval at highest layer is not necessarily 1.

Item 4: This would not be consistent with expression of similar syntax elements in other places (e.g. HRD). No action.

Item 5: Alignment to syntax expression in different standards – agreed.

Item 6: It was commented that the syntax element timing\_type could have benefit independent of the subsequent timing\_flag, though both have somewhat overlap in meaning. It is kind of descriptive metadata. It was later agreed in the context of discussing JVET-AE0069 to adopt item 6, however using the **spti\_source\_type\_present\_flag** for possible gating of the syntax element.

[JVET-AF0069](https://jvet-experts.org/doc_end_user/current_document.php?id=13317) AHG9: On source picture timing information SEI message [S. McCarthy, G. J. Sullivan, P. Yin (Dolby)]

It was agreed at the previous meeting of JVET to include the source picture timing information (SPTI) SEI message in the TuC for VSEI extensions. This contribution proposes modifications to the SPTI SEI message to address comments from the previous JVET meeting.

This contribution proposes the following items to address the comments from the previous meeting and to generally improve the SPTI SEI message:

1. Add syntax element, spti\_source\_type\_present\_flag, to condition the presence of syntax elements indicating the timing relationship between source pictures and corresponding decoded output pictures.
2. Simplify syntax element name spti\_source\_picture\_timing\_type to spti\_source\_type.
3. Change the precision of spti\_source\_type from u(8) to u(16).
4. Simplify the syntax element name spti\_num\_units\_in\_elemental\_source\_picture\_interval to spti\_num\_units\_in\_elemental\_interval.
5. Simplify the syntax element name spti\_sublayer\_source\_picture\_interval\_scale\_factor[ i ] to spti\_sublayer\_interval\_scale\_factor[ i ].
6. Change the precision of spti\_num\_units\_in\_elemental\_interval from u(32) to u(18)
7. Add text to clarify the descriptions of the various source types, e.g., “slow motion”, “high-speed imaging”, etc.
8. Add semantic constraints to prevent mutually exclusive timing relationships between source pictures and corresponding decoded output pictures. Specifically, prevent the combination of “high-speed imaging” and “time-lapse imaging”.
9. Replace syntax element spti\_source\_timing\_equals\_output\_timing\_flag with spti\_source\_timing\_info\_present\_flag and add corresponding semantics.
10. Move specification of the variable temporalReversalFlag to the semantics following spti\_sublayer\_interval\_scale\_factor[ i ].
11. Integrate the variable temporalReversalFlag in the equation for SourcePictureInterval[ i ] and remove the equation for SourcePictureTime[ i ] (i.e., the absolute source picture time).

On item 1: It was agreed that the name of the previous **spti\_source\_timing\_equals\_output\_timing\_flag** should not be modified into **spti\_source\_timing\_info\_present\_flag.** If it is zero, the **spti\_source\_type** shall be sent (with its gating flag).

All other items 2-11 are agreed.

[JVET-AF0097](https://jvet-experts.org/doc_end_user/current_document.php?id=13344) [AHG9] On description of picture 0 in source picture timing information SEI message [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

It is asserted that the current description regarding picture 0 in the draft text of source picture timing information (SPTI) SEI message is not clear which may cause confusion. This contribution proposes some updates to the current text as follows:

1. Specify that picture 0 for computation of SourcePictureTime[ 0 ] of an SPTI SEI message is the picture in the same access unit that contains the SEI message.
2. Add a constraint that picture 0 shall not be a synthesized picture since it does not have corresponding source picture.

Item 1: It was commented that usally picture 0 is understood as being the first picture in the bitstream. On the other hand, in terms of the SEI message, only relative distances between pictures are relevant.

According to the proponent, the timing before the SEI message may be unknown. Therefore, it would be more reasonable to start counting at that point. This is agreed.

The denotation as “picture 0” is confusing and should be discarded. The proponent was asked to modify the text accordingly. The updated text is included in v2 of the contribution. It was agreed that the following text in that version shall be included in the next version of the SPTI SEI.

The information provided by the SPTI SEI message pertains only for picture(s) starting from the picture in the current layer in the access unit that contains the SPTI SEI message and all subsequent pictures of the current layer in output order based on its persistence.

Item 2: It was commented that there may be cases where this is inappropriate.

[JVET-AF0098](https://jvet-experts.org/doc_end_user/current_document.php?id=13345) [AHG9] On temporal reversal feature in source picture timing information SEI message [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

It is asserted that the current methods for assigning the value of SourcePictureTime[ 0 ] in source picture timing information (SPTI) SEI message does not work. While the method of assigning it with externally provided value may work, the other method by setting the source picture timing for that first output picture to be equal to 0 does not work since it is simply wrong.

This contribution proposes two options to overcome the asserted problem:

**Option 1:** when type includes temporal reversal, allow possibility to explicitly signal the elemental distance between the first picture and the last picture for the temporal reversal and the following applies:

1. When the elemental distance between the first picture and the last picture is provided, the SourcePictureTime[ 0 ] is set to this value.
2. When the elemental distance between the first picture and the last picture is not provided, it is constrained that the value is provided externally.
3. The distance information shall not be present when the SPTI SEI message does not include temporal reversal type.

**Option 2:** specify a constraint such that when the SPTI SEI message includes temporal reversal picture timing type (i.e., the value of temporalReversalFlag is equal to 1), the value of SourcePictureTime[ 0 ] shall be provided by external means.

It was commented that it may not be necessary to make a change, as the picture where the SEI message occurs is the starting point where instead of counting time forward it is counted backward. An application would be able to manage this (perhaps an editorial note on this).

No action necessary.

[JVET-AF0099](https://jvet-experts.org/doc_end_user/current_document.php?id=13346) [AHG9] On signaling of source picture interval scale factor and sublayer synthesized flag in source picture timing information SEI message [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

It is asserted that syntax elements spti\_max\_sublayers\_minus1, spti\_sublayer\_source\_picture\_interval\_scale\_factor[ i ] and spti\_sublayer\_synthesized\_picture\_flag[ i ] are not needed when the SEI message persists only for one picture (i.e., for picture in the same access unit that contains the SEI message). Furthermore, when the persistence of the SPTI SEI applies only to the current picture, the computation of SourcePictureInterval[ ] and SourcePictureTime[ ] is not necessary and may cause confusion.

This contribution proposes the following modifications to overcome the asserted problem:

1. Signal syntax elements spti\_max\_sublayers\_minus1, spti\_sublayer\_source\_picture\_interval\_scale\_factor[ i ] and spti\_sublayer\_synthesized\_picture\_flag[ i ] only when the value of spti\_persistence\_flag is equal to 1.
2. Update the description related to SourcePictureInterval[ i ] and SourcePictureTime[ n ] such that they apply only when spti\_persistence\_flag is equal to 1.
3. Specify that when persistence of the SPTI SEI applies only to the current picture, SourcePictureInterval is set to be equal to ElementalSourcePictureInterval.

Item 1: Partially similar as item 2 in JVET-AF0055 – it had been agreed that spti\_max\_sublayers\_minus1 can be gated based on persistence flag, and derived.

On Items 2 and 3, it was commented that picture interval and picture time would also be needed for pictures following the current picture, similar to HRD. The only case where this would not be the case is when the SEI only applies to a single picture. No action.

[JVET-AF0100](https://jvet-experts.org/doc_end_user/current_document.php?id=13347) [AHG9] On miscellaneous aspects of source picture timing information SEI message [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

This contribution proposes several modifications that are asserted to provide improvement to some aspects of the source picture timing information (SPTI) SEI message design.

Item 1: value of ElementalSourcePictureInterval: Similar comment as for item 4 of JVET-AF0055. No action, the valid value range is also specified.

Item 2: Time reversal and persistence: No reason to disallow persistence, when time is counted down (see similar discussion under JVET-AF0098)

Item 3: Synthesized temporal sublayer: No reason to impose a constraint.

## AHG9: Encoder optimization information SEI message aspects (4)

Contributions in this area were discussed at 1700–1815 on Sunday 15 Oct. 2023 (chaired by JRO).

[JVET-AF0052](https://jvet-experts.org/doc_end_user/current_document.php?id=13300) AHG9: Proposed updates on the encoder optimization information SEI message [M. M. Hannuksela, F. Cricri (Nokia)]

This contribution proposes the following updates to the encoder optimization information SEI message:

1. Editorial updates to the phrasing of the cancellation and persistence.
2. A syntax element eoi\_object\_based\_idc to indicate the type of object-based optimization, including blurring, quantization adjustment, and overwriting sample values of areas outside the detected objects.
3. Replacement of the 16-bit optimization\_type by individual flags followed by reserved zero bits.
4. Replacement of the optimization\_ prefix in the syntax element names with eoi\_ to obtain shorter syntax element names.
5. Sensibility constraints that when the persistence is for the current picture only, the temporal optimizations (eoi\_temporal\_subsampling\_flag and eoi\_temporal\_quality\_flag) are required to be off.
6. For temporal subsampling, addition of eoi\_num\_int\_pics, which is indicative of the count of pictures that the encoding systems excluded between each pair of coded pictures in output order.
7. For temporal quality optimization, a clarification of the semantics and addition of a related NOTE.

Agreed on all items to be included in TuC except item 3 (editorial, JVET-AF0107 has an alternative formulation)

[JVET-AF0068](https://jvet-experts.org/doc_end_user/current_document.php?id=13316) AHG8/AHG9: Signalling of encoder preprocessing information [W. Jia, Y. Li, Y.-K. Wang, K. Zhang, L. Zhang (Bytedance)]

This contribution proposes to add, to the encoder optimization information SEI message in the TuC document JVET-AE2032, the signalling of the type of the spatial subsampling filters used in encoder preprocessing when the optimization\_type indicates spatial subsampling.

It is asserted that this information could be useful for both human and machine consumption. It is well known that pairs of down- and up-sampling filters need to be matched for optimum results (e.g. in spatial scalability). Beyond the filter types proposed in the contribution, filters specified by coefficients, information about subsampling phase, and decimation factor would be relevant, as well as potential different processing for luma and chroma. Also, the description in the table is very generic, more explicit definition of the meaning of the different filter types. The meaning of “mask for interpolation types” is unclear.

More study on these aspects recommended.

[JVET-AF0105](https://jvet-experts.org/doc_end_user/current_document.php?id=13352) AHG9: On color vision deficiency optimization type for encoder optimization info SEI message [C. Kim, Hendry, D. Gwak, J. Lim, S. Kim (LGE)]

This contribution proposes an optimization type that describes optimization applied by encoder related to color modification for human viewing. This optimization may be intended for human viewing with color vision deficiency. When a video is optimized for individuals with color vision deficiency, it may be uncomfortable for those with normal color vision to watch, thus, such information provided in the SEI may be helpful. The proposed SEI message contains the following information:

* Indication whether the optimiztion for the picture(s) is for color vision deficiency.

Presented Monday 16 Oct.

It was asked for what purpose this is useful. Several experts pointed out that such a color correction would rather be done at the decoder side.

No action.

[JVET-AF0107](https://jvet-experts.org/doc_end_user/current_document.php?id=13354) AHG9: On miscellaneous aspects for encoder optimization info SEI message [C. Kim, Hendry, D. Gwak, J. Lim, S. Kim (LGE)]

This contribution proposes the following modifications to the encoder optimization information SEI message:

1. Updates on the optimization\_type.

* Simplification of optimization\_type table(editorial)
* Changing temporal/spatial subsampling to temporal/spatial resampling

1. Add a constraint on optimization\_for\_human\_viewing\_flag and optimization\_for\_machine\_analysis\_flag such that they cannot be equal to 0 at the same time.

For the change suggested on item 1, semantics would need to be added how the bitmask is to be combined with the optimization\_type syntax element. It was commented that in terms of software implementation, an explicit expression via flags is better understandable.

The proponent was asked to provide a formulation of semantics for the interpretation of optimization\_type. This was presented from v2 of the contribution on Thu. 18 Oct. It was agreed that the corresponding editorial change (using bitmask in the table, and inferring flags from the syntax element) should be made (yellow highlighted in the contribution).

Agreed on re-formulation spatial/temporal resampling (which would also apply to upsampling).

Item 2 is not deemed necessary, as same state of the flags indicates that it is not specifically optimized for one of them, regardless if both are 0 or both are 1.

It is further reported that in the combination with the adoption from JVET-AF0052 and JVET-AF0107, a discrepancy was found that is resolved in the green highlighted parts of v2 of JVET-0107. It was also agreed to include these changes in the updates of the EOI SEI in the TuC.

It is noted that v2 of JVET-0107 in section 2.1 has an integration of the texts from the different proposals that were adopted.

## AHG9: Object mask information SEI message aspects (2)

Contributions in this area were discussed at 0840–0930 on Tuesday 17 Oct. 2023 (chaired by JRO).

[JVET-AF0087](https://jvet-experts.org/doc_end_user/current_document.php?id=13335) AHG9: Experimental results of object mask auxiliary picture coding [Z. Zhang, [J. Chen](mailto:jiechen.cj@alibaba-inc.com), Y. Ye, S. Wang (Alibaba)]

The object mask information (OMI) SEI message was included in the first draft of technology under considerations for future extensions of VSEI at the last meeting. With OMI SEI message, the object masks are sent as auxiliary pictures. This contribution reports the experimental results of coding object masks as auxiliary pictures with VTM layered coding. According to the coding results, it is reported that it is practical to code the mask pictures as auxiliary pictures.

The percentages given in the contribution are reporting the part of the mask bit rate vs. overall bit rate.

The difference between lossy and lossless mask coding is not large, but dependent per sequence

Lossless coding uses low QP (as per CTC config files)

It was suggested to investigate whether palette mode could further reduce the rate.

It was found that inter prediction does not have much benefit for the masks.

Generally, it was commented that results are very interesting, indicating that the concept works with reasonably low bit rate, at least in the case of the sequences shown. The test cases were masks generated from classes A-E, with maximum of 4 masks. It was also commented that for more objects higher bit rates would be necessary.

In lossy coding case, the number of wrong samples is very small (10-3 at QP 37 at most).

It was asked to make the mask sequences available via the JVET ftp. Agreed.

It was also asked to provide the config files in an update of the contribution.

[JVET-AF0088](https://jvet-experts.org/doc_end_user/current_document.php?id=13336) AHG9: Software implementation and further fixes of the object mask information SEI message [J. Chen, Z. Zhang, Y. Ye, S. Wang (Alibaba)]

This contribution provides the software implementation of the object mask information (OMI) SEI message and also proposes fixes and further changes to the OMI SEI message in the draft 1 of technologies under consideration for future extensions of VSEI.

The proposed fixes and changes include the following four aspects:

* Aspect 1: The text related to bounding box parameters was fixed and refined.
* Aspect 2: The binarization of bounding box parameters was changed from ue(v) to u(16)
* Aspect 3: A gating flag for bounding box parameters was added to give signaling flexibility to the encoder
* Aspect 4: The parsing dependency among different OMI SEI messages was removed by always signaling omi\_mask\_cancel[ i ][ j ][ k ]

It was commented that aspect 4 is not a parsing dependency as normally understood. It is more for simplifying storage at decoder.

Agreed that the changes 1-4 are appropriate, and shall be integrated into the next TuC.

It was proposed to create a branch in VTM software for this and possible future contributions on SEI messages in TuC. It was commented that providing software for SEI messages in TuC is strongly encouraged, to enable other experts studying them.

Agreed to include the software from JVET-AF0088 into the TuC branch of VTM.

## AHG9: Other SEI topics (11)

Contributions in this area were discussed at 0930–1055 and 1530–XXXX on Tuesday 17 Oct. 2023 (chaired by JRO).

[JVET-AF0138](https://jvet-experts.org/doc_end_user/current_document.php?id=13396) AHG8/AHG9: Truncated bit depth support SEI messages [D. Ding, X. Zhao, S. Liu, G. Teniou, S. Wenger (Tencent)]

Currently, video content is typically represented with eight or higher bit depth, such as 8-bit or 10-bit, which has been widely adopted for video communication perceived by human vision. However, it is observed through experimental results that, for machine vision/tasks, such as object tracking and detection, the quality score, e.g., measured by detection rate, shows lower sensitivity to the bit depth. Therefore, a lower sample value range, which can be represented by fewer bits per sample than 8 bits, can achieve a better trade-off between bitrate and machine task score.

In this contribution, it is proposed to include SEI messages signaling the actual sample value range (or bit depth) of the input video, which is represented by 8-bit or higher bit depth.

The idea is that only the n LSBs are used when an n-bit video is input into the decoder. It is verbally reported that this provides benefit in terms of improving BD-mAP.

From the experience with IBDI, such an approach would generally not have benefit for video that is encoded for human consumption.

It appears that the benefit of this proposal would be mainly for VCM. It was negotiated in joint meetings with WG 4 that JVET would not define metadata that are specifically useful only for machine vision applications. This would be a typical case of an SEI message that should be managed within WG 4.

[JVET-AF0140](https://jvet-experts.org/doc_end_user/current_document.php?id=13398) [AHG9] Inter Picture Dependency SEI message [G. Teniou, S. Wenger, S. Liu (Tencent)]

Nowadays, video applications rely more and more on mixed hardware/software-based implementations, particularly for decoders/transcoders located in the Cloud. Such cloud media processing benefits from massive parallel computing resources such as GPUs. In many scenarios, the decoding process needs to happen as fast as possible. As an example, live TV programs that are later proposed as on-demand contents, need to be made available as soon as possible. The TV program from the live stream is decoded and re-encoded into multiple formats and resolutions suitable for on-demand streaming. This contribution proposes to address the fast-decoding mode by signaling pictures that can be decoded independently.

It was commented that the SEI message would somehow duplicate the information from RPL. The proponent argues that the purpose is knowing the structure beforehand to allocate ressources. This is particularly useful in cloud applications, where resources are shared between multiple bitstreams, and some of them come with referencing structures allowing more parallelism at the picture level.

It was commented that HEVC has a structure of pictures SEI message that has very similar concepts.

Proponents are asked to study commonalities with the existing SEI message of HEVC, and possible usage of the latter for their purpose. Also, more evidence should be brought about the actual benefit in terms of sharing ressources by imposing the restrictions, including some more examples of restricted picture structures.

[JVET-AF0141](https://jvet-experts.org/doc_end_user/current_document.php?id=13399) AHG9: SEI messages for common image metadata formats [A. Hinds, S. Wenger (Tencent)]

Neural network (NN) post filtering shows increasing potential for emerging use cases (e.g., video coding for machines) with its pairing of SEI metadata and imagery suitable for a congruently trained NN in a post filtering process. Likewise, generative AI applications for the creation and or capture of still and moving imagery, are increasingly growing both in popularity and sophistication. However, for generative AI applications, the number and different types of inputs that these applications require to achieve satisfactory results can vary widely.

This contribution proposes specifications for the carriage of popular image metadata formats in SEI messages to facilitate access to additional information about the content source and context of image the capture process for use by, e.g., generative AI applications.

It was pointed out that referencing EXIF in an ITU standard would currently problematic. XMP and JFIF could be more viable, being ISO standards.

Though there are other places in systems layer (e.g. file format) where such metadata could be transported, a certain benefit could be seen of carrying them in the video stream.

JFIF also allows carrying thumbnails. This would be restricted to 64 kByte, which appears manageable.

Generally agreed that having such an SEI message would have advantages. Further investigation necessary how the proposed formats could be referenced in an ITU-T/ISO standard.

Agreed to include in TuC with an approprioate note about that problem.

[JVET-AF0142](https://jvet-experts.org/doc_end_user/current_document.php?id=13400) [AHG9] Local Film Grain Synthesis using Annotated Regions SEI [G. Teniou, S. Wenger (Tencent)]

See section 4.12

[JVET-AF0144](https://jvet-experts.org/doc_end_user/current_document.php?id=13402) [AHG9] Adaptive Film grain synthesis using Alpha Channel Information SEI message [G. Teniou, S. Wenger (Tencent)]

See section 4.12

[JVET-AF0145](https://jvet-experts.org/doc_end_user/current_document.php?id=13403) AHG9: SEI message for text for generative AI [A. Hinds, S. Wenger (Tencent)]

Text prompts serve as an important element to generative AI image and video generator applications. The sequence of images or video frames to be used as the source (inputs) visual basis for the resulting image or video generated outputs, from a generative AI process, can be carried in a coded video stream. Likewise, for the purposes of provenance, it may be important that such applications be able to record that generative AI operations were performed to create the output. This contribution proposes that a new SEI be created to store the text data for two use cases related to generative AI.

The intent is to have a synchronized way of adding an animation on top of a video. The animation would need to be generated externally (typically in the cloud) by a mechanism not specified in the SEI message. Typically, a user would say what he wants to do with the video at a certain time position, then it is sent to a cloud server, and after the animation is generated, it is sent back in a processed way.

It is commented that often generative AI services have their own interfaces, and it is not clear if the SEI message would be useful for them, as ofte they would build their own service interface.

It was asked if it would be reasonable to add information about actions to be generated at certain parts of the picture.

One comment is given that the semantics has a “shall” in a wrong place.

Unclear if an SEI is needed for that use case.

The other use case is to transport a text along with a video that was itself generated by generative AI.

It was asked why that would only be done for AI?

It was was requested to investigate whether the copyright SEI might be sufficient to signal that a certain part of the video was generated by a specific generative AI.

[JVET-AF0147](https://jvet-experts.org/doc_end_user/current_document.php?id=13405) AHG8/AHG9: On Picture Modality Type [J. Gao, H.-B. Teo, C.-S. Lim, K. Abe (Panasonic)]

This contribution proposes to add picture modality types in VSEI. The picture modality type such as visible picture and infrared picture tells the nature of the picture data. This distinction can be crucial for applications where the processing, interpretation, or display of the decoded picture is influenced by its spectral properties. Hence, it is proposed to add three picture modality types (visible, infrared and thermal infrared) in VSEI by one of the following options:

- Option 1: Signal picture modality type in Encoder Optimization Information (EOI) SEI message in JVET-AE2032.

- Option 2: Signal picture modality type in VUI parameters.

- Option 3: Signal picture modality type in Scalability Dimensional Information (SDI) SEI message in VSEI.

It is proposed to adopt one of the proposed options to VSEI.

It was commented that infrared imaging has many applications, signalling the source properties at a high level in VVC would be interesting.

Option 1 would not be appropriate, as the EOI SEI is about encoder action, not the source.

Option 2 (VUI) could be done without defining a new profile – several experts expressed support for that. This would for example allow representing an IR video standalone (monochrome), or associate it with an RGB source video with an infrared video as additional layer (therefore, the functionality would not be different from option 3).

Agreed to include option 2 in TuC.

In a follow-up discussion on Thu. 18 Oct., the question was raised if this additional VUI signalling would also be beneficial to be included in the next versions of AVC and HEVC. This was left for further study (not in DIS/DAM), and it is noted that NB comments would be required to make such a change at FDIS/FDAM stage.

[JVET-AF0148](https://jvet-experts.org/doc_end_user/current_document.php?id=13406) [AHG9] Large SEI message signalling [G. Teniou, S. Wenger (Tencent)]

In nowadays video coding standards SEI messages serve multiple purposes e.g., improving the rendering quality or helping the receiver to apply the correct post-processing techniques.... Those SEI messages, particularly the most recent ones, defined in VSEI have a significant size impacting the signalling of the associated header. The present contribution proposes a complementary approach suitable for large SEI messages.

Target is to reduce the transmission overhead of current SEI bitstream packaging (signalling FF at every 256th byte position).

It was commented by several experts that looking for a change of the design which dates back to AVC is useful. The question is if this is appropriate to do for existing standards, or starting cleanly with a new standard.

Relation of syntax element for position with prefix/suffix SEI NAL units?

Relation of syntax element for relevance with manifest SEI?

Introducing this mechanism for existing SEI messages would have the consequence that they could not be interpreted by legacy devices. In existing standards, this should – if at all – be done for new messages.

The amount of bit saving is not huge, therefore no need for immediate action.

Agreed to include in TuC.

[JVET-AF0149](https://jvet-experts.org/doc_end_user/current_document.php?id=13407) [AHG9] Application-required SEI NAL Units [G. Teniou, S. Wenger (Tencent)]

Many audiovisual service architectures rely on media-aware network elements that in under certain conditions might drop SEI NAL Units, as they are, by definition, not required for the luma and chroma decoding process. However, some of those SEI messages are also required by some applications to ensure a consistent quality of service and experience across the users. This contribution proposes a solution to signal the SEI messages required to be maintained in the bitstream by an application.

A new NAL unit type is proposed which is by itself empty, but signals that the subsequent NAL unit should not be dropped by a middle box.

If it is present, this would not mean that it is mandatory to use the subsequent SEI message at the decoder. The purpose is that a middlebox would not drop it, when an SEI message would be mandatory in a certain application domain.

It was asked if this could not be achieved by the manifest SEI? The proponent claims that this is transported as any other SEI, and is difficult to detect in a middlebox.

The proponent is requested to upload an update of the contribution. The contribution presented was describing another version of the proposal with a non-empty NAL unit.

The contribution was again presented after the availability of the intended variant in v2 of the document (v4 in the website, where three identical copies had been uploaded, and two of them were therefore rejected).

Three different options were presented, where option 2 is the one that had been verbally described. Option 1 is the one that had been contained in the first version of the contribution.

The new version also includes a new option 3, where the new NAL unit type also has a payload that expresses that the non-discarding shall be applied to n subsequent SEI NAL units.

Due to the fact that new possible considerations were proposed very late, and option 3 might be interesting but would again require more complexity in the MANE further study seems to be necessary.

The proposal would only be appropriate if the new NAL unit type would not have any impact on decoder operations (neither legacy decoders than newly defined ones). For the decoder, existing mechanisms such as manifest, prefix, etc. are already defined for the given purpose.

[JVET-AF0167](https://jvet-experts.org/doc_end_user/current_document.php?id=13425) AHG9: On the Proposed Multiplane Image Information SEI Message [T. Lu, P. Yin, S. Oh, S. McCarthy, W. Husak, G. J. Sullivan (Dolby)]

was presented Thursday 19 Oct. 0900.

This contribution provides an analysis and comparison of the capabilities and limitations of MIV (V3C) and the multiplane image information (MPII) SEI message proposed in JVET-AE0066. It is reported that the MIV specification provides flexibility to enable a variety of high-quality immersive video representations on devices that implement an MIV format decoder and one or more conventional 2D video decoders. In contrast, the proposed MPII SEI message is designed for the specific use case of distributing single-view MPI content to existing smartphones and other consumer devices equipped with a single conventional 2D decoder. It is also reported that the current MIV specification and TMIV reference software lack features required for the decoding of MPI content by devices equipped with only a single conventional 2D decoder. Temporal interleaved MPI representations are discussed, with commentary on relative bit rate overhead and implementation complexity of MIV and the proposed MPII SEI message. A key point is that a bitstream carrying an MPII SEI message can be decoded by an ordinary 2D video decoder, as contrasted with extraction of sub-bitstream NAL unit fragments from the MIV bitstream, followed by one or more ordinary 2D video decoders operating in parallel on the extracted data. It is suggested that an appropriate justification for applying the more general MIV approach is when the number of layers in an MIV representation is sufficiently large and the occupancy of each layer is sufficiently small to justify using an MIV atlas to reduce the amount of data processed through the video decoding processes.

This contribution also provides a software implementation of the proposed MPII SEI message for the VTM reference software.

The v2 version adds a high level summary table comparing the MPII SEI and the MIV extended restricted geometry profile for the use case of distributing single-view MPI content to existing devices equipped with a single conventional 2D decoder.

It was commented that for investigating the software it would be desirable to have test material available. The proponents agreed to provide test material, and also some test material from WG 4 could be used.

Three different packing schemes are proposed: top/bottom, left/right, temporal interleaving. It was asked if all of them are useful.

The proponents explain that for temporal interleaving larger picture sizes could be supported in the same level of a profile. It was pointed out that a consequence would be that less reference pictures could be used at maximum, such that there might be a loss of compression efficiency.

Could the software for the synthesis also be provided? The proponents confirmed this would be possible, it is open source. A link where it can be retrieved, and short description how to run it should be provided.

It was asked if a quality analysis had been made using similar metrics as used in MIV? The contribution JVET-AE0066 contained such results, but did not compare directly against MIV.

In the experiments, 16 layers had been used.

The software will be provided in an update of the contribution.

It was agreed to include JVET-AF0167 in the TuC, and the software in the VTM SW branch that will be created for SEI messages in the TuC.

[JVET-AF0255](https://jvet-experts.org/doc_end_user/current_document.php?id=13518) Crosscheck of JVET-AF0167 (AHG9: On the Proposed Multiplane Image Information SEI Message) [R.-L. Liao (Alibaba)] [late]

[JVET-AF0306](https://jvet-experts.org/doc_end_user/current_document.php?id=13570) Essentiality of manifest SEI for film grain [A. Tourapis, D. Podborski (Apple)] [late]

See section 4.12

## Non-SEI HLS aspects (0)

Section kept as a template for future use.

# Plenary meetings, joint meetings, BoG reports, and liaison communications

## JVET plenaries

An intermediate plenary was held on Monday 16 Oct. 1600. The following items were discussed:

* Short report and confirmation of decisions from the tracks on EE2 related (Y. Ye) and HLS (JRO) that had been held on Saturday and Sunday
* Planning of schedule for remaining week
* Expert viewing
* Joint meetings
* Planning of outputs: Liaison, ballot responses

Further detail on scheduling is recorded in section 2.15.

Joint meetings involving JVET were held as follows:

* JVET, XXX on XXX topics, on XXday XX Oct. at XXXX–XXXX
* Further detail about these sessions with other groups is provided in the other subsections of this section.

General plenary wrap-up discussions are recorded under sections 8, 9, and 10.

## Information sharing meetings

Information sharing sessions with other WGs and AGs of the MPEG community were held on Monday 16 Oct. 0900–1130, Wednesday 18 Oct. 0900–1030, and Friday 20 Oct. 1400–1600.

The status and plans for the work in the MPEG WGs and AGs was reviewed at these information sharing sessions.

## Joint session 11am Tuesday 17 Oct on Generative face video: WG 2 Requirements, WG 5 / JVET, WG 7 3DG, VCEG

(notes by GJS)

Relevant Convenors & Rapporteurs present: I. Curcio (WG 2), J.-R. Ohm (WG 5 / JVET), M. Preda (WG 7), G. Sullivan (VCEG), Y. Ye (VCEG, but as presenter); also J. Ostermann of AG 2.

Presentation of M64987/VCEG-BT04 [Yan Ye (Alibaba), Sean McCarthy (Dolby), Han Boon Teo (Panasonic), Zhuoyi Lv (vivo), Shiqi Wang (CityU), Kai Zhang (Bytedance), Marta Karczewicz (Qualcomm), Iole Moccagatta (Intel)] **On VVC-assisted ultra-low rate generative face video coding**

This contribution provides an overview of generative face video coding algorithms and asserts that they can achieve significantly better (perceptual) reconstruction quality than the state-of-the-art VVC standard at ultra-low bitrate ranges. JVET has had inputs studying these algorithms since the beginning of this year. The proponents of this contribution suggest that JVET should be tasked with the exploration and potential standardization of ultra-low bitrate 2D generative face video coding methods.

The demonstration uses key pictures encoded using a conventional 2D video coder (e.g., VVC) plus some parameters extracted from individual frames to generate corresponding pictures at the decoder that are generated using the decoded key pictures and the parameters.

Low-delay parameter generation is used (i.e. no picture look-ahead), so this could be used for low-delay conversational services (given sufficient computing resources).

The contribution proposes to explore and standardize such technology in JVET.

Possibilities of SEI or profile specification are suggested.

The focus is on ultra-low-bit-rate 2D content.

As tested, the decoder is fixed (not using a model customized for the particular content sequence).

See also: [JVET-AF0234](https://urldefense.com/v3/__https:/jvet-experts.org/doc_end_user/current_document.php?id=13497__;!!HOHtwYw!CfTLqHXCf8tjcq1eYxeGnb96WoaYAuuzKyznAdUE-zU4LYApY0hfr2WFsZSUZFFPAOY8Rkz6FUps_n-4O3uGxiv1$) / m65470 [B. Chen, J. Chen, Y. Ye (Alibaba), S. Wang (CityU), S. McCarthy, P. Yin, G.-M. Su, A. K. Choudhury, W. Husak, G. J. Sullivan (Dolby)] AHG9: Common text for proposed generative face video SEI message

The input contributions refer to available software for experimentation.

This is presented as a 2-D video coding technology rather than an “avatar” or 3-D modeling technology.

It was asked what type of video content is considered, and the presenter responded that it is “head and shoulders”. Other content may not be handled well.

There was a question about interoperability – especially how clearly specified the decoder would be.

It was commented that this is a “tool” for a particular type of content. Is it only for faces or potentially for other content. Studies so far have focused on faces.

The presenter responded that the basic codec would be generic while an SEI message could be used for the additional purpose.

The possibility of “fusion” of generative content with coded pictures is also described in the related proposals. This, or more frequent key frames, was suggested as a way to help if the source content contains non-head-and-shoulders moving objects.

As an SEI message, it could potentially be used with different video coding technologies (e.g., AVC, HEVC, VVC, etc.).

There was a comment about studying the relationship with VVC, whether there would be a coupling of the core 2D video coding technology with the generative content.

There was a remark that a 1994 MPEG contribution (from Homer Chen) had reported on video generation using audio, which can be considered an early example of a proposal for a similar concept.

It was suggested that the target bit rate should not just consider extreme low bit rates and low resolutions.

It was commented that this is, at a basic level, a report of a coding efficiency benefit.

It was asked whether these reported experiments were coded as 4:4:4 RGB or 4:2:0 YUV. The proponent said currently the tests used 4:4:4. It was commented that it is hard to convince any mobile manufacturer to deploy 4:4:4 video codecs; that here is very little deployment of 4:4:4 video codec deployment because not only of coding performance but because also of complexity. Using 4:4:4 was thus suggested to be undesirable. It was replied that 4:4:4 is not a fundamental necessity and this could easily also be done with 4:2:0.

A comment was one example technology with a somewhat similar spirit is LCEVC approach in which the “base layer” can be any codec.

There was discussion of whether this should be considered layered coding, e.g., a type of temporal scalability.

The main application target is suggested to be videoconferencing. It was commented that this application is currently dominated by end-to-end non-interoperable products, saying that for apps which control both encoder and decoder, no interoperability may be needed, thus no standardization. A rationale for standard video coding is custom hardware, and it was asked whether a generic GPU could do this without custom hardware and whether interop standardization is helpful. Another participant commented that the custom hardware needed for this is not insignificant. One would need to implement on mobile devices hardware encoders and decoders for generic video codecs, and this would add considerable area on a mobile phone to support this extra functionality. It is hard to convince companies to add more area on their devices for a limited-use application. It was said that this technique could fail under certain scenarios when motion may exceed the limitations of the modelling scheme.

It was commented however, that inside of the non-interoperable apps are often standardized elements, and thereby taking advantage of the economies of scale of standardized technology.

From a requirements perspective, the use case is suggested to be videoconferencing. Study of market needs analysis was suggested before or while JVET studies the technical potential. It was noted that SEI messages (including relatively challenging ones) are often developed without formal requirements study.

VCEG and others indicated that it would OK for JVET to study this as limited scope/time subject for study, although not as major initiative. If a new normative standard would be needed (e.g. rather than an SEI message), requirements analysis would be needed. If further study determines that more than an SEI message is needed, more disciplined requirements analysis should be conducted to determine what should be done. This was agreed.

MPEG Requirements was also suggested to study the potential and use cases of generative video in general, in parallel with the technical studies in JVET. This was agreed.

Ended 12:08.

## Joint meeting of JVET, VCEG, and WG 4 on MIP/MIV related topics

(notes by JRO)

This joint session was held on Wednesday 18 Oct. at 1030–1130. About 80 people were present in the room and about XXX were connected on Zoom (substantially overlapping with those present in the room).

Documents considered:

[JVET-AF0167](https://jvet-experts.org/doc_end_user/current_document.php?id=13425) / m64998 / m65401 AHG9: On the Proposed Multiplane Image Information SEI Message [T. Lu, P. Yin, S. Oh, S. McCarthy, W. Husak, G. J. Sullivan (Dolby)]

[m65002](https://dms.mpeg.expert/doc_end_user/current_document.php?id=89623&id_meeting=196) [MIV] On MIV Extended Restricted Geometry Profile and MPI coding [T. Lu, S. Oh, P. Yin, S. McCarthy, W. Husak, G. J. Sullivan (Dolby)]

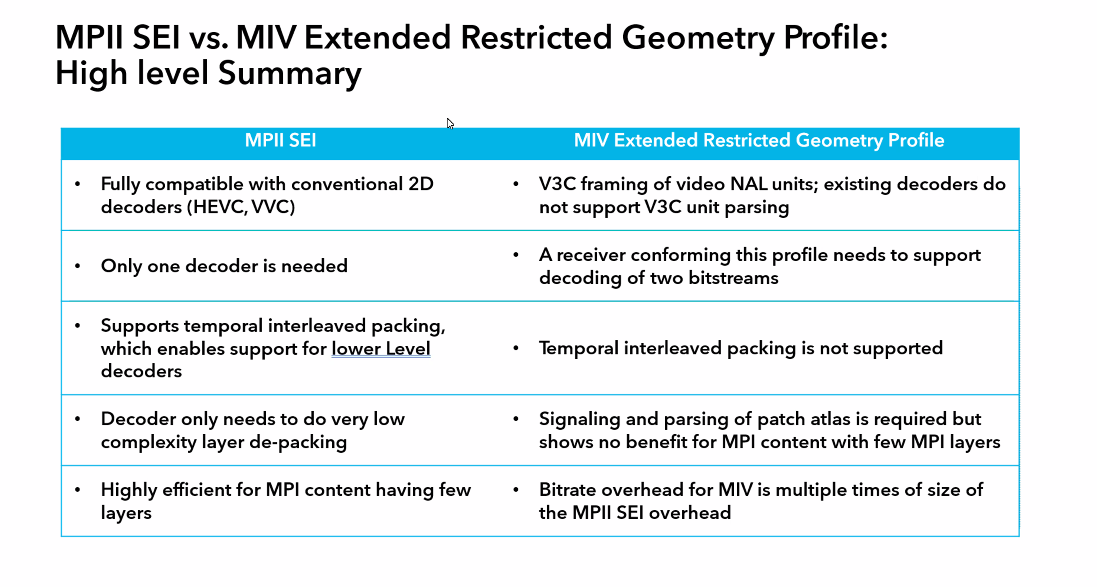
[m65158](https://dms.mpeg.expert/doc_end_user/current_document.php?id=89779&id_meeting=196) [MIV] How to achieve a low complexity MPI with the MIV standard [Bertrand Chupeau, Gaelle Martin-Cocher (InterDigital)]

The considered use case is the use of one video bitstream for one view.

It was commented by the JVET chair that a concept equivalent to MPI (called layered depth imaging by that time) had been proposed long ago in MPEG and JCT-3V but had not been pursued, as there had been more interest in pursuing autostereoscopic capabilities.

The following documents the progressing of discussions:

* Discuss the two possible options: SEI message, MIV profile
* The proponents of JVET-AF0167 presented the following comparison table of the two options:



* An analysis of the proposals had been performed in WG 4. It was found that the desirable functionality is already supported by an existing MIV proposal (except of temporal interleaving). This would require availability of two VVC decoders (or one decoder with higher level).
* It was commented by proponents of the SEI message that the time to market would be shorter with the SEI approach.
* Several experts commented that it would not require new hardware. Implementation of rendering would be necessary, but that would be the case with both approaches.
* It was commented from a service provider that a minimum set of functionality and implementation effort might be desirable
* It was commented that having different ways of implementing the same would be confusing for the market
* The JVET chair commented that implementing a subset of an MIV profile might not be a desirable solution, as this would not be conformant with that profile. The clean solution would be defining a new profile (which could also include temporal interleaving). For an SEI message, there is no conformance, but that is clear from the concept of SEI. It was agreed that with the MIV approach, the only reasonable option would be defining a new profile. This would have the advantage over the SEI message that a decoder is conformant in reconstructing the final output.
* It was commented by the WG 4 convenor that the quality from the demo of the simple solution was not convincing.
* One expert commented that using auxiliary pictures rather than TI would be a more clean solution (which however would not be possible with main profile)
* It was commented that the bar of defining a new profile is usually higher than for an SEI message (market needs analysis, quality analysis, etc.)
* It was commented by one expert that it would be damaging for MIV if JVET would define an SEI message.
* It was commented by another expert that it would be damaging to MPEG and VCEG if a proprietary format would be defined outside of standardization.
* It was commented that there are plenty of cases where the same functionality can be achieved in different ways (e.g. frame packing vs. multiview coding). Ultimately, the market decides in such cases what is taken up.
* Though there was still objection raised by several participants against defining an SEI message in JVET, and arguments were iterating, it was finally concluded that JVET should investigate possible development and practicality of an SEI message for MPI.

## Joint session 11:45am Wednesday 18 Oct on Multi-layer coding tests AG 5 Visual Quality, WG 5 / JVET, VCEG

(Notes by GJS)

Convenors / rapporteurs: J.-R. Ohm (WG 5 / JVET), G. Sullivan (VCEG), M. Wien (AG 5), Y. Ye (VCEG); also J. Ostermann of AG 2 and L. Yu of WG 4.

[JVET-AF0075](https://jvet-experts.org/doc_end_user/current_document.php?id=13323) to JVET, [m65560](https://dms.mpeg.expert/doc_end_user/current_document.php?id=90181&id_meeting=196) to WG 4 (cross-checked in [JVET-AF0224](https://jvet-experts.org/doc_end_user/current_document.php?id=13487), JVET-AF0228, [JVET-AF0277](https://jvet-experts.org/doc_end_user/current_document.php?id=13541))

This contribution presents evaluation results of Low Complexity Enhancement Video Codec (LCEVC) using its reference implementation (LTM-5.4.1) when combined with HEVC (HM-16.25) and VVC (VTM-19.0) as base layer codecs. Tests were performed for 4K content (Class A1 and A2 test sequences), following JVET Random Access (RA) Common Testing Conditions (CTC). The anchor is the result of single layer coding, i.e., encoding 4K videos by VTM or HM only, and the test is the result of base layer codec, which encodes quarter-resolution version of 4K video by using VTM or HM as the base layer, followed by enhancement layer coded using LTM-5.4.1, where the enhancement layer is obtained as a difference between original 4K video and upsampled reconstructed frames from the base layer. PSNR BD-rate was used for objective evaluation.

On top of VTM-19.0, the simulation results of LCEVC, VTM-19.0 + LTM-5.4.1,were reported as:

* **RA CTC**: {Y BD-rate = +31.59%, Cb BD-rate = +99.55%, Cr BD-rate = +70.60%}

On top of HM-16.25, the simulation results of LCEVC, HM-16.25 + LTM-5.4.1, were reported as:

* **RA CTC**: {Y BD-rate = +7.89%, Cb BD-rate = +78.01%, Cr BD-rate = +51.42%}

Visual artefacts were also reported and illustrated in the contribution using still-frame captures from the sequences. Subjective viewing of the results reportedly showed some loss of details in some LCEVC encoded areas.

[m65556](https://dms.mpeg.expert/doc_end_user/current_document.php?id=90177&id_meeting=196) (partly in [JVET-AF0292](https://jvet-experts.org/doc_end_user/current_document.php?id=13556), which was to be revised with the additional material from m65556) Considerations on LCEVC performances in relation to input document JVET-AF0075/m65560 [Simone Ferrara, Lorenzo Ciccarelli, Guendalina Cobianchi, Guido Meardi]

This document comments on the methodology and the conclusions of input document JVET-AF0075 and the expected performance of MPEG-5 Low Complexity Enhancement Video Coding (LCEVC) on objective metrics vs. formal MOS. Prior documents (about 30) were highlighted including LCEVC’s verification tests (m54455) and several evaluations of LCEVC using the same sequences (such as WG 11 N 19160, m53796, m52997 and m53523). A summary of LCEVC verification tests (VTs) can be found in WG 11 N19571. A relatively low correlation of PSNR with subjective quality was noted. The selection of bit rates for testing was also highlighted. The CTC for LCEVC, including configurations of LTM and anchors, was focused around a different range of bit rates, typically below 15 Mbps for UHD sequences – with the HM anchors ranging from QP 26–27 to QP 39–40, depending on the sequence – see also WG 11 N 18988. For the verification tests the QP values chosen for the anchors ranged between QP 27 and QP 42. Some results were shown with comparison to the x265 encoder for HEVC and the MainConcept encoder for VVC. Various metrics are discussed. It was emphasized that LCEVC has low complexity as a significant part of its design intent.

It was emphasized that PSNR has a relatively low correlation with subjective quality; subjective quality is what really matters. This contribution discussed various other objective metrics as well as PSNR, and in some cases those other metrics tend to have better correlation with subjective quality.

Chroma edges are less impactful on visual quality but the chroma fidelity balance with luma quality can be changed.

It was commented that the bit rate allocation between the base layer and enhancement layer, including temporal effects, is very important and has often not been optimized when performing comparisons. This includes per-frame, per-layer, bit rate allocation issues.

It was asked what kind of quantization was performed for the LCEVC VTs. The default quantization matrices (QMs) were reported to have been used for LCEVC. These are not flat, but are fixed-value matrices. It was said that flat matrices may have been used for the HEVC and VVC references. It was emphasized that although LCEVC used a non-flat matrix, it was a fixed-value matrix defined as the default rather than something customized by the encoder.

It was asked if dithering filtering was used in the verification test, and reported that it was not used.

There was a comment about x265 in relation to the HEVC HM encoder, saying x265 had significantly lower compression performance. It was commented that the x265 encoder configuration settings that were used in the tests reported in this contribution results in a large loss compared to the HM (in fact more than 80% loss in luma, more than 100% loss in chroma) and is worse than the AVC JM, saying having gains on top of a weak encoder may not compensate for the loss to the results obtained with a stronger single-layer encoder.

Another participant responded that what should be analysed is the delta between the layered and non-layered coding using the same base encoder rather than the relative performance with different base encoders. However, it was remarked that the delta does not seem to be invariant to the question of which base layer encoder is used.

It was commented that enforcing the use of three colour planes in LTM5.4.1 results in worse results compared to the tests reported in JVET-AF0075. Using a newer LTM version has different results but still was reported to have a loss.

It was remarked that carefully considering and studying the quality of an upscaled base layer encoding is very important. At low bit rates, an upscaled lower-resolution encoding may actually look better than (or as good as) a dual-layer or single-layer encoding of higher resolution video. It was reported that such an upscaled reference had been included in the LCEVC verification tests.

It was commented that VMAF is a trained metric; only luma was used for that training, and as a result, VMAF only considers the luma distortion.

It was asked whether the settings used for the LCEVC LTM5.4.1 encoder were used according to the settings reported in the verification testing, and the contributors of JVET-AF0075 confirmed that the settings were the same as reported in the WG 4 verification test report on the compression performance of LCEVC (WG 4 N 0076). This is also confirmed for the other contribution.

It was indicated that in JVET-AF0075 the quantization parameter (QP) values configured in its test followed the formula provided in the verification test report.

It was reported that the bitstreams from JVET-AF0075 were available and could be viewed upon request.

It was indicated that there are two enhancement layers in LCEVC, one of which is an upscaling layer.

It was commented that some parts of the video have better visual quality in the LCEVC layered coding approach than in the single-layer encodings, pointing to examples provided in the m65556 input contribution.

[JVET-AF0186](https://jvet-experts.org/doc_end_user/current_document.php?id=13444) AHG4: experiments related to VVC spatial scalability visual testing [P. de Lagrange, F. Urban (InterDigital)]

This document reports about some experiments related to visual testing of VVC spatial scalability, as planned in JVET-AE2021, and after observations from expert viewing tests made during the previous JVET meeting, where the upscaled anchor was often found better than the dual-layer test. After investigation, it was found that moving the inter-layer reference to be the first picture in the reference picture lists (instead of the last picture) improved the situation. In addition, video sequences from LCEVC verification tests have been tested.

[JVET-AF0311](https://jvet-experts.org/doc_end_user/current_document.php?id=13575) Results of expert viewing at AF meeting for the spatial scalability category of the VVC multilayer VT M. Wien (RWTH Aachen University)

Some results of experts viewing experiments conducted during the meeting were shown.

Using the referencing settings proposed in JVET-AF0186 were reported to help.

At the low rates, an upsampled lower resolution video seemed better. [add notes from contrib]

There is some more coherence in the results than had been obtained previously, but it is not yet clear how to resolve the issues.

It was commented that the encoder still seems not mature and needs further work.

Generally it seems needed to conduct further study of the issues raised in these VVC multilayer comparisons reported in JVET-AF0311.

End 13:00

## BoGs (1)

The following break-out groups were established at this meeting to conduct discussion and develop recommendations on particular subjects.

[JVET-AF0307](https://jvet-experts.org/doc_end_user/current_document.php?id=13571) BoG on assessment of HOP proposals [E. Alshina, F. Galpin]

Was presented in JVET Wed. 18 Oct. 1305 (chaired by JRO)

Following categories of proposals have been discussed:

1. Filter usage aspects (section 1);
2. Training aspects (section 2);
3. Design elements of HOP for future NNVC (section 3).

***Recommendations***

* Adoption to NNVC-7.0 (classify SW/CTC/train)
  + Decision: Adopt [JVET-AF0085](about:blank) (EE1-1.2.1), enable by default for LOP and HOP;
  + Decision: Adopt [JVET-AF0172](about:blank) (redundant signalling removal);
  + Decision: Adopt JVET-AF0193 (disable by default);
  + Decision: Adopt [JVET-AF0180](about:blank) and [JVET-AF0155](about:blank) (12:1:1 distortion weight in HOP training)
  + Decision: Adopt [JVET-AF0155](about:blank) (Chroma QP offset +1 for HOP inference)
* Next EE1:
  + Agreement: stay with single models for HOP
  + EE1-0: HOP re-train with 12:1:1 loss (in inference Chroma QP offset +1 and [JVET-AF0085](about:blank));
    - [JVET-AF0150](about:blank) and [JVET-AF0296](about:blank) (batch size 32 for HOP training and new learning rate)
    - Training must be completed by Nov. 17
  + EE1-1,2:
    - Training as EE1-0
      * NN design: NNVC-6 HOP with following changes:
        + EE1-1.1.2b (alternation of 3x1 and 1x3 CONV, C31=48)
        + EE1-1.1.3 (1x1 CONV in Head Block for IPB and BS)
        + One of the 2 variants:

(1) Group convolution in BBB, location and size as in EE1-1.1.4b

(2) Group convolution in BBB, location size as in EE1-1.1.1-t1

* + - Complexity:
      * kMAC/pxl <=477
      * adjustment of number of BBBs, if mode parameters modified must be same form in (1) and (2)
    - Final training starts after Nov. 20
  + EE1-3:
    - Only inference test for HOP-6 (single model) and EE1-1.1.4c (two models), no retraining foreseen, memory usage, accurate time measurement w/o AVX512, same inference/usage
  + EE1-4: Study joint design of [JVET-AF0154](about:blank) (rotation) and [JVET-AF0086](about:blank) (flipping)
  + EE1-5: Deep Reference Frame Generation for Inter Prediction Enhancement, training is crosschecked.
  + + EE tests discussed during JVET plenary

Yellow highlighted bullet point to be updated from new version.

Additional Decision: Adopt JVET-AF0205 (was decided Thu 19 Oct. 1630, after confirmation of successful completion of the crosscheck).

All recommendations of the BoG were confirmed by JVET plenary.

## Liaison communications (2)

JPEG liaison [m65594](https://dms.mpeg.expert/doc_end_user/current_document.php?id=90215&id_meeting=196) provided …. (Elena)

A reply was drafted as WG 5 N XXX.

3GPP liaison [m65645](https://dms.mpeg.expert/doc_end_user/current_document.php?id=90266&id_meeting=196) provided …. (Walt)

A reply was drafted as WG 5 N XXX.

Liaison responses were reviewed in JVET on Thursday 19 Oct. 0830-1000. The draft replies were also presented in the AG 3 meeting Thursday 19 Oct. at 1000.

# Project planning

## Software timeline

ECM 11 software (including all adoptions) was planned to be available 3 weeks after the meeting.

The NNVC 7.0 codebase software was planned to be available 2 weeks after the meeting. An update 7.1 (including potentially the reduced-complexity intra prediction, and an update ofHOP) was planned to be available 5 weeks after the meeting.

VTM23.0 software was planned to be available on 2023-11-24. (Note that further updates may be released later)

Updates on top of HM17.0 software were not planned, but might be released after merging pending requests, as appropriate.

As a general rule in software development, a person who is executing a merge shall not be from the same company as the person who submitted that merge request.

## Core experiment and exploration experiment planning

An EE on neural network-based video coding was established, as recorded in output document JVET-AF2023.

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-AF2024.

Initial versions of these documents were presented and approved.

## Drafting of specification text, encoder algorithm descriptions, and software

The following agreement has been established: the editorial team has the discretion to not integrate recorded adoptions for which the available text is grossly inadequate (and cannot be fixed with a reasonable degree of effort), if such a situation hypothetically arises. In such an event, the text would record the intent expressed by the committee without including a full integration of the available inadequate text.

## Plans for improved efficiency and contribution consideration

The group considered it important to have the full design of proposals documented to enable proper study.

Adoptions need to be based on properly drafted working draft text (on normative elements) and HM/VTM encoder algorithm descriptions – relative to the existing drafts. Proposal contributions should also provide a software implementation (or at least such software should be made available for study and testing by other participants at the meeting, and software must be made available to cross-checkers in EEs).

Suggestions for future meetings included the following generally-supported principles:

* Normative contributions (relating to changes in bitstream/decoder) shall include draft specification text
* Proposals shall contain all details relevant for understanding and be self-contained. In cases where the document is a follow-up of a previous contribution, the overall concept and the novelties should be highlighted at minimum
* Coding tool and encoder optimization proposals shall contain Excel sheets that allow assessment on a per-sequence basis
* Algorithm description text is strongly encouraged for non-normative contributions that are intended to be included in model description documents (VTM, ECM, etc.), and that is required for inclusion in TR drafts.
* Early upload deadline to enable substantial study prior to the meeting
* Using a clock timer to ensure efficient proposal presentations (5 min) and discussions (nott exercised currently)

As general guidance, it was suggested to avoid usage of company names in document titles, software modules etc., and not to describe a technology by using a company name.

## General issues for experiments

It was emphasized that those rules which had been set up or refined during the 12th JVET meeting should be observed. In particular, for some CEs of some previous meetings, results were available late, and some changes in the experimental setup had not been sufficiently discussed on the JVET reflector.

Group coordinated experiments have been planned as follows:

* “Core experiments” (CEs) are the coordinated experiments on coding tools which are deemed to be interesting but require more investigation and could potentially become part of a draft standard by the next meeting or in the near future.
* “Exploration experiments” (EEs) are also coordinated experiments. These are conducted on technology which is not foreseen to become part of a draft standard in the near future. The investigating methodology for assessment of such technology can also be an important part of an EE. (Further general rules for EEs, as far as deviating from the CE rules below, should be discussed in a future meeting. For the current meeting, procedures as described in the EE description document are deemed to be sufficient.)
* A CE is a test of a specific fully described technology in a specific agreed way. It is not a forum for thinking of new ideas (like an AHG). The CE coordinators are responsible for making sure that the CE description is complete and correct and has adequate detail. Reflector discussions about CE description clarity and other aspects of CE plans are encouraged.
* A description of each experiment is to be approved at the meeting at which the experiment plan is established. This should include the issues that were raised by other experts when the tool was presented, e.g., interference with other tools, contribution of different elements that are part of a package, etc. The experiment description document should provide the names of individual people, not just company names.
* Software for tools investigated in a CE will be provided in one or more separate branches of the software repository. Each CE will have a “fork” of the software, and within the CE there may be multiple branches established by the CE coordinator. The software coordinator will help coordinate the creation of these forks and branches and their naming. All JVET members will have read access to the CE software branches (using shared read-only credentials as described below).
* During the experiment, revisions of the experiment plans can be made, but not substantial changes to the proposed technology. Withdrawing parts of experiments that were intended to show the individual benefits of a tool or parts of a tool is strongly discouraged. Combination tests may not be considered in such cases. Any changes made to individual tools in a combination shall be documented.
* The CE description must match the CE testing that is done. The CE description needs to be revised if there has been some change of plans.
* The CE summary report must describe any changes that were made in the process of finalizing the CE.
* By the next meeting it is expected that at least one independent cross-checker will report a detailed analysis of each proposed feature that has been tested and confirm that the implementation is correct. Commentary on the potential benefits and disadvantages of the proposed technology in cross-checking reports is highly encouraged. Having multiple cross-checking reports is also highly encouraged (especially if the cross-checking involves more than confirmation of correct test results). The reports of cross-checking activities may (and generally should) be integrated into the CE report rather than submitted as separate documents.
* It is mandatory to report encoder optimizations made for the benefit of a tool, and if an equivalent optimization could be applied on the anchor, a comparison against the improved anchor shall be provided.
* A new proposal can be included in a CE based on group decision, regardless if an independent party has already performed a cross-check in the meeting when it was first proposed.

It is possible to define sub-experiments within particular CEs, for example designated as CEX.a, CEX.b, etc., where X is the basic CE number.

As a general rule, it was agreed that each CE should be run under the same testing conditions using one software codebase, which should be based on the group test model software codebase. An experiment is not to be established as a CE unless there is access given to the participants in (any part of) the CE to the software used to perform the experiments.

The general agreed common conditions for single-layer coding efficiency experiments for SDR video are described in the prior output document JVET-T2010.

Experiment descriptions should be written in a way such that it is understood as a JVET output document (written from an objective “third party perspective”, not a proponent perspective – e.g. not referring to methods as “improved”, “optimized”, “enhanced”, etc.). The experiment descriptions should generally not express opinions or suggest conclusions – rather, they should just describe what technology will be tested, how it will be tested, who will participate, etc. Responsibilities for contributions to CE work should identify individuals in addition to company names.

CE descriptions contain a basic description of the technology under test, but should not contain excessively verbose descriptions of a technology (at least not unless the technology is not adequately documented elsewhere). Instead, the CE descriptions should refer to the relevant proposal contributions for any necessary further detail. However, the complete detail of what technology will be tested must be available – either in the CE description itself or in documents that are referenced in the CE description that are also available in the JVET document archive.

Any technology must have at least one cross-check partner to establish a CE – a single proponent is not enough. It is highly desirable have more than just one proponent and one cross-checker.

The CE development workflow is described at:

<https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/wikis/Core-experiment-development-workflow>

CE read access is available using shared accounts: One account exists for MPEG members, which uses the usual MPEG account data. A second account exists for VCEG members with account information available in the TIES informal ftp area (IFA) system at:

<https://www.itu.int/ifa/t/2017/sg16/exchange/wp3/q06/vceg_account.txt>

Some agreements relating to CE activities were established as follows:

* Only qualified JVET members can participate in a CE.
* Participation in a CE is possible without a commitment of submitting an input document to the next meeting. Participation 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 Friday 20 Oct. 2023 at 0935.

|  |  |  |
| --- | --- | --- |
| **Title and Email Reflector** | **Chairs** | **Mtg** |
| **Project Management (AHG1)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Coordinate overall JVET interim efforts. * Supervise AHG and experiment studies. * Report on project status to JVET reflector. * Provide a report to the next meeting on project coordination status. * Supervise processing and delivery of output documents | J.-R. Ohm (chair), G. J. Sullivan (vice‑chair) | N |
| **Draft text and test model algorithm description editing (AHG2)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Produce and finalize draft text outputs of the meeting (JVET-AF1006, JVET-AF1016, and JVET-AF2027. * Collect reports of errata for the VVC, VSEI, HEVC, AVC, CICP, and the published related technical reports and produce the JVET-AF1004 errata output collection. * Coordinate with the test model software development AhG to address issues relating to mismatches between software and text. * Collect and consider errata reports on the texts. | B. Bross, C. Rosewarne (co-chairs), F. Bossen, A. Browne, S. Kim, S. Liu, J.‑R. Ohm, G. J. Sullivan, A. Tourapis, Y.-K. Wang, Y. Ye (vice‑chairs) | N |
| **Test model software development (AHG3)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Coordinate development of test models (VTM, HM, SCM, SHM, HTM, MFC, MFCD, JM, JSVM, JMVM, 3DV-ATM, 360Lib, and HDRTools) software and associated configuration files. * Produce documentation of software usage for distribution with the software. * Enable software support for recently standardized additional SEI messages, and SEI messages in TuC (the latter in a separate branch of VTM). * Discuss and make recommendations on the software development process. * Perform comparative tests of test model behaviour using common test conditions, including HDR, high bit depth and high bit rate. * Suggest configuration files for additional testing of tools. * Investigate how to minimize the number of separate codebases maintained for group reference software. * Coordinate with AHG on Draft text and test model algorithm description editing (AHG2) to identify any mismatches between software and text, and make further updates and cleanups to the software as appropriate. * Prepare drafts of merged and updated CTC documents for HM and VTM, as applicable. | F. Bossen, X. Li, K. Sühring (co-chairs), E. François, Y. He, K. Sharman, V. Seregin, A. Tourapis (vice‑chairs) | N |
| **Test material and visual assessment (AHG4)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Consider plans for additional verification testing of VVC capability, particularly target conducting a first test for VVC multi-layer features by the next meeting, and update the test plan according to subsequent tests. * Coordinate with AHG13 on assessing new test material and investigating metrics that could be used to assess quality of synthesized film grain; improve and update the draft test plan for subjective quality testing of the FGC SEI message. * Maintain the video sequence test material database for testing the VVC and HEVC standards and potential future extensions, as well as exploration activities. * Study coding performance and characteristics of available and proposed video test material. * Identify and recommend appropriate test material for testing the VVC standard and potential future extensions, as well as exploration activities. * Identify and characterize missing types of video material, solicit contributions, collect, and make available a variety of video sequence test material, in coordination with other AHGs, as appropriate. * Maintain and update the directory structure for the test sequence repository, as necessary. * Collect information about test sequences that have been made available by other organizations. * Prepare and conduct expert viewing for purposes of subjective quality evaluation. * Coordinate with AG 5 in studying and developing further methods of subjective quality evaluation, e.g. based on crowd sourcing. * Coordinate with AHG15 on investigating sequences with gaming content, and make such sequences available for study. * Prepare availability of viewing equipment and facilities arrangements for future meetings. | V. Baroncini, T. Suzuki, M. Wien (co-chairs), W. Husak, S. Iwamura, P. de Lagrange, S. Liu, X. Meng, S. Puri, A. Segall, S. Wenger (vice-chairs) | Y (tel., 2 weeks notice) |
| **Conformance testing (AHG5)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the draft of additional conformance bitstreams for VVC multilayer configurations JVET-AE2028, and investigate the need for future improvements of conformance testing specifications. * Study the conformance needs for HEVC multi-view profiles, and develop a set of conformance bitstreams as appropriate. * Study the requirements of VVC, HEVC, and AVC conformance testing to ensure interoperability. * Maintain and update the conformance bitstream database, and contribute to report problems, and suggest actions to resolve these. * Study additional testing methodologies to fulfil the needs for VVC conformance testing. | I. Moccagatta (chair), F. Bossen, K. Kawamura, P. de Lagrange, T. Ikai, S. Iwamura, H.-J. Jhu, S. Paluri, K. Sühring, Y. Yu (vice‑chairs) | N |
| **ECM software development (AHG6)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Coordinate development of the ECM software and associated configuration files. * Produce documentation of software usage for distribution with the software. * Prepare and deliver ECM-11.0 software version and the reference configuration encodings according to the ECM common test conditions. * Investigate encoder speedup and other software optimization such as reduction of memory consumption. * Coordinate with ECM algorithm description editors to identify any mismatches between software and text, make further updates and cleanups to the software as appropriate. | V. Seregin (chair), J. Chen, R. Chernyak, F. Le Léannec, K. Zhang (vice-chairs) | N |
| **ECM tool assessment (AHG7)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Investigate methodology of tool assessment. * Coordinate with AHG6 on resolving tool-off test related software issues (missing tool controls and software bugs). * Prepare configuration files and generate bitstreams and results of tool-on/tool-off testing. * Prepare reporting of tool assessment results. * Collect simulation results on non-CTC sequences (e.g., those used in previous verification tests), and identify a set of non-CTC sequences that would be appropriate for additional testing. * Investigate the possibility of conducting subjective tests on subsets of tools in coordination with AHG4 and AG 5. * Develop methodology of more reliable runtime measurement | X. Li (chair), L.-F. Chen, Z. Deng, J. Gan, E. François, H.-J. Jhu, X. Li, H. Wang (vice‑chairs) | N |
| **Optimization of encoders and receiving systems for machine analysis of coded video content (AHG8)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Solicit and study non-normative encoder and receiving systems technologies that enhance performance of machine analysis tasks on coded video content. * Identify and collect test materials that are suitable to be used by JVET for machine analysis tasks. * Generate anchors according to the common test conditions JVET-AF2031. * Discuss improvements on the evaluation framework, including evaluation procedures and methodologies. * Coordinate software development, and continue to migrate the software basis used in AHG8 to newest VTM version. * Coordinate experiments on optimization of encoders and receiving systems for machine analysis of coded video content. * Maintain the software implementation example algorithms in the repository, including sufficient documentation in terms of operation and performance. * Evaluate proposed technologies and their suitability for machine analysis applications. * Propose improvements to the draft technical report JVET-AE2030 on optimization of encoders and receiving systems for machine analysis of coded video content. * Study the potentials of using SEI messages for the purpose of machine analysis in coordination with AHG9. * Coordinate with WG 4 VCM AHG on aspects such as common test conditions, evaluation metrics, test and training materials, usage of SEI messages, and on studying characteristics and requirements of targeted machine analysis tasks, etc. | C. Hollmann, S. Liu, S. Wang, M. Zhou (AHG chairs) | N |
| **SEI message studies (AHG9)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the SEI messages in VSEI, VVC, HEVC and AVC. * Discuss the document for technologies under consideration for VSEI JVET-AF2032, and propose improvements as appropriate. * Collect software and showcase information for SEI messages, including encoder and decoder implementations and bitstreams for demonstration and testing. * Identify potential needs for additional SEI messages, including the study of SEI messages defined in HEVC and AVC for potential use in the VVC context. * Study the alignments of the same SEI messages in different standards * Coordinate with AHG8 and WG 4 to study mechanisms for signalling metadata in the context of machine analysis of coded video content. * Coordinate with AHG3 for software support of SEI messages. | S. McCarthy, Y.-K. Wang (co-chairs), T. Chujoh, S. Deshpande, C. Fogg, M. M. Hannuksela, Hendry, P. de Lagrange, G. J. Sullivan, A. Tourapis, S. Wenger (vice-chairs) | N |
| **Encoding algorithm optimization (AHG10)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the impact of using techniques such as tool adaptation and configuration, and perceptually optimized adaptive quantization for encoder optimization. * Study the impact of non-normative techniques of preprocessing for the benefit of encoder optimization. * Study encoding techniques of optimization for objective quality metrics and their relationship to subjective quality. * Study optimized encoding for reference picture resampling and scalability modes in VTM. * Study optimized encoding and tool combinations for low latency and low complexity. * Consider neural network-based encoding optimization technologies for video coding standards. * Investigate other methods of improving objective and/or subjective quality, including adaptive coding structures and multi-pass encoding. * Study methods of rate control and rate-distortion optimization and their impact on performance, subjective and objective quality. * Study the potential of defining default or alternate software configuration settings and test conditions optimized for either subjective quality, or higher objective quality, and coordinate such efforts with AHG3 and AHG6. * Study the effect of varying configuration parameters depending on temporal layer, such as those related to deblocking, partitioning, chroma QP. | P. de Lagrange, A. Duenas, R. Sjöberg, A. Tourapis (AHG chairs) | N |
| **Neural network-based video coding (AHG11)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Evaluate and quantify the performance improvement potential of NN-based video coding technologies compared to existing video coding standards such as VVC, including both individual coding tools, architectures and content adaptation with NN parameters overfitting. * Discuss potential refinements of the test conditions for NN-based video coding in JVET-AF2016. Generate and distribute anchor encoding, and develop supporting software as needed. * Study the impact of training (including the impact of loss functions) on the performance of candidate technologies and identify suitable material for testing and training. * Analyse complexity characteristics for technologies under study, including transformers, perform complexity analysis, and develop complexity reductions of candidate technology. * Finalize and discuss the EE on neural network-based video coding. * Coordinate with other groups, including SC29/AG5 on the evaluation and assessment of visual quality, and AHG12 on the interaction with ECM coding tools. If possible, prepare encodings with combinations of tools included in the NNVC software for visual quality assessment at the next meeting. * Coordinate with AHG14 on items related to NNVC software development. | E. Alshina, F. Galpin, S. Liu, A. Segall (co-chairs), J. Li, R.-L. Liao, D. Rusanovskyy, M. Santamaria, T. Shao, M. Wien, P. Wu (vice chairs) | Y (tel., 2 weeks notice), first on Nov. 16 |
| **Enhanced compression beyond VVC capability (AHG12)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Solicit and study non-neural-network video coding tools with enhanced compression capabilities beyond VVC. * Discuss and propose refinements to the ECM11 algorithm description JVET-AF2025. * Coordinate with AHG7 to study the performance and complexity tradeoff of these video coding tools. * Coordinate with AHG6 on ECM software development. * Support AHG6 in generating anchors according to the test conditions in JVET-AF2017. * Analyse the results of exploration experiments described in JVET-AF2024 in coordination with the EE coordinators. * Coordinate with AHG11 to study the interaction with neural network-based coding tools. | M. Karczewicz, Y. Ye, L. Zhang (co-chairs), B. Bross, R. Chernyak, X. Li, K. Naser, Y. Yu (vice-chairs) | N |
| **Film grain technologies (AHG13)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the benefits and characteristics of film grain technologies, including autoregressive and frequency-filtering technologies. * Study alternative film grain models and their associated documentation. * In consultation with AHG4, study and define content characteristics and test conditions that are desirable for the study and testing of film grain technologies, and perform an assessment of newly available test materials in that regard. * Investigate metrics for measuring film grain fidelity in itself, or as present in a video. * Discuss the potential need for film grain conformance guidelines. * Given the study of desirable content characteristics, solicit or create new test material for further determining the operational characteristics of, testing, and developing any related technologies. * Study preprocessing and encoder technologies for determining values for FGC (Film Grain Characteristics) SEI message syntax elements. * Identify potential need for additional film grain technology and signalling, if needed. * Coordinate development of film grain technology software and configuration files. * Coordinate with AG 5 on improving the draft plan for subjective quality testing of the FGC SEI message JVET-AD2022, and conduct preparations for such testing. * Coordinate with AHG3 for software support of the FGC SEI message. | W. Husak, P. de Lagrange (co-chairs), A. Duenas, D. Grois, Y. He, X. Meng, M. Radosavljević, A. Segall, G. Teniou, A. Tourapis (vice-chairs) | Y (tel., 2 weeks notice) |
| **NNVC software development (AHG14)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Coordinate development of the NNVC software and associated configuration files. * Prepare and deliver NNVC-7.0 and NNVC-7.1 software versions and the reference configuration encodings according to the NNVC common test conditions as described in JVET-AF2016. * Investigate combinations of tools included in the NNVC software, prepare and release anchor data for all configurations of the software, including anchors for High and Low Operation Point (HOP/LOP) configurations. * Study and maintain the SADL (Small Adhoc Deep-Learning Library). Identify gaps in functionality and develop improvements as needed. * Coordinate with NNVC algorithm and software description (JVET-AF2019) editors to identify any mismatches between software and description document, suggest further updates to the description document as appropriate. * Coordinate with AHG11 on items related to NNVC activities. | F. Galpin (chair), Y. Li, Y. Li, J. Shingala, L. Wang, Z. Xie (vice chairs) | Y (tel., 2 weeks notice), first on Nov. 16 |
| **Gaming content compression (AHG15)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Identify gaming content application scenarios and their requirements for codec operation. * Identify and characterize required types of content; solicit contributions, collect, and make a variety of gaming content available, in coordination with AHG4 and AG 5. * Propose test conditions appropriate for gaming applications. * Evaluate JVET test models (such as ECM, VTM, NNVC, etc.) under the proposed test conditions. * Investigate possibilities to enhance compression capability for gaming content. | S. Puri, J. Sauer (co-chairs), R. Chernyak, A. Duenas, L. Wang (vice chairs) | Y (tel., 2 weeks notice) |
| **Generative face video compression (AHG16)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Establish testing conditions for evaluating the compression performance of generative face video compression (GFVC). * Identify and develop software tools for experimentation on GFVC, and make a software package available. * Study interoperability requirements, including study of the compression performance impact due to GFVC parameter translation. * Study compression performance using the VVC Main 10 profile. * Develop a document summarizing GFVC technologies. | Y. Ye (chair), H.-B. Teo, Z. Lyu, S. McCarthy, S. Wang (vice chairs) | Y (tel., 2 weeks notice) |

It was confirmed that the rules which can be found in document ISO/IEC JTC 1/‌SC 29/‌AG 2 [N 046](https://www.mpegstandards.org/wp-content/uploads/2022/01/ISO-IECJTC1-SC29-AG2_N0046_AhG.pdf) “Ad hoc group rules for MPEG AGs and WGs” (available at <https://www.mpegstandards.org/adhoc/>), are consistent with the operation mode of JVET AHGs. It is pointed out that JVET does not maintain separate AHG reflectors, such that any JVET member is implicitly a member of any AHG. This shall be mentioned in the related WG Recommendations. The list above was also issued as a separate WG 5 document (ISO/IEC JTC 1/‌SC 29/‌WG 5 N 253) 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 253, as noted in section 9.

[JVET-AF1000](https://jvet-experts.org/doc_end_user/current_document.php?id=13582) Meeting Report of the 32nd JVET Meeting [J.-R. Ohm] [WG 5 N 239] (2023-08-16)

Initial versions of the meeting notes (d0 … d8) were made available on a daily basis during the meeting.

Remains valid – not updated: [JVET-AC1001](https://jvet-experts.org/doc_end_user/current_document.php?id=12566) Guidelines for HM-based software development [K. Sühring, F. Bossen, X. Li (software coordinators)]

Remains valid – not updated: [JVET-Y1002](https://jvet-experts.org/doc_end_user/current_document.php?id=11463) High Efficiency Video Coding (HEVC) Test Model 16 (HM 16) Encoder Description Update 16 [C. Rosewarne (primary editor), K. Sharman, R. Sjöberg, G. J. Sullivan (co-editors)] [WG 5 [N 103](https://dms.mpeg.expert/doc_end_user/current_document.php?id=82085&id_meeting=189)]

Remains valid – not updated: [JVET-AD1003](https://jvet-experts.org/doc_end_user/current_document.php?id=12970) Coding-independent code points for video signal type identification (Draft 2 of 3rd edition) [WG 5 preliminary FDIS N 206] [G. J. Sullivan, A. Tourapis] (2023-06-30)

The technical content was submitted for ITU consent (but will not be published until ST 2128 is available); ISO FDIS was to be delayed until it is available.

Post-meeting note: Expected *de facto* primary editor for ITU consent text: G. J. Sullivan.

[JVET-AF1004](https://jvet-experts.org/doc_end_user/current_document.php?id=13583) Errata report items for VVC, VSEI, HEVC, AVC, and Video CICP [Y.-K. Wang, B. Bross, I. Moccagatta, C. Rosewarne, G. J. Sullivan] (2024-01-10, near next meeting)

Post-meeting note: *De facto* primary editor: Y.-K. Wang.

Errata from JVET-AF0064, and one item on VSEI that was brought up by FNB in the context of ballot comments on AVC and HEVC (missing parenthesis).

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 (along with some errata corrections from JVET-AD1004) was included in a new edition of HEVC submitted for ITU consent (and had previously been included in the FDIS submitted as WG 5 N 179 issued from the January 2023 meeting).

Post-meeting note: *De facto* primary editor for ITU consent text: Y.-K. Wang.

(JVET-Z1005 can be removed after publication of the new edition of ISO/IEC 23008-2.)

[JVET-AF1006](https://jvet-experts.org/doc_end_user/current_document.php?id=13584) New profiles, colour descriptors, and SEI messages for HEVC (draft 2) [WG 5 DAM N 244] [Y.-K. Wang, B. Bross, T. Ikai, G. J. Sullivan, A. Tourapis] (2023-11-10)

Primary editor for this document and WG 5 N 244: Y.-K. Wang.

Text updates from JVET-AF0063 (except those relating to the MV profile structure), and some errata fixes.

A DoCR on ISO/IEC 23008-2/CDAM1 was issued as WG 5 N 243 (reviewed Thursday 19 Oct 1530).

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]

Remains valid – not updated: [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)

Remains valid – not updated: [JCTVC-AC1009](https://jvet-experts.org/doc_end_user/current_document.php?id=12569) Common test conditions for SHVC [K. Sühring]

Remains valid – not updated [JCTVC-O1010](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=8511) Guidelines for Conformance Testing Bitstream Preparation [T. Suzuki, W. Wan]

Remains valid – not updated: [JVET-AE1011](https://jvet-experts.org/doc_end_user/current_document.php?id=13267) HEVC multiview profiles supporting extended bit depth (draft 2) [S. Paluri, W. Husak, A. Tourapis] [2023-08-11]

From JVET-AE0296. The specification of these profiles was also included in [JVET-AE1006](https://jvet-experts.org/doc_end_user/current_document.php?id=13266) and WG 5 N 226. This document, basically duplicating part of JVET-AE1006, was not issued as a separate WG 5 N document. See JVET-AE1006 for editorship note.

Draft 1 had been issued as preliminary WD WG 5 N 143.

(Number 1011 can be re-used when JVET-AE1006 progresses.)

Remains valid – not updated: JVET-[AD1012](https://jvet-experts.org/doc_end_user/current_document.php?id=12973) Overview of IT systems used in JVET [J.-R. Ohm, I. Moccagatta, K. Sühring, M. Wien] (2023-05-19)

Remains valid – not updated: [JCT3V-G1003](http://phenix.int-evry.fr/jct3v/doc_end_user/current_document.php?id=1884) 3D-AVC Test Model 9 [ D. Rusanovskyy, F. C. Chen, L. Zhang, T. Suzuki] [WG 11 N 14239]

Remains valid – not updated: [JCT3V-K1003](http://phenix.int-evry.fr/jct3v/doc_end_user/current_document.php?id=2499) Test Model 11 of 3D-HEVC and MV-HEVC [Y. Chen, G. Tech, K. Wegner, S. Yea] [WG 11 N 15141]

Remains valid – not updated: [JVET-AE1013](https://jvet-experts.org/doc_end_user/current_document.php?id=13268) Common test conditions of 3DV experiments [K. Sühring, M. Wien] [2023-09-01]

New licensing available from JVET-AE0179. Other owners have not responded, therefore we can assume that they don’t have a problem that the sequences are used.

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.

[JVET-AF1016](https://jvet-experts.org/doc_end_user/current_document.php?id=13585) AVC with extensions and corrections (draft 2) [WG5 DIS of 11th ed. N 241] [B. Bross, T. Ikai, G. J. Sullivan, A. Tourapis, Y.-K. Wang] [2023-11-10]

Primary editor of this document and WG 5 N 241: B. Bross.

Target 15th edition of ITU-T H.264 in April 2024.

Changes from JVET-AF0045, and some errata item fixes.

A DoCR on ISO/IEC CD 14496-10 was issued as WG 5 N 240 (reviewed Thursday 19 Oct 1610).

No output: JVET-Axx1017 through JVET-Axx1099

Remains valid – not updated [JVET-AA1100](https://jvet-experts.org/doc_end_user/current_document.php?id=11944) Common Test Conditions for HM Video Coding Experiments [K. Sühring, K. Sharman]

This specifies only the CTC for non-4:2:0 colour formats. The corresponding document for VVC is JVET-T2013, with no unification yet.

**No output: JVET-Axx2001**

[JVET-AF2002](https://jvet-experts.org/doc_end_user/current_document.php?id=13586) Algorithm description for Versatile Video Coding and Test Model 21 (VTM 21) [A. Browne, Y. Ye, S. Kim] [WG 5 N 245] (2024-01-12, near next meeting)

JVET-AF0111, JVET-AF0122

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]

Remains valid – not updated: [JVET-AE2005](https://jvet-experts.org/doc_end_user/current_document.php?id=13270) New level and systems-related supplemental enhancement information for VVC (Draft 6) [B. Bross, E. François, M. M. Hannuksela, A. Tourapis, Y.-K. Wang] (2023-08-18)

Primary editor for this document, the corresponding ITU consent text, and the corresponding FDIS text WG 5 N 228: B. Bross.

Remains valid – not updated: [JVET-AE2006](https://jvet-experts.org/doc_end_user/current_document.php?id=13271) Additional SEI messages for VSEI (Draft 5) [S. McCarthy, T. Chujoh, M. M. Hannuksela, G. J. Sullivan, Y.-K. Wang] (2023-08-18)

Primary editor for this document, the corresponding ITU consent text, and the corresponding FDIS text WG 5 N 220: Y.-K. Wang.

Remains valid – not updated: [JVET-AD2007](https://jvet-experts.org/doc_end_user/current_document.php?id=12977) Guidelines for NNVC software development [F. Galpin, S. Eadie, L. Wang, Z. Xie, Y. Li] (2023-05-26)

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-AF2016](https://jvet-experts.org/doc_end_user/current_document.php?id=13587) Common test conditions and evaluation procedures for neural network-based video coding technology [E. Alshina, R.-L. Liao, S. Liu, A. Segall] (2023-10-26)

This includes some editorial updates.

[JVET-AF2017](https://jvet-experts.org/doc_end_user/current_document.php?id=13588) Common test conditions and evaluation procedures for enhanced compression tool testing [M. Karczewicz, Y. Ye] (2023-11-03)

This is to include a modified Excel sheet, and mention the requirement of documenting decoder memory consumption (see discussion under JVET-AF0201).

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-AF2019](https://jvet-experts.org/doc_end_user/current_document.php?id=13580) Description of algorithms and software in neural network-based video coding (NNVC) version 5 [F. Galpin, Y. Li, D. Rusanovskyy, J. Ström, L. Wang] [WG 5 N 248] (2023-12-15)

New elements from notes elsewhere in this report:

….

Remains valid – not updated: [JVET-AE2020](https://jvet-experts.org/doc_end_user/current_document.php?id=13275) Film grain synthesis technology for video applications (Draft 5) [D. Grois, Y. He, W. Husak, P. de Lagrange, A. Norkin, M. Radosavljević, A. Tourapis, W. Wan] [WG 5 DTR N 223] (2023-09-08)

[JVET-AF2021](https://jvet-experts.org/doc_end_user/current_document.php?id=13589) Verification test plan for VVC multilayer coding (update 2) [S. Iwamura, P. de Lagrange, M. Wien] (2023-12-22)

See notes under JVET-AF0311 for updates.

Remains valid – not updated: [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)

According to discussions under section 4.12, more investigations are necessary on the new sequences. An update of this document was agreed to be postponed.

[JVET-AF2023](https://jvet-experts.org/doc_end_user/current_document.php?id=13578) Exploration experiment on neural network-based video coding (EE1) [E. Alshina, F. Galpin, Y. Li, D. Rusanovskyy, M. Santamaria, J. Ström, R. Chang, Z. Xie] [WG 5 N 247] (2023-11-03)

An initial draft of this document was reviewed and approved at 0900-0925 on Friday 20 Oct.

This round of EE1 tests will include:

* EE1-1: HOP
  + EE1-1.0: HOP re-training and luma/chroma balance changes ([JVET-AF0155](about:blank), [JVET-AF0180](about:blank)), possible tuning from JVET-AF0150, JVET-AF0296
  + EE1-1.1: EE1-0 with architecture change variant 1 (JVET-AF0102, JVET-AF0103, JVET-AF0182), as defined in [JVET-AF0307](https://jvet-experts.org/doc_end_user/current_document.php?id=13571).
  + EE1-1.2: EE1-0 with architecture change variant 2 (JVET-AF0102, JVET-AF0103, JVET-AF0153), as defined in [JVET-AF0307](https://jvet-experts.org/doc_end_user/current_document.php?id=13571).
  + EE1-1.3: Comparison test for HOP single model and two models (JVET-AF0183)
  + EE1-1.4: Study joint inference design of [JVET-AF0154](about:blank) (rotation) and [JVET-AF0086](about:blank) (flipping)
  + EE1-1.5: HOP In-loop filter with transformer blocks (JVET-AF0158)
* EE1-2: LOP
  + EE1-2.1: LOP/HOP fast training (JVET-AF0043, fast Stage III training)
  + EE1-2.2: LOP Content adaptive (JVET-AF0056)
  + EE1-2.3: Further complexity reduction of LOP (JVET-AF0071)
* EE1-3: inter prediction
  + EE1-3.1: Deep Reference Frame Generation for Inter Prediction Enhancement (JVET-AF0208)
* EE1-4: super-resolution
  + EE1-4.1: Unified CNN super resolution for resampling-based video coding (JVET-AF0143)

All tests in EE1 to use NNVC-7.0 as code base and to follow NNVC CTC, unless it is explicitly specified by the test description. The anchor for EE1 test is the default configuration of NNVC-7.0 as defined by AhG11/AhG14 (NN-intra and low complexity NN-filter enabled by default) in JVET-AE2016. Anchor performance and reference point for HOP NN-filters will be provided by AhG14.

All HOP tests should be done using stage 3 HOP official dataset. LOP tests uses official LOP stage 3 dataset, except EE1-2.1 which uses official HOP stage 3 dataset.

LOP-related tests will report results vs NNVC-7.0 Anchor (LOP.2 filter and NN-Intra are enabled by default). HOP-related tests to report results in comparison to HOP Anchor (HOP.1 filter and NN-Intra enabled), unless alternative anchor is specified in the test description.

[JVET-AF2024](https://jvet-experts.org/doc_end_user/current_document.php?id=13579) Exploration experiment on enhanced compression beyond VVC capability (EE2) [V. Seregin, J. Chen, R. Chernyak, K. Naser, J. Ström, F. Wang, M. Winken, X. Xiu, K. Zhang] [WG 5 N 249] (2023-11-17)

An initial draft of this document was reviewed and approved at 0925-0935 on Friday 20 Oct.

Categories and experiments are listed in the subsequent table:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Tests** | **Tester** | **Cross-checker** |
| **1 Intra prediction** | | | |
| 1.1a | Decoder derived CCP mode | Y.-J. Chang (Qualcomm) |  |
| 1.1b | Test 1.1a with decoder derived CCP fusion modes | Y.-J. Chang (Qualcomm) |  |
| 1.2 | IntraTMP with merge candidates | K. Naser (InterDigital) |  |
| 1.3 | SGPM with IntraTMP and IBC | K. Naser  (InterDigital) |  |
| 1.4 | IntraTMP extension to DIMD | K. Naser  (InterDigital) |  |
| 1.5 | IntraTMP extension to LIC | F. Le Léannec  (InterDigital) |  |
| 1.6 | Test 1.4 + Test 1.5 | F. Le Léannec  (InterDigital) |  |
| 1.7 | Test 1.6 + Test 1.3 | K. Naser  (InterDigital) |  |
| 1.8a | The length of the auto-relocated BVP trace path is 1 (i.e, n=1) | N. Zhang  (Bytedance) |  |
| 1.8b | The length of the auto-relocated BVP trace path is 2 (i.e, n=2) | N. Zhang  (Bytedance) |  |
| 1.8c | No constraint for the length of the auto-relocated BVP trace path | N. Zhang  (Bytedance) |  |
| 1.9 | Intra TMP fusion probing | J.-L. Lin  (Qualcomm) |  |
| 1.10 | Bilateral filtering for intra prediction | W. Yin  (Bytedance) |  |
| 1.11a | IntraTMP search area extension with sampling factor proportional to CU distance | D. Ruiz Coll  (Ofinno) |  |
| 1.11b | IntraTMP search area extension with sampling factor proportional to search area dimension | K. Naser  (InterDigital) |  |
| 1.11c | Test 1.11a + Test 1.11b | D. Ruiz Coll  (Ofinno )  K. Naser  (InterDigital) |  |
| 1.12a | IBC with extended reference area up to the picture’s upper side boundary | Y. Kidani  (KDDI) |  |
| 1.12b | IBC with four CTU rows instead of two CTU rows in HD resolution or less | Y. Kidani  (KDDI) |  |
| 1.13a | Test 1.11c + Test 1.12a | Y. Kidani  (KDDI)  D. Ruiz Coll  (Ofinno )  K. Naser  (InterDigital) |  |
| 1.13b | Test 1.11c + Test 1.12b | Y. Kidani  (KDDI)  D. Ruiz Coll  (Ofinno )  K. Naser  (InterDigital) |  |
| 1.14a | Encoder run-time reduction methods for extrapolation filter-based intra prediction mode only | L. Xu (OPPO) |  |
| 1.14b | Encoder run-time reduction methods for extrapolation filter-based intra prediction and its merge mode | L. Xu (OPPO) |  |
| 1.15a | CCP merge fusion | H. Huang  (OPPO)  Z. Deng  (Bytedance) |  |
| 1.15b | Inheriting LB-CCP flag in CCP merge mode | H. Huang  (OPPO)  Z. Deng  (Bytedance) |  |
| 1.15c | Test1.15a + Test1.15b | H. Huang  (OPPO)  Z. Deng  (Bytedance) |  |
| 1.16 | Slope adjustment for IBC LIC | C. Ma  (Kwai) |  |
| 1.17a | IBC GPM with block vector difference | C. Ma  (Kwai) |  |
| 1.17b | IBC GPM with split mode reordering | C. Ma  (Kwai) |  |
| 1.17c | Test 1.17a + Test 1.17b | C. Ma  (Kwai) |  |
| 1.18a | DIMD mode derivation from spatial blocks | J. Huo, J. Fan  (Xidian Univ.)  M. Li  (OPPO) |  |
| 1.18b | DIMD mode derivation with reduced complexity | J. Huo, J. Fan  (Xidian Univ.)  M. Li  (OPPO) |  |
| 1.19 | IBC-LIC model merge mode | L. Zhang  (OPPO) |  |
| 1.20a | TIMD fusion with non-angular predictor | P. Andrivon (Ofinno) |  |
| 1.20b | Test 1.20a + TIMD sample-based fusion | P. Andrivon (Ofinno) |  |
| 1.21 | TIMD fusion reference line determination | C. Zhou (vivo) |  |
| 1.22a | Test 1.20a + Test 1.21 | P. Andrivon (Ofinno)  C. Zhou  (vivo) |  |
| 1.22b | Test 1.20b + Test 1.21 | P. Andrivon (Ofinno)  C. Zhou (vivo) |  |
| **2 Inter prediction** | | | |
| 2.1 | A LIC flag for inter prediction merge modes | Y. Zhang (Qualcomm) |  |
| 2.2 | Enable PU level BDMVR and BDOF for bi-predicted LIC | Y. Zhang (Qualcomm) | A. Robert (InterDigital) |
| 2.3 | LIC with multiple templates | Y. Wang  (Bytedance)  X. Xiu  (Kwai) |  |
| 2.4 | LIC with slope adjustment | Y. Wang  (Bytedance) |  |
| 2.5 | Non-local illumination compensation | X. Xiu  (Kwai) |  |
| 2.6a | Test 2.1 + Test 2.2 | Y. Zhang (Qualcomm) |  |
| 2.6b | Test 2.3 + Test 2.4 | Y. Wang  (Bytedance)  X. Xiu  (Kwai) |  |
| 2.6c | Test 2.3 + Test2.5 | X. Xiu  (Kwai)  Y. Wang  (Bytedance) |  |
| 2.7 | AMVP with SbTMVP mode | R.-L. Liao  (Alibaba) |  |
| 2.8a | DMVR for non-equal POC distance cases | M. Salehifar  (Bytedance) | P. Le Guyadec (InterDigital) |
| 2.8b | DMVR for non-equal POC distance cases with constraint | M. Salehifar  (Bytedance) |  |
| 2.8c | BDOF layer with new subblock sizes including 16×16 | M. Salehifar  (Bytedance) |  |
| 2.8d | Mean removed formula for BDOF | M. Salehifar  (Bytedance) |  |
| 2.8e | Test 2.8b + Test 2.8c + Test 2.8d | M. Salehifar  (Bytedance) |  |
| 2.9 | CIIP with subblock-based motion compensation | L. Zhao  (Bytedance) |  |
| 2.10 | GPM with affine prediction | K. Zhang  (Bytedance) |  |
| 2.11 | Regression-based GPM blending | P. Bordes  (InterDigital ) |  |
| 2.12 | Utilizing LFNST/NSPT for inter coding | F. Wang  (OPPO) | F. Le Léannec (InterDigital) |
| 2.13 | Adjusting out-of-boundary prediction samples | P. Astola  (Nokia) |  |
| **3** **Transform and coefficients coding** | | | |
| 3.1a | Intra contexts for intra CUs in inter slices | F. Lo Bianco (InterDigital) |  |
| 3.1b | Test 3.1a + inter contexts for IntraTMP/IBC CUs | F. Lo Bianco (InterDigital) |  |
| 3.1c | Test 3.1b without context switch for CBF | F. Lo Bianco (InterDigital) |  |
| 3.1d | Test 3.1 with retrained context initialization of the affected by the test contexts | F. Lo Bianco (InterDigital) |  |
| 3.2a | Transform coefficient coding | P. Nikitin  (Qualcomm) |  |
| 3.2b | Test 3.2a with retrained context retrained context initialization | P. Nikitin  (Qualcomm) |  |
| **4 In-loop filtering** | | | |
| 4.1a | Adaptive clipping with signalled min/max values from MCTF prefiltered picture | K. Cui  (Qualcomm) |  |
| 4.1b | Adaptive clipping with signalled min/max values from original picture | K. Cui  (Qualcomm) |  |
| 4.2 | Fixed filter for chroma ALF | N. Hu  (Qualcomm) |  |
| 4.3 | Coefficient precision adjustment for ALF | W. Yin  (Bytedance)  N. Hu  (Qualcomm) |  |
| **5 Entropy coding** | | | |
| 5.1 | Retrain CABAC initialization parameters for all slices | F. Galpin  (InterDigital)  V. Seregin  (Qualcomm) |  |
| 5.2 | Window slice type-dependent offsets initialization | V. Seregin  (Qualcomm)  F. Galpin (InterDigital) |  |
| 5.3a | Spatial CABAC tuning | J. Lainema  (Nokia) |  |
| 5.3b | Spatial CABAC tuning in combination with retrained context initialization for inter slices | J. Lainema  (Nokia) |  |

[JVET-AF2025](https://jvet-experts.org/doc_end_user/current_document.php?id=13581) Algorithm description of Enhanced Compression Model 11 (ECM 11) [M. Coban, R.-L. Liao, K. Naser, J. Ström, L. Zhang] [WG 5 N 250] (2023-12-15)

New elements from notes elsewhere in this report:

* ….

It is noted that the list above may not be complete; if some adoption is missing that is recorded somewhere else in the meeting notes it shall also be considered included.

Remains valid – not updated: [JVET-AC2026](https://jvet-experts.org/doc_end_user/current_document.php?id=12581) Conformance testing for VVC operation range extensions (Draft 4) [D. Rusanovskyy, T. Hashimoto, H.-J. Jhu, I. Moccagatta, Y. Yu] (2023-04-14)

This was integrated with v1 (JVET-X2008) and delivered for ITU-T consent as H.266.1 2nd ed.

Primary editor for the ITU consent text: I. Moccagatta.

[JVET-AF2027](https://jvet-experts.org/doc_end_user/current_document.php?id=13590) SEI processing order and processing order nesting SEI messages in VVC (draft 6) [G. J. Sullivan, M. M. Hannuksela, Y.-K. Wang] [WG 5 preliminary WD 5 N 246] (2023-12-01)

Primary editor: G. J. Sullivan

Updated from JVET-AF0049, JVET-AF0062, JVET-AF0070, JVET-AF0189, JVET-AF0310.

Remains valid – not updated: [JVET-AE2028](https://jvet-experts.org/doc_end_user/current_document.php?id=13279) Additional conformance bitstreams for VVC multilayer configurations [S. Iwamura, P. de Lagrange, I. Moccagatta] (2023-09-01)

Remains valid – not updated: [JVET-AB2029](https://jvet-experts.org/doc_end_user/current_document.php?id=12225) Visual quality comparison of ECM/VTM encoding [V. Baroncini, J.-R. Ohm, M. Wien] [AG 5 N 75]

Remains valid – not updated: [JVET-AE2030](https://jvet-experts.org/doc_end_user/current_document.php?id=13280) Optimization of encoders and receiving systems for machine analysis of coded video content (draft 3) [J. Chen, C. Hollmann, S. Liu] [WG 5 N 224)] (2023-09-15)

Expect CDTR April 2024. Could later be defined as part of MPEG-AI, or MPEG-C

[JVET-AF2031](https://jvet-experts.org/doc_end_user/current_document.php?id=13591) Common test conditions for optimization of encoders and receiving systems for machine analysis of coded video content [S. Liu, C. Hollmann] (2023-11-17)

Updates: Only one anchor (VTM20 or newest VTM), and prescribe to use an RA period of 1 s in case of temporal subsampling, when comparing against an anchor without subsampling.

[JVET-AF2032](https://jvet-experts.org/doc_end_user/current_document.php?id=13592) Technologies under consideration for future extensions of VSEI (version 2) [M. M. Hannuksela, J. Chen, S. Deshpande, Hendry, S. McCarthy] [WG 5 N 242)] (2023-12-01)

Elements from XXXX; see detailed notes under these document numbers.

[JVET-AF2033](https://jvet-experts.org/doc_end_user/current_document.php?id=13593) Report of verification test on VVC multi-layer coding: Content layering [S. Iwamura, P. de Lagrange, M. Wien] [AG 5 N 105)] (2023-12-22)

Includes a description of the content layering functional test conducted during the 31st meeting, and the outcome.

# Future meeting plans, expressions of thanks, and closing of the meeting

Future meeting plans were established according to the following guidelines (assuming face-to-face meetings):

* Meeting under ITU-T SG16 auspices when it meets (ordinarily starting meetings on the Tuesday or Wednesday of the first week and closing it on the Wednesday of the second week of the SG16 meeting – a total of 8-9 meeting days), and
* Otherwise meeting under ISO/IEC JTC 1/‌SC 29 auspices when its MPEG WGs meet (ordinarily starting meetings on the Thursday or Friday prior to the main week of such meetings and closing it on the same day as other MPEG WGs – a total of 8-9 meeting days).

In cases where an exceptionally high workload is expected for a meeting, an earlier starting date may be defined. In cases of online meetings, no sessions should be held on weekend days, such that meetings would typically start two days earlier.

Some specific future meeting plans (to be confirmed) were established as follows:

* During Wed. 17 – 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 Wed. 17 – Wed. 24 April 2024, 34th meeting under ITU-T SG16 auspices in Rennes, FR,
* During Fri. 12 – Fri. 19 July 2024, 35th meeting under ISO/IEC JTC 1/‌SC 29 auspices in Sapporo, JP,
* During Fri. 1 – Fri. 8 November 2024, 36th meeting under ISO/IEC JTC 1/‌SC 29 auspices, location t.b.d. (expected face-to-face in Turkey or teleconference),
* 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,
* During October 2025, 40th meeting under ITU-T SG16 auspices, date and location t.b.d.

The agreed document deadline for the 33rd JVET meeting was planned to be Wednesday 10 January 2024.

Apple was thanked for providing additional test materials with synthesized film grain that can be used in assessing and developing technology in the context of film grain synthesis applications.

Marius Preda was thanked for maintaining the document site jvet-experts.org, as well as the document sites of JCT-VC and JCT-3V. Institut Mines-Télécom was thanked for hosting the sites.

Silke Kenzler, as well as the teams of KCM and HCC, were thanked for the support provided during the meeting.

Institut für Informationsverarbeitung of Leibniz Universität Hannover was thanked for sponsoring the social event.

The 32nd JVET meeting was closed at approximately 1724 hours CEST (UTC+2) on Friday 20 Oct. 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-second meeting of the JVET, according to confirming in a sign-in sheet regularly circulated in the JVET meeting rooms (approximately XXX people in total), were as follows:

1. XXX (XXX – XX)

# Annex B2 to JVET report: List of meeting participants attending remotely

The remote participants of the thirty-second meeting of the JVET, according to the participation records from the Zoom teleconferencing tool used for the meeting sessions (approximately XXX people in total, not including those who had attended the meeting in person at least part-time (see annex B1), and not including those who attended only the joint sessions with other groups), were as follows:

1. XXX (XXX – XX)

# Annex C to JVET report: Recommendations of the 13th meeting of ISO/IEC JTC 1/SC 29/WG 5 MPEG Joint Video Experts Team with ITU-T SG 16

**ISO/IEC JTC 1/SC 29/WG 5 N XXX**