|  |  |
| --- | --- |
| **Joint Video Experts Team (JVET)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29**  28th Meeting: Mainz, DE, 20–28 October 2022 | Document: JVET-AB\_Notes\_d9 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Meeting Report of the 28th Meeting of the Joint Video Experts Team (JVET), Mainz, DE, 20–28 October 2022** | | |
| *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 twenty-eighth meeting face-to-face during 20–28 October 2022 in Mainz, Germany, at Erbacher Hof (Grebenstr. 24-26, D-55116, Mainz, Tel: +49 (0) 6131 257-0, web <http://www.ebh-mainz.de>). Remote participation was provided for experts who were unable to travel.

The arrangements for the 28th meeting had not been clear yet by the time of closing the last meeting, but it had been agreed to give the chair the discretion to clarify the situation and communicate via the JVET reflector as soon as possible. The meeting had originally been planned to be held under ITU-T SG16 auspices during Fri. 21 – Fri. 28 October 2022 in Antalya, TR, but that plan had to be cancelled due to difficulties of identifying a suitable meeting venue. ITU-T SG16 later decided to hold their meeting as face-to-face meeting one week earlier (during Mon. 17 – Fri. 28 October 2022) in Geneva, CH, but indicated there was not sufficient room space available in the ITU premises to also host JVET. Therefore, it was not possible to hold the 28th meeting as face-to-face meeting under ITU-T SG16 auspices, and SG16 management agreed that JVET could meet together with other MPEG WGs under SC 29 auspices in Mainz, DE. This also implied that JVET would not need to finish the meeting by Wed. 26 Oct. (before SG16 would start their final WP and SG plenaries); however, it would be necessary to approve and finalize documents to be submitted for ITU-T consent already by that date, such that starting the meeting one day earlier than (by Thu. 20 Oct.) was considered to be necessary.

For ISO/IEC purposes, JVET is alternatively designated ISO/IEC JTC 1/‌SC 29/‌WG 5, and this was the ninth 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 2.14 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 hours CEST on Thursday 20 October 2022. Meeting sessions were held on all days, including the weekend days of Saturday and Sunday 22 and 23 October 2022, until the meeting was closed at approximately XXXX hours CEST on Friday 28 October 2022. Approximately XXX people attended the JVET meeting (XXX in presence, and XXX remotely), and approximately XXX input documents (not counting crosschecks), 13 AHG reports, 2 EE summary reports, X BoG reports, and X liaison documents were discussed. The meeting took place in coordination with a meeting of various SC29 Working Groups and Advisory Groups – where WG 5 is representing the Joint Video Coding Team(s) and their activities from the perspective of the SC 29 parent body, under whose auspices this JVET meeting was held. The subject matter of the JVET meeting activities consisted of work on further development and maintenance of the twin-text video coding technology standards *Advanced Video Coding* (AVC), *High Efficiency Video Coding* (HEVC), *Versatile Video Coding* (VVC)*, Coding-independent Code Points (Video)* (CICP), and *Versatile Supplemental Enhancement Information Messages for Coded Video Bitstreams* (VSEI), as well as related technical reports, reference software and conformance testing packages. Further important goals were reviewing the results of the Exploration Experiment (EE) on Neural Network-based Video Coding, of the EE on Enhanced Compression beyond VVC capability, of other technical input on novel aspects of video coding technology, and to plan next steps for investigation of candidate technology towards further standard development.

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

* JVET-AA1004 Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR
* JVET-AA1011 HEVC multiview profiles supporting extended bit depth (draft 1)
* JVET-AA1100 Common Test Conditions for HM Video Coding Experiments
* JVET-AA2005 New level and systems-related supplemental enhancement information for VVC (Draft 3), also issued as WG 5 DAM
* JVET-AA2006 Additional SEI messages for VSEI (Draft 2), also issued as WG 5 CDAM
* JVET-AA2016 Common Test Conditions and evaluation procedures for neural network-based video coding technology
* JVET-AA2018 Common test conditions for high bit depth and high bit rate video coding
* JVET-AA2020 Film grain synthesis technology for video applications (Draft 2), also issued as WG 5 WD
* JVET-AA2023 Exploration Experiment on neural network-based video coding (EE1)
* JVET-AA2024 Exploration Experiment on enhanced compression beyond VVC capability (EE2)
* JVET-AA2025 Algorithm description of Enhanced Compression Model 6 (ECM 6)
* JVET-AA2027 SEI processing order SEI message in VVC (draft 1)

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

* JVET-AA1004 Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR
* JVET-AA1011 HEVC multiview profiles supporting extended bit depth (draft 1)
* JVET-AA1100 Common Test Conditions for HM Video Coding Experiments
* JVET-AA2005 New level and systems-related supplemental enhancement information for VVC (Draft 3), also issued as WG 5 DAM
* JVET-AA2006 Additional SEI messages for VSEI (Draft 2), also issued as WG 5 CDAM
* JVET-AA2016 Common Test Conditions and evaluation procedures for neural network-based video coding technology
* JVET-AA2018 Common test conditions for high bit depth and high bit rate video coding
* JVET-AA2020 Film grain synthesis technology for video applications (Draft 2), also issued as WG 5 WD
* JVET-AA2023 Exploration Experiment on neural network-based video coding (EE1)
* JVET-AA2024 Exploration Experiment on enhanced compression beyond VVC capability (EE2)
* JVET-AA2025 Algorithm description of Enhanced Compression Model 6 (ECM 6)
* JVET-AA2027 SEI processing order SEI message in VVC (draft 1)

For the organization and planning of its future work, the JVET established XX “ad hoc groups” (AHGs) to progress the work on particular subject areas. At this meeting, 2 Exploration Experiments (EE) were defined. The next eight JVET meetings were planned for 11 – 13 and 16 – 20 January 2023 under ISO/IEC JTC 1/‌SC 29 auspices, to be conducted as a teleconference meeting; during 21 – 28 April 2023 under ISO/IEC JTC 1/‌SC 29 auspices, in Antalya, TR; during 14 – 21July 2023 under ITU-T SG16 auspices in Geneva, CH; during October 2023 under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.; during January 2024 under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.; during April 2024 under ITU-T SG16 auspices, date and location t.b.d.; during 12 – 19 July 2024 under ISO/IEC JTC 1/‌SC 29 auspices in Sapporo, JP; and during October 2024 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 twenty-eighth meeting face-to-face during 20–28 October 2022 in Mainz, Germany, at Erbacher Hof (Grebenstr. 24-26, D-55116, Mainz, Tel: +49 (0) 6131 257-0, web <http://www.ebh-mainz.de>). For ISO/IEC purposes, JVET is alternatively designated ISO/IEC JTC 1/‌SC 29/‌WG 5, and this was the ninth meeting as WG 5. The JVET meeting was held under the chairmanship of Dr Jens-Rainer Ohm (RWTH Aachen/Germany).

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

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

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

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

This report contains three important annexes, as follows:

* Annex A contains a list of the documents of the JVET meeting
* Annex B contains a list of the meeting participants, as recorded by the teleconferencing tool used for the meeting
* Annex C contains the meeting recommendations of ISO/IEC JTC 1/‌SC 29/‌WG 5 for purposes of results reporting to ISO/IEC.

## Meeting logistics

Information regarding logistics arrangements for the meeting had been provided via the email reflector [jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de) and at <http://wftp3.itu.int/av-arch/jvet-site/2022_10_AB_Mainz/>.

## Primary goals

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

* JVET-AA1004 Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR
* JVET-AA1011 HEVC multiview profiles supporting extended bit depth (draft 1)
* JVET-AA1100 Common Test Conditions for HM Video Coding Experiments
* JVET-AA2005 New level and systems-related supplemental enhancement information for VVC (Draft 3), also issued as WG 5 DAM
* JVET-AA2006 Additional SEI messages for VSEI (Draft 2), also issued as WG 5 CDAM
* JVET-AA2016 Common Test Conditions and evaluation procedures for neural network-based video coding technology
* JVET-AA2018 Common test conditions for high bit depth and high bit rate video coding
* JVET-AA2020 Film grain synthesis technology for video applications (Draft 2), also issued as WG 5 WD
* JVET-AA2023 Exploration Experiment on neural network-based video coding (EE1)
* JVET-AA2024 Exploration Experiment on enhanced compression beyond VVC capability (EE2)
* JVET-AA2025 Algorithm description of Enhanced Compression Model 6 (ECM 6)
* JVET-AA2027 SEI processing order SEI message in VVC (draft 1)

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.

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 also follow the CEST timezone (local time in Mainz).

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

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

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

### Late and incomplete document considerations

The formal deadline for registering and uploading non-administrative contributions had been announced as Friday, 14 October 2022. Any documents uploaded after 1159 hours Paris/Geneva time on Saturday 15 October 2022 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-AB0192 were registered after the “officially late” deadline (and therefore were also uploaded late). However, some documents in the “late” range might include break-out activity reports that were generated during the meetings, 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-AB0XXX (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-AB0XXX (a document on …), uploaded 10-XX,
* …

Most cross-verification reports at this meeting were registered late, and/or uploaded late (except for JVET-AB0063, JVET-AB0088, JVET-AB0137, and JVET-AB0150). In the interest of brevity, the late ones 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-AB0113, JVET-AB0167, JVET-AB0250, JVET-AB0278.

The following cross-verification reports were still missing three weeks after the end of the meeting: JVET-AB0XXX, … . They were thus considered to become withdrawn.

“Placeholder” contribution documents that were basically empty of content, or lacking any results showing benefit for the proposed technology, and obviously uploaded with an intent to provide a more complete submission as a revision, had been agreed to be considered unacceptable and to be rejected in the document management system until a more complete version was available (which would then be counted as a late contribution if the update was after the document deadline). At the current meeting, this situation did apply with documents JVET-AB0XXX and …, which were both categorized as late in the list above, based on the time of the first reasonable document upload.

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

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

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

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

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

### Outputs of the preceding meeting

All output documents of the previous meeting, particularly the meeting report JVET-AA1000, the Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR JVET-AA1004, the HEVC multiview profiles supporting extended bit depth (Draft 1) JVET-AA1011, the Common test conditions for HM video coding experiments JVET-AA1100, the New level and systems-related supplemental enhancement information for VVC (Draft 3) JVET-AA2005, the Additional SEI messages for VSEI (Draft 2) JVET-AA2006, the Common test conditions and evaluation procedures for neural network-based video coding technology JVET-AA2016, the Film grain synthesis technology for video applications (Draft 2) JVET-AA2020, the Description of the EE on Neural Network-based Video Coding JVET-AA2023, the Description of the EE on Enhanced Compression beyond VVC capability JVET-AA2024, the Algorithm description of Enhanced Compression Model 6 (ECM 6) JVET-AA2025, and the SEI processing order SEI message in VVC (Draft 1) JVET-AA2027, had been completed and were approved. The software implementations of HM (version 16.26), VTM (versions 17.1, 17.2, 18.0, and 18.1), and ECM (version 5.1 and 6.0) were also approved. Furthermore, in the context of the AHG11 report, the software implementation of “NCS1.0” (neural compression software) was approved. It was suggested to discuss about a name for that software during the meeting (see discussion under section 4.8).

Only minor editorial issues were found in the meeting report JVET-AA1000; no need to produce an update was identified (see section 2.12 for details).

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

## Attendance

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

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

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

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

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

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

## Agenda

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

* Opening remarks and review of meeting logistics and communication practices
* Roll call of participants
* Adoption of the agenda
* Code of conduct policy reminder
* IPR policy reminder and declarations
* Contribution document allocation
* Review of results of the previous meeting
* Review of target dates
* Reports of ad hoc group (AHG) activities
* Report of exploration experiments on neural-network-based video coding
* Report of exploration experiments on enhanced compression beyond VVC capability
* Consideration of contributions on high-level syntax
* Consideration of contributions and communications on project guidance
* Consideration of video coding technology contributions
* Consideration of contributions on conformance and reference software development
* Consideration of contributions on coding-independent code points for video signal type identification
* Consideration of contributions on film grain synthesis technology
* Consideration of contributions on 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 will follow 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 will be 0900-1900 CEST with coffee breaks and lunch breaks as appropriate, however some early morning or late night sessions may be necessary. Sessions will be announced in the JVET calendar in advance as far as possible, but it might happen that some activities (such as breakout sessions) will be held at short notice.

* 1. ***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/PUB100397.html>.

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

These include points relating to:

* Respecting others
* Behaving ethically
* Escalating and resolving disputes
* Working for the net benefit of the international community
* Upholding consensus and governance
* Agreeing to a clear purpose and scope
* Participating actively and managing effective representation

## IPR policy reminder

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

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

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

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

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

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

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

## Software copyright disclaimer header reminder

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

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

These considerations apply to the 360Lib video conversion software and HDRTools as well. It is noted that the SADL and NCS? 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 1185 (as of 17 October 2022). 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).

## Terminology

* **ACT**: Adaptive colour transform
* **AFF**: Adaptive frame-field
* **AI**: All-intra
* **AIF**: Adaptive interpolation filtering
* **ALF**: Adaptive loop filter
* **AMP**: Asymmetric motion partitioning – a motion prediction partitioning for which the sub-regions of a region are not equal in size (in HEVC, being N/2x2N and 3N/2x2N or 2NxN/2 and 2Nx3N/2 with 2N equal to 16 or 32 for the luma component)
* **AMVP**: Adaptive motion vector prediction
* **AMT or MTS**: Adaptive multi-core transform, or multiple transform selection
* **AMVR**: (Locally) adaptive motion vector resolution
* **APS**: Adaptation parameter set
* **ARC**: Adaptive resolution conversion (synonymous with DRC, and a form of RPR)
* **ARMC**: Adaptive re-ordering of merge candidates
* **ARSS**: Adaptive reference sample smoothing
* **ATM**: AVC-based multiview and 3D test model
* **ATMVP** or “subblock-based temporal merging candidates”: Alternative temporal motion vector prediction
* **AU**: Access unit
* **AUD**: Access unit delimiter
* **AVC**: Advanced video coding – the video coding standard formally published as ITU-T Recommendation H.264 and ISO/IEC 14496-10
* **BA**: Block adaptive
* **BC**: See CPR or IBC
* **BCW**: Biprediction with CU based weighting
* **BD**: Bjøntegaard-delta – a method for measuring percentage bit rate savings at equal PSNR or decibels of PSNR benefit at equal bit rate (e.g., as described in document VCEG-M33 of April 2001)
* **BDOF**: Bi-directional optical flow (formerly known as **BIO**)
* **BDPCM**: Block-wise DPCM
* **BL**: Base layer
* **BMS**: Benchmark set (no longer used), a former preliminary compilation of coding tools on top of VTM, which provide somewhat better compression performance, but are not deemed mature for standardzation
* **BoG**: Break-out group
* **BR**: Bit rate
* **BT**: Binary tree
* **BV**: Block vector (used for intra BC prediction)
* **CABAC**: Context-adaptive binary arithmetic coding
* **CBF**: Coded block flag(s)
* **CC**: May refer to context-coded, common (test) conditions, or cross-component
* **CCALF**: Cross-component ALF
* **CCLM**: Cross-component linear model
* **CCCM**: Cross-component convolutional model
* **CCP**: Cross-component prediction
* **CCSAO**:Cross-component SAO
* **CE**: Core Experiment – a coordinated experiment conducted toward assessment of coding technology
* **CG**: Coefficient group
* **CGS**: Colour gamut scalability (historically, coarse-grained scalability)
* **CIIP**: Combined inter/intra prediction
* **CIPF**: CABAC initialization from the previous frame
* **CL-RAS**: Cross-layer random-access skip
* **CPB**: Coded picture buffer
* **CPMV**: Control-point motion vector
* **CPMVP**: Control-point motion vector prediction (used in affine motion model)
* **CPR**: Current-picture referencing, also known as IBC – a technique by which sample values are predicted from other samples in the same picture by means of a displacement vector called a block vector, in a manner conceptually similar to motion-compensated prediction
* **CST**: Chroma separate tree
* **CTC**: Common test conditions
* **CVS**: Coded video sequence
* **DCI**: Decoder capability information
* **DCT**: Discrete cosine transform (sometimes used loosely to refer to other transforms with conceptually similar characteristics)
* **DCTIF**: DCT-derived interpolation filter
* **DF**: Deblocking filter
* **DIMD**: Decoder intra mode derivation
* **DMVR**: Decoder motion vector refinement
* **DoCR**: Disposition of comments report
* **DPB**: Decoded picture buffer
* **DPCM**: Differential pulse-code modulation
* **DPS**: Decoding parameter sets
* **DRC**: Dynamic resolution conversion (synonymous with ARC, and a form of RPR)
* **DT**: Decoding time
* **DQ**: Dependent quantization
* **ECS**: Entropy coding synchronization (typically synonymous with WPP)
* **EMT**: Explicit multiple-core transform
* **EOTF**: Electro-optical transfer function – a function that converts a representation value to a quantity of output light (e.g., light emitted by a display
* **EPB**: Emulation prevention byte (as in the emulation\_prevention\_byte syntax element)
* **ECM**: Enhanced compression model – a software codebase for future video coding exploration
* **ECV**: Extended Colour Volume (up to WCG)
* **EL**: Enhancement layer
* **EOS**: End of (coded video) sequence
* **ET**: Encoding time
* **FRUC**: Frame rate up conversion (pattern matched motion vector derivation)
* **GCI**: General constraints information
* **GDR**: Gradual decoding refresh
* **GLM**: Gradient linear model
* **GOP**: Group of pictures (somewhat ambiguous)
* **GPM**: Geometry partitioning mode
* **GRA**: Gradual random access
* **HBD**: High bit depth
* **HDR**: High dynamic range
* **HEVC**: High Efficiency Video Coding – the video coding standard developed and extended by the JCT-VC, formalized by ITU-T as Rec. ITU-T H.265 and by ISO/IEC as ISO/IEC 23008-2
* **HLS**: High-level syntax
* **HM**: HEVC Test Model – a video coding design containing selected coding tools that conforms to the HEVC standard design (possibly with under-development extensions) – now also used especially in reference to the (non-normative) encoder algorithms (see WD and TM)
* **HMVP**: History based motion vector prediction
* **HRD**: Hypothetical reference decoder
* **HTM**: HEVC-based multiview and 3D test model (developed by JCT-3V)
* **HyGT**: Hyper-cube Givens transform (a type of NSST)
* **IBC** (also **Intra BC**): Intra block copy, also known as CPR – a technique by which sample values are predicted from other samples in the same picture by means of a displacement vector called a block vector, in a manner conceptually similar to motion-compensated prediction
* **IBDI**: Internal bit-depth increase – a technique by which lower bit-depth (8 bits per sample) source video is encoded using higher bit-depth signal processing, ordinarily including higher bit-depth reference picture storage (ordinarily 12 bits per sample)
* **IBF**: Intra boundary filtering
* **ILP**: Inter-layer prediction (in scalable coding)
* **ILRP**: Inter-layer reference picture
* **IPCM**: Intra pulse-code modulation (similar in spirit to IPCM in AVC and HEVC)
* **IRAP**: Intra random access picture
* **ISP**: Intra subblock partitioning
* **JCCR**: Joint coding of chroma residuals
* **JCT-3V**: Joint collaborative team on 3D video (for AVC and HEVC)
* **JCT-VC**: Joint collaborative team on video coding (for HEVC)
* **JEM**: Joint exploration model – a software codebase previously used for video coding exploration
* **JM**: Joint model – the primary software codebase that has been developed for the AVC standard
* **JSVM**: Joint scalable video model – another software codebase that has been developed for the AVC standard, which includes support for scalable video coding extensions
* **JVET**: Joint video experts team (initially for VVC, later expanded)
* **JVT**: Joint video team (for AVC)
* **KLT**: Karhunen-Loève transform
* **LB** or **LDB**: Low-delay B – the variant of the LD conditions that uses B pictures
* **LD**: Low delay – one of two sets of coding conditions designed to enable interactive real-time communication, with less emphasis on ease of random access (contrast with RA). Typically refers to LB, although also applies to LP
* **LFNST**: Low-frequency non-separable transform
* **LIC**: Local illumination compensation
* **LM**: Linear model
* **LMCS**: Luma mapping with chroma scaling (formerly sometimes called “in-loop reshaping”)
* **LP** or **LDP**: Low-delay P – the variant of the LD conditions that uses P frames
* **LUT**: Look-up table
* **LTRP**: Long-term reference picture
* **MANE**: Media-aware network element
* **MC**: Motion compensation
* **MCP**: Motion compensated prediction
* **MCTF**: Motion compensated temporal pre-filtering
* **MDNSST**: Mode dependent non-separable secondary transform
* **MIP**: Matrix-based intra prediction
* **MMLM**: Multi-model (cross component) linear mode
* **MMVD**: Merge with MVD
* **MPEG**: Moving picture experts group (an alliance of working groups and advisory groups in ISO/IEC JTC 1/‌SC 29, one of the two parent bodies of the JVET)
* **MPM**: Most probable mode (in intra prediction)
* **MRL**: Multiple reference line intra prediction
* **MV**: Motion vector
* **MVD**: Motion vector difference
* **NAL**: Network abstraction layer
* **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).
* **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).
* **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.

## Opening remarks

Remarks during the opening session of the meeting Thursday 20 October at 0900 CEST were as follows.

* Timing and organization of the F2F meetings and online access, calendar posting of session plans
* Plan to deal only with ITU deliverables (to be approved October 26 at latest) and related AHG reports and input docs on the first day
* Other AHG reports, and EE review to start on Friday 21 Oct.
* Subjective viewing of ECM vs. VTM planned for Sunday 23 Oct. and Monday 24 Oct. – avoid extensive JVET sessions in parallel
* Standards, TRs, supplements and technical papers approval and publication status
  + AVC
    - H.264 V14 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 14496-10:202X (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, currently in consultation between ISO EPM and project editors
    - Preliminary draft text for YCgCo-Re and YCgCo-Ro issued at 26th meeting (not yet formally requested as a project)
    - Conformance testing
      * 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 43:20153D-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
      * H.264.2 V7 Approved 2016-02-13, published 2016-05-30
      * Various amendments of ISO/IEC 14496-5:2001, 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
    - 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
    - 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 April 2022, text submitted for DIS ballot 2022-07-10
    - Preliminary draft text for YCgCo-Re and YCgCo-Ro issued at 26th meeting (not yet formally requested as a project)
    - Conformance testing
      * 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
      * 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
    - Reference software
      * 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 2017-02
      * ISO/IEC 23008-5:2017/AMD 1:2017 Reference software for screen content coding extensions, published 2017-10
  + VVC
    - H.266 V1 approved 2020-08-29, published 2020-11-10
    - ISO/IEC 23090-3:2021 (Ed. 1) published 2021-02-16
    - 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 23090-3:202x (Ed. 2) with operation range extensions, approval at WG level to proceed to FDIS 2022-01-21, was published 2022-09-XX
    - ISO/IEC 23090-3:202x (Ed. 2) / CDAM 1 New level and systems-related supplemental enhancement information issued from 26th meeting, ballot to close 2022-07-14
    - Conformance testing
      * 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 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
      * ISO/IEC 23090-15 DAM 1 Operation range extensions – DAM 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 closes 2022-11-15, no action at this meeting
    - Reference software
      * 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 V1 approval at WG level to proceed to FDIS 2022-01-21, FDIS ballot pending, upgraded to “DIS approved for registration” status in ISO Projects system 2022-04-21, upgraded to “FDIS registered for formal approval” 2022-04-22
  + VSEI
    - H.274 V1 approved 2020-08-29, published 2020-11-10
    - ISO/IEC 23002-7:2021 (Ed. 1) published 2021-01-28
    - 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
    - ISO/IEC 23002-7:202x (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 pending
    - 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
  + CICP
    - ISO/IEC 23091-2:2021 (Ed. 2) had been forwarded from DIS directly for publication in 2021-04 and published 2021-10-18
    - 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
  + Conversion and coding practices for HDR/WCG Y′CbCr 4:2:0 video with PQ transfer characteristics
    - 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
    - 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
    - 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
    - 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
    - ISO/IEC TR 23002-9 Request for subdivision and WD 1 issued at 25th meeting 2022-01-21, uploaded 2022-04-20, WD 2 issued at 27th meeting
  + The following freely available standards are published here in ISO/IEC:  
    <https://standards.iso.org/ittf/PubliclyAvailableStandards/index.html>
    - ISO/IEC 14496-10:2020 (Ed. 9) AVC
    - ISO/IEC 23002-7:2021 (Ed. 1) VSEI
    - ISO/IEC 23008-2:2020 (Ed. 4) HEVC
    - ISO/IEC 23090-3:2021 (Ed. 1) VVC
  + 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 2022-07-12. (Please see below for record of previously issued requests.)
    - ISO/IEC 23091-2:2021 (Ed. 2) Video CICP (was requested in April 2021, and the 2019 previous edition was also not made available there)
    - ISO/IEC 23008-2:2020 (Ed. 4) Amd.1:2021: Shutter interval information SEI message (has not been requested but separate publication may not be necessary if it is promptly included in a next edition)
    - ISO/IEC 14496-10:202X – AVC 10th edition – final text issued and public availability requested at the 25th meeting (January 2022)
    - ISO/IEC 23002-7:202X – VSEI 2nd edition – FDIS issued and public availability requested at the 25th meeting (January 2022)
    - ISO/IEC 23090-3:202X – VVC 2nd edition – FDIS issued and public availability requested at the 25th meeting (January 2022)
    - ISO/IEC 23090-15:202X – VVC conformance – FDIS issued and public availability requested at the 24th meeting (October 2021)
    - ISO/IEC 23090-16:202X – VVC reference software – FDIS issued and public availability requested at the 25th meeting (January 2022)
  + 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 left tbd until the next meeting – perhaps best to compile a list of all relevant software and conformance parts of AVC, HEVC, MPEG-2akaH.262?, CICP, and request in bulk)
* Draft standards progression status
  + New levels (from JVET-Z1005) – ISO/IEC 23008-2 DIS of new edition of HEVC was issued from 26th meeting, incorporating Amd.1 and corrigenda items (ballot closes 2023-01-10, expecting FDIS in January 2023); 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 draft, multiview profiles draft) which might potentially be included based on ballot comments, better defer ITU-T consent for July, as the aspects that could be consented from this meeting are not highly relevant. It was noted that the referencing of VSEI is also somewhat different in the ITU-T and ISO/IEC versions of HEVC and AVC, which might be aligned at the next convenient time (basically editorial).
  + New level and systems-related supplemental enhancement information (from JVET-AA2005) – VVC DAM was issued from 27th meeting, ballot to close 2023-01-03. JVET draft 4 seems not necessary (from the status known at opening plenary) The plan is to convert this into FDIS of new edition in January (provided that the 2nd edition would still be published in 2022 – check status). ITU-T consent would be premature, better wait for July 2023 when new edition of VSEI is consented, to keep VVC and VSEI aligned.
  + Additional SEI messages (from JVET-Z2006) – VSEI CDAM ballot comments in [m60678](https://dms.mpeg.expert/doc_end_user/current_document.php?id=84173&id_meeting=192). DAM (JVET draft 3) to be issued from current meeting. To be converted into FDIS of new edition in April, ITU-T consent in July.
  + Film grain synthesis technology for video applications (from JVET-AA2020 and JVET-AB0042) – JVET draft 3 to be issued at the current meeting, also ISO/IEC 23002-9 PDTR was planned to be issued in October 2022 (request made at 25th meeting) – discuss Friday
  + VVC Conformance testing for operation range extensions – (from JVET-Y2026) – original plan had been for ISO/IEC 23090-16 FDAM and ITU-T consent from current meeting – the progression plan on this was clarified during the AHG5 presentation: As the DAM ballot only closes by 2022-11-15, and it is unknown which comments would be made on some known issues, it would be premature to progress this to ITU consent.
  + Video CICP new edition with for YCgCo-Re and YCgCo-Ro (from JVET-Z1003) – JVET draft 3 to be issued at the current meeting (any changes?), CD ballot comments in [m60676](https://dms.mpeg.expert/doc_end_user/current_document.php?id=84171&id_meeting=192), ISO/IEC 23091-2 DIS to be issued, ITU-T consent from current meeting?
  + Ballot responses on final standards: FDIS 23090-16 [m60682](https://dms.mpeg.expert/doc_end_user/current_document.php?id=84177&id_meeting=192), FDIS 23002-7 (2nd ed.) [m60683](https://dms.mpeg.expert/doc_end_user/current_document.php?id=84178&id_meeting=192).
  + 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
* The meeting logistics, agenda, working practices, policies, and document allocation considerations were reviewed.
  + Remote access to the meeting was provided using Zoom. This requires discipline in the meeting room (no microphone to be switched on, podium and room microphones to be under central control).
  + Having text and software available is crucial (and not just arriving at the end of the meeting).
  + There were no objections voiced in the opening plenary to the consideration of late contributions.
* The results of the previous meeting and the meeting report JVET-AA1000 were reviewed. The following small issues in the meeting report were noted and were not considered sufficient to warrant issuing a revision. These are obviously left over from a previous report, and the correct information can be found in other places of the report:
  + The sentence “This case did not happen at this meeting” in the context of cross-verification reports not becoming available is wrong. Actually, JVET-AA0221 and JVET-AA0230 were never provided and later marked as withdrawn by the chair.
  + In section 2.4.3, the document numbers JVET-Y2023 and JVET-Y2024 should read as the corresponding “Z” numbers.
  + In section 2.12, the date 20 April should read as 13 July.
  + Regarding section 7.3, only one joint meeting was held with AG 5 and VCEG, and this meeting was held on Monday 18 July 1400-1500 UTC. The topics discussed related to 4.3 test conditions and 4.5 test material (see 2.13 scheduling, where the correct information had been included in the report)
  + Regarding section 7.4, only one BoG was held withn the topic of neural network-based video coding (see 5.2.1 and particularly JVET-AA0247 for further details).
  + In section 11, the agreed deadline is for the 28th meeting (not 27th).
* There was somewhat less of a problem of late non-cross-check documents.
* There were again a few documents registered where authors’ given names were not abbreviated, and/or company affiliation was missing in the authors’ list. Participants were reminded to stick to JVET’s conventions.
* Experts are asked to inform the chair when the title of a document is changed, or if authors are added. Otherwise, that might not be correct in the meeting notes.
* The primary goals of the meeting were
  + ITU-T deliverables, including errata
  + Conformance testing for version 2 of VVC (FDAM to be issued, also ITU-T deliverable)
  + New edition for HEVC (FDIS to be issued, also new edition H.265 as ITU-T deliverable)
  + New level and systems-related SEI for VVC (under DAM ballot)
  + Additional SEI in VSEI (DAM to be issued)
  + Preparation of TR for film grain (draft 3) (new version of WG 5 WD, PDTR in January)
  + New edition video CICP (DIS, also ITU-T deliverable?)
  + Any action items on reference software HM/VTM?
  + Additional colour type identifiers for AVC and HEVC (Draft 1 in JVET-Z1008 was not updated in last meeting) – could be included in new edition of H.265 (23008-2 DIS needs ballot comment on this), no action for AVC at this moment. It was suggested considering new editions of H.264 and 14496-10 in the near future.
  + Guide to IT systems used in JVET (email reflectors, guide to MPEG and JVET web sites, calendar, ITU-T ftp, ftp for test materials, software git, conformance repositories, ticket reporting system, etc. This will be developed by J. Ohm (general), K. Sühring (software), I. Moccagatta (conformance), M. Wien (test materials).
  + It was suggested to discuss public availability of experimental software packages – it was later discussed in the context of AHG6 report. It was decided that the ECM branch (not the EE branches) shall be made publicly available.
  + Exploration Experiments
    - Neural network-based video coding
    - Enhanced compression beyond VVC
* Liaison communication: the possibility of a liaison letter to JPEG was mentioned
* The number of documents was significantly higher than for the previous meeting (115->150)
* Scheduling was discussed, and it was agreed to avoid conducting “track” sessions in parallel (some BoG parallelism could occur)
  + Planning for the first two days: Only items related to current standard developments (ITU-T deliverables require early action until Wednesday Oct. 26), review of other AHGs and EEs to start on Friday.
  + The JVET plenary room is available only until 16:00 on Monday Oct. 24. It is therefore suggested to perform the ECM/VTM subjective viewing comparison partially on Monday afternoon, rather than doing all on Sunday.
  + Proposed scheduling: JVET hods a Sunday morning session 9-13, no JVET session Sunday afternoon; first part of ECM vs. VTM viewing sessions (12 volunteers needed) Sunday 13:30-16:30, second part (another 12 volunteers) Monday 16:30-19:30.
* Principles of standards development were discussed.
* Meeting plans need to be discussed – next F2F meeting in Antalya, April 2023.

## Scheduling of discussions

The times of meeting sessions will follow 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 will be 0900-1900 CEST with coffee breaks and lunch breaks as appropriate, however some early morning or late night sessions may be necessary. Sessions will be announced in the JVET calendar in advance as far as possible, but it might happen that some activities (such as breakout sessions) will be held at short notice.

Particular scheduling notes are shown below, although not necessarily 100% accurate or complete:

* Thu. 20 Oct., 1st day
  + Morning session:
    - 0900–1015 Opening remarks, review of practices, agenda, IPR reminder (section 2)
    - 1040–1155 Reports of AHGs 1–3, 5, 8, 9 (section 3)
    - 1155–1210 Review of section 4.2
    - 1210-1330 Review of NNPF (6.1)
  + Afternoon session:
    - 1500–1930 Review of NNPF (section 6.1)
* Fri. 21 Oct., 2nd day
  + Morning session:
    - 0900–0920 Additional opening remarks (if necessary, section 2)
    - 0920–1100 Reports of AHGs 4, 6, 7, 10-13 (section 3)
    - 1100–1300 Review of EE1 (section 5.2)
  + Afternoon session:
    - 1400–1915 Review of EE2 (section 5.3.1/5.3.2)
    - 1915–2000 Review of EE2 related (section 5.3.3)
    - 1400-1800 BoG on 6.1: NNPF (S. Deshpande)
* Sat. 22 Oct., 3rd day
  + Morning session:
    - 0900–1300 Review of EE1 & EE1 related (sections 5.2.2-5.2.3)
  + Afternoon session:
    - 1400–1640 Review of EE2 related (section 5.3.3)
    - 1700–2010 Review of other ECM (section 5.3.4)
* Sun. 23 Oct., 4th day
  + Morning session:
    - 0900–1300 Review of BoG on NNPF, and HLS topics (6.x)
  + Afternoon session:
    - 1400-1900 Review of EE1 related (section 5.2.3), and other NNVC (section 5.2.4) (chaired by A. Segall from 1800)
  + 1400-1730 ECM/VTM subjective viewing comparison (12 experts)
* Mon. 24 Oct., 5th day
  + 0900–1130 MPEG information sharing session
  + Afternoon session:
    - 1200–1330 JVET plenary: Coordination; Review of HLS (section 6.2, 6.4)
    - 1430–1600 Review of HLS (section 6.2/6.4/6.3)
  + 1615-1745 VCEG meeting (outside JVET)
  + 1630-2000 ECM/VTM subjective viewing comparison (12 experts)
  + 1900-2000 Joint WG 2/4/5 on VCM (outside JVET)
* Tue. 25 Oct., 6th day
  + Morning session:
    - 0900–1100 Review of documents on project development (section 4.x except 4.2, 4.4)
    - 1115-1230 Joint VCEG, JVET, WG 2/4 on VCM
  + Afternoon session:
    - 1330–1400 Continue review of documents on project development (section 4.x except 4.2, 4.4)
    - 1400–1530 Joint with AG 5 Review of visual testing (section 4.4)
    - 1530–2000 Review of remaining from ECM/EE2 category (JVET-AB0150, section 5.3.4) (chaired by Y. Ye after 1800)
* Wed. 26 Oct., 7th day
  + 0900–1000 MPEG information sharing session
  + Morning session:
    - 1030–1330 Remaining docs 4.7, 4.9, 4.10, 4.11, 4.13, review comments on CICP CD, further planning
  + 1430-1545 VCEG meeting (outside JVET)
  + Afternoon session:
    - 1600-XXXX BoG (with AG 5) on planning verification tests
    - 1600–1730 Document review 4.2, revisits, further planning
    - 1745 (at latest) Leave for social event
* Thu. 27 Oct., 8th day
  + Morning session:
    - 0900–1300 HLS revisits, review DoC, output document planning, EE planning, liaisons
  + 1400-1430 Joint with AG 5, WG 2, WG 4 on subjective tests ECM vs. VTM
  + Afternoon session:
    - 1500–1700 Plenary: Wrapup, output doc reviews/approvals, draft recommendations, AHG setup
* Fri. 28 Oct., 9th day
  + Plenary:
    - 0900–1200 Plenary: Wrapup, output doc reviews/approvals, AHG13/14 updates, draft recommendations, liaison updates, meeting planning, AoB
  + 1300–1500 MPEG information sharing session
  + XXXX–XXXX WG 5 Plenary: Approval of meeting recommendations, closing

## 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 may not be completely accurate):

* AHG reports (13) (section 3)
* Project development (section 4)
  + Deployment and advertisement of standards (2)
  + Text development and errata reporting (2)
  + Test conditions (2)
  + Subjective quality testing and verification testing (4)
  + Test Material (0)
  + Quality assessment (0)
  + Conformance test development (1)
  + Software development (6)
  + Implementation studies and complexity analysis (2)
  + AHG7: Low latency and constrained complexity (1)
  + Encoding algorithm optimization (3)
  + Profile/tier/level specification (0)
  + Proposed modification of system interface (0)
  + Use cases of standards related to specific applications (3)
* Low-level tool technology proposals (section 5) with subtopics (number counts excluding BoG and summary reports)
  + AHG8: High bit depth and high bit rate coding (0) (section 5.1)
  + AHG11 and EE1: Neural network-based video coding (31) (section 5.2)
  + AHG12 and EE2: Enhanced compression beyond VVC capability (72) (section 5.3)
* AHG9: High-level syntax (HLS) proposals (section 6) with subtopics
  + SEI messages on neural-network post filter (15) (section 6.1)
  + SEI messages on topics other than NNPF (9) (section 6.2)
  + Film grain synthesis (1) (section 6.3)
  + Non-SEI HLS aspects (2) (section 6.4)
* Joint meetings, plenary discussions, BoG reports (X), liaison (2 TBP), 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 (13)

These reports were discussed on Thursday 20 October 2022 during 1040–1155 CEST and on Friday 21 October 2022 during 0920–1100 (chaired by JRO).

[JVET-AB0001](https://jvet-experts.org/doc_end_user/current_document.php?id=11994) JVET AHG report: Project management (AHG1) [J.-R. Ohm (chair), G. J. Sullivan (vice-chair)]

The number of subscribers (by the beginning of the current meeting) was 1185. Furthermore, the previous lists of joint teams (which were still kept open as archives) had the following number of subscribers:

* JCT-VC – 1158 subscribers
* JCT-3V – 683 subscribers
* JVT-experts – 2063 subscribers

1. **Goals and activity**

The work of the JVET overall had proceeded well in the interim period with significantly 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 all output documents from the preceding meeting had been produced.

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/2022_07_AA_Virtual/>). It is noted that the previous document sites <http://phenix.int-evry.fr/jvet/>, <http://phenix.int-evry.fr/jct/>, and <http://phenix.int-evry.fr/jct3v/> are still accessible, but were converted to read-only.

The list of documents produced included the following, particularly:

* The meeting report (JVET-AA1000) [Posted 2022-08-18, also submitted as WG 5 N 140]
* Errata report items for VVC, HEVC, AVC, Video CICP, and CP usage TR (JVET-AA1004) [Posted 2022-09-14]
* HEVC multiview profiles supporting extended bit depth (Draft 1) (JVET-AA1011) [Posted 2022-10-19, also submitted as WG 5 WD N 143]
* Common Test Conditions for HM video coding experiments (JVET-AA1100) [Posted 2022-10-19]
* New level and systems-related supplemental enhancement information for VVC (Draft 3) (JVET-AA2005) [Posted 2022-07-27, last update 2022-08-08, also submitted as WG 5 DAM N 145]
* Additional SEI messages for VSEI (Draft 2) (JVET-AA2006) [Posted 2022-08-09, last update 2022-08-19, also submitted as WG 5 CDAM N 141]
* Common Test Conditions and evaluation procedures for neural network-based video coding technology (JVET-AA2016) [Posted 2022-08-06]
* Common Test Conditions for high bit depth and high bit rate video coding (JVET-AA2018) [Posted 2022-08-19]
* Film grain synthesis technology for video applications (Draft 2) (JVET-AA2020) [Posted 2022-09-30, last update 2022-10-01, also submitted as WG 5 WD N 142]
* Exploration experiment on Neural Network-based Video Coding (EE1) (JVET-AA2023) [Posted 2022-07-22, last update 2022-08-06, also submitted as WG 5 N 148]
* Exploration experiment on Enhanced Compression beyond VVC capability (EE2) (JVET-AA2024) [Posted 2022-07-222, also submitted as WG 5 N 149]
* Algorithm description of Enhanced Compression Model 6 (ECM 6) (JVET-AA2025) [Posted 2022-10-11, also submitted as WG 5 N 150]
* SEI processing order SEI message in VVC (Draft 1) (JVET-AA2025) [Posted 2022-08-05, also submitted as WG 5 WD N 146]

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

* Recommendations of the 8th WG 5 meeting (WG 5 N 139)
* Disposition of comments received on ISO/IEC 23090-3:202x (2nd Ed.) CDAM 1 (WG 5 N 144)
* Text of ISO/IEC CD 23091-2:202x Coding-independent code points - Part 2: Video (3rd edition) (WG 5 N 147)
* Request for ISO/IEC 23002-7:202x (2nd Ed.) Amd.1 Additional SEI messages (WG 5 N 125)
* List of AHGs established at the 8th WG 5 meeting (WG 5 N 151)
* Liaison statement to ISO/IEC JTC 1/SC 29/WG 1 (JPEG) on JPEG AI and NNVC (WG 5 N 152)

The thirteen *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 28th meeting had not been clear yet by the time of closing the last meeting, but it had been agreed to give the chair the discretion to clarify the situation and communicate via the JVET reflector as soon as possible. The meeting had originally been planned to be held under ITU-T SG16 auspices during Fri. 21 – Fri. 28 October 2022 in Antalya, TR, but that plan had to be cancelled due to difficulties of identifying a suitable meeting venue. ITU-T SG16 later decided to hold their meeting as face-to-face meeting one week earlier (during Mon. 17 – Fri. 28 October 2022) in Geneva, CH, but indicated there was not sufficient room space available in the ITU premises to also host JVET. Therefore, it was not possible to hold the 28th meeting as face-to-face meeting under ITU-T SG16 auspices, and SG16 management agreed that JVET could meet together with other MPEG WGs under SC 29 auspices in Mainz, DE. This also implied that JVET would not need to finish the meeting by Wed. 26 Oct. (before SG16 would start their final WP and SG plenaries); however, it would be necessary to approve and finalize documents to be submitted for ITU-T consent already by that date, such that starting the meeting one day earlier than (by Thu. 20 Oct.) was considered to be necessary. This was accordingly announced in the JVET reflector, in the JVET logistics document (<https://www.itu.int/wftp3/av-arch/jvet-site/2022_10_AB_Mainz/JVET-AB_Logistics.docx>), and in the WG 5 calling notice (N 153) and agenda (N 154) for the 9th WG 5 meeting. Remote participation was provided for experts who were unable to travel.

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 150 input contributions (not counting the AHG, CE and EE 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 27th meeting had been made publicly available on the ITU-hosted ftp site <http://wftp3.itu.int/av-arch/jvet-site/2022_10_AB_Mainz/>.

1. **Recommendations**

The AHG recommends its continuation.

[JVET-AB0002](https://jvet-experts.org/doc_end_user/current_document.php?id=11995) JVET AHG report: Draft text and test model algorithm description editing (AHG2) [B. Bross, C. Rosewarne (co-chairs), F. Bossen, J. Boyce, A. Browne, S. Kim, S. Liu, J. R. Ohm, G. J. Sullivan, A. Tourapis, Y.-K. Wang, Y. Ye (vice-chairs)]

2.1 Output documents produced

2.1.1 JVET-Z1003 Coding-independent code points for video signal type identification (Draft 1 of 3rd edition)

[CR: The document register has JVET-Z1003-v1, which I thought released before the July meeting, yet it appears in the mandate for the October meeting too?]

This document contains the draft text toward a 3rd edition of Coding-independent code points for video signal type identification (Rec. ITU-T H.273 | ISO/IEC 23091-2). Text modifications are provided for specification of code point identifiers for YCoCg-R colour representation with equal luma and chroma bit depths. The new code points are referred to as YCgCo-Re and YCgCo-Ro, where the number of bits added to a source RGB bit depth is 2 (i.e., even) and 1 (odd), respectively. Revision marking is provided to show modifications relative to the basis text (based on the 2021-07 edition of Rec. ITU-T H.273). Equation numbers and their cross-references that are maintained automatically have been updated without revision marking.

2.1.2 JVET-AA1004 Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR

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, while some other items have not yet been confirmed. This document also provides publication status backgrounds of these standards.

Incorporated items at the JVET-A meeting:

• For VVC: Added an errata item for VVC conformance.

• For VSEI: Added a pointer of some HEVC errata items that also apply for VSEI in Subsection 4.2.

• For HEVC: Removed various items that have been addressed during the editing of the DIS text of HEVC 5th edition in WG05 N00128, and added some editor's notes for the remaining items.

• For AVC:

o Added changes on motion vector value ranges, based on JVET meeting minutes noted under JVET-AA0222.

o Added a pointer of some HEVC errata items that also apply for AVC in Subsection 4.2.

2.1.3 JVET-AA1011 HEVC multiview profiles supporting extended bit depth (Draft 1)

This document contains the draft text for changes to the High Efficiency Video Coding (HEVC) standard (ITU T H.265 | ISO/IEC 23008-2), for the addition of new multiview profiles supporting extended bitdetph.

Draft 1 incorporated items:

• Addition of new multiview profiles supporting extended bitdetph (JVET-AA0239)

2.1.4 JVET-AA2005 New level and systems-related supplemental enhancement information for VVC (Draft 3)

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 3 incorporated items (among some very minor editorial changes that are not mentioned below):

• Removed the constraints on the values of BinCountsInPicNalUnits and BinCountsInSubpicNalUnits for the unlimited level (i.e., level 15.5) (JVET-AA0048 and US-001 in the DoCR in WG5 N0144).

• Replaced "subpicture level index" with "subpicture index" in subclause 7.4.2.4.5 (Order of VCL NAL units and their association to coded pictures) (JVET-AA0048 and bug ticket #1556).

• Removed the two rows in Table D.1 for the green metadata SEI message and the VDI envelop SEI message, because only those SEI messages that are actually specified in the VVC spec are listed in the table.

• Removed the subclauses on syntax and semantics of the green metadata SEI message and the VDI envelop SEI message, because only those SEI messages that are actually specified in the VVC spec have such subclauses. Like the SEI messages that are specified in the VSEI spec, when some VVC-specific usage of the SEI messages need to be specified in the VVC spec, then those are specified.

• Updated the description of the subclauses on use of VSEI SEI messages to include also the use of SEI messages that are specified in ISO/IEC 23001-11 and ISO/IEC 23090-13.

• Included, from JVET-Z2006, the VVC specific texts, including the texts specifying the use of the shutter interval information SEI message and the neural network post-filter characteristics SEI message.

• Corrected the NOTEs in subclauses A.3.1 and A.3.2 to reflect the possibility that a bitstream of a video profile can indicate the unlimited level (US-002 in the DoCR in WG5 N0144).

• Fixed tickets #1555 and #1560 (US-005 in the DoCR in WG5 N0144).

• Allocated/reallocated the payloadType values for some SEI messages, including replacing the value 205 for the scalability dimension information SEI message with the value 208 (US-011 in the DoCR in WG5 N0144).

• Fixed ticket #1558 (DE-03 and US-005 in the DoCR in WG5 N0144).

• Fixed the typo in the mis-spelled word "messag" in Table D.1.

2.1.5 JVET-AA2006 Additional SEI messages for VSEI (Draft 2)

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 characteristics SEI message, and phase indication SEI message.

2.1.6 JVET-AA2027 SEI processing order SEI message in VVC (Draft 1)

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.

3 Related input contributions

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

• JVET-AB0223 AHG2: Text improvement for Timing / DU information SEI message in HEVC and VVC

4 Remaining bug tickets

• #1564 Intra prediction ref pixel array bounds too small for wide angle

• #1568 Semantics correction for pps\_sao\_info\_in\_ph\_flag, pps\_alf\_info\_in\_ph\_flag, and pps\_wp\_info\_in\_ph\_flag

• #1569 Minor whitespace issues in coding unit syntax

• #1572 Sub clause C.1 -- Regarding number of bitstream conformance tests to be performed

5 Recommendations

The AHG recommends to:

• Approve JVET-Z1003, JVET-AA1004, JVET-AA1011, JVET-AA2005, JVET-AA2006, and JVET-AA2027. 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 AHG2 related contributions, bug tickets, and other AHG2 related inputs and act on them if found to be necessary.

It was noted that JVET-Z1003 was the document to be submitted as WG 5 CD of the new CICP edition.

[JVET-AB0003](https://jvet-experts.org/doc_end_user/current_document.php?id=11996) 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 18.0 (Sept. 2022)

• HM-16.26 (Oct. 2022)

• HM-16.21+SCM-8.8 (Mar. 2020)

• SHM 12.4 (Jan. 2018)

• HTM 16.3 (Jul. 2018)

• JM 19.0

• JSVM 9.19.15

• JMVC 8.5

• 3DV ATM 15.0 (no version history)

• HDRTools 0.23 (October 2021)

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.

2 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.

3 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 17.1 and 17.2 were tagged on Jul. 27, and VTM version 18.0 was tagged on Sept. 20. VTM 18.1 is expected during the 28th JVET meeting.

VTM 17.1 was tagged on Jul. 27, 2022. Changes include:

• JVET-Z0047: Improved film grain analysis

• Remove unnecessary mrgTypeNeighbours field

• Fix which ptl\_mulltlayer\_enabled\_flag is used

• Fix setNumOutputLayerSets() to accept value 256

• Use chromaFormatIDC stored inside Picture

• Fix: lintf undef

• Avoid floating-point operations when computing HAD

• Rename BcwIdx to bcwIdx, indentation fixes

• Fix #1557: change dependencies for writing ccv\_max\_luminance\_value and ccv\_avg\_luminance\_value

• Fix #1551: check coefficient range

• Avoid using std::vector and use static\_vector instead

• Avoid goto in SEIFilmGrainAnalyzer.cpp

• Reduce size of temporary memory for deblocking

• Precalculate fixed-point weights for MSE

• Y4M support at both encoder and decoder (#206)

• Consistently use helper functions and define them inline

• Don't use reference for simple bool parameters

• Clean up InterPrediction::xPredAffineBlk and associated functions

• Enable inlining of simple PelStorage methods

• Replace #define for constants in interpolation filter with constexpr

• Use constexpr instead of const for compile time constants

• Branchless version of Area::contains

• Clean up filter management for MC interpolation

• Clean up SIMD interpolation filter code and add support of horizontal 6-tap filter

• Fix CHECKD logic in prof

• Fix for PQ when EncDbOpt is used

• Fix population of m\_multiLayerOlsIdxToOlsIdx

• Fix lookup of vps\_ols\_dpb\_params\_idx

• Ticket #1561: Fix for userdata SEI parsing

• Add check whether writing to bitstream succeeded

• Fix for ticket #1559

• Fix for ticket #1562

• Fix copyRefPicList to copy ILRP entries

• Fix variable names, indentation and braces

VTM 17.2 was tagged on Jul. 27, 2022. Changes include:

• Remove macros from previous cycle

VTM 18.0 was tagged Sept. 20, 2022. Changes include:

• JVET-AA0056: AHG9: On syntax gating in the neural-network post-filter characteristics SEI message

• JVET-AA0100-fix: Fixing TypeDef guard issue for AA0100 and AA0056

• JVET-AA0067\_NNPFC\_SEI\_FIX

• JVET-A0110: Adding support for Phase indication SEI message

• JVET-AA0102 and JVET-AA2027: SEI Processing Order

• JVET-AA0054\_PROPOSAL2: Signal 1 flag instead of 2 flags to specify output chroma format information

• JVET-AA0055 Proposal A and B

• JVET-AA0054 Proposal 1: Signaling of external URI information in neural network post-filter characteristics SEI message

• Fix #1565: Invalid encoder memory access for EncDbOpt=1 and CHROMA\_400

• Use BitDepths stored inside Picture

• Fix #1566: run ISP related check only when in ISP mode

• Check layer IDs in addition to POCs

• Reduce compile time for VC

• Avoid "using namespace" in header files

• Simplify covariance computation in ALF by premultiplying constants

• Fix variable names and remove unnecessary check

• Use strong enum type for filter index

• Fix weighting when processing chroma components in ALF

• Clean up motion interpolation functions

• Fix compiling when enabling HDR tools

• Fix: RPR crash caused by picHeader creation/deletion

• Fix: Related to RPR and GDR

• Reset list of Prefix SEI NALUs if frame is skipped

• Fix #1571: tap number for vertical affine RPR

• Fix regression related to high bit depth motion compensation

• Add SIMD implementation of 6-tap filter for HBD motion compensation

• Fix compiling issues for GCC 12.1.0

• Fix Y4M picture size (considering conformance window)

• Fix GDR crash: Move deletion of picHeader from destroyCoeffs() to destroy()

• Fix: crash of Move deletion of picHeader from destroyCoeffs() to destroy()

VTM 18.1 is expected to be tagged during the 28th JVET meeting. Changes include so far:

• Fix memory leaks

• Add the missing setting of temporal filter to RA GOP16 cfg

• fix active picture number for L0 in reference picture list creation

• Fix #1574: SPS ID and PPS ID in encoding writeoutput()

• Add filler payload SEI message

• Fix RPL-based marking and RPL checks in Multilayer context

• Fix a Y4M bug (chroma scaling was not considered)

• fix reference picture lists checking for multilayer scalable

• Use bool constants and operators where appropriate

• Fix variable names

• Fix #1576: use correct interpolation filter size

• Fix DeblockingFilterDisable = 1

• Fix variable name and compile issue on Xcode

• Fix compile error with tracing is enabled

• Fix indentation and braces

• Fix indentation and braces

• Fix #1575: Use per-layer APS ids range for ALF APSs

• Avoid compile issue when using address sanitizer

• Remove JVET\_X0143\_ALF\_APS\_ID\_OFFSET related code

• Fix GDR code to avoid invalid reference

• Fix #1580: L1 RPL filling for incomplete GOP

• Fix: avoid redundant coding of RPL in both SPS and SH in multilayer

3.1 CTC Performance

The following tables shows VTM 18.0 performance over HM 16.26 for a fair comparison with aligned tool configurations.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **All Intra Main10** |  |  |
|  |  |  | **Over HM-16.26** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | -29.50% | -32.84% | -33.81% | 1316% | 166% |
| Class A2 | -29.73% | -24.40% | -21.58% | 2200% | 181% |
| Class B | -22.33% | -27.21% | -31.01% | 2414% | 182% |
| Class C | -22.90% | -19.56% | -23.23% | 3485% | 194% |
| Class E | -26.05% | -25.89% | -24.34% | 1946% | 168% |
| **Overall** | -25.50% | -25.76% | -27.06% | 2249% | 179% |
| Class D | -18.79% | -13.79% | -13.86% | 4061% | 177% |
| Class F | -39.48% | -40.18% | -42.90% | 4895% | 181% |
|  |  |  |  |  |  |
|  |  |  | **Random access Main10** |  |  |
|  |  |  | **Over HM-16.26** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | -40.60% | -40.73% | -47.09% | 570% | 160% |
| Class A2 | -44.00% | -41.70% | -40.79% | 683% | 174% |
| Class B | -37.44% | -50.14% | -48.49% | 654% | 158% |
| Class C | -33.90% | -36.28% | -38.29% | 903% | 168% |
| Class E |  |  |  |  |  |
| **Overall** | -38.44% | -42.88% | -43.95% | 700% | 164% |
| Class D | -31.72% | -32.70% | -31.97% | 1034% | 164% |
| Class F | -46.08% | -49.81% | -50.71% | 503% | 144% |
|  |  |  |  |  |  |
|  |  |  | **Low delay B Main10** |  |  |
|  |  |  | **Over HM-16.26** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -30.43% | -37.43% | -34.85% | 633% | 153% |
| Class C | -28.21% | -21.28% | -21.06% | 791% | 158% |
| Class E | -31.11% | -35.62% | -29.30% | 290% | 124% |
| **Overall** | -29.86% | -31.59% | -28.87% | 561% | 147% |
| Class D | -26.80% | -16.11% | -15.60% | 855% | 160% |
| Class F | -42.22% | -44.10% | -44.16% | 423% | 134% |
|  |  |  |  |  |  |
|  |  |  | **Low delay P Main10** |  |  |
|  |  |  | **Over HM-16.26** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -34.54% | -38.33% | -36.21% | 589% | 165% |
| Class C | -28.15% | -17.66% | -18.25% | 724% | 167% |
| Class E | -33.00% | -38.21% | -31.27% | 283% | 134% |
| **Overall** | -32.03% | -31.41% | -28.99% | 525% | 157% |
| Class D | -26.70% | -12.74% | -11.61% | 821% | 182% |
| Class F | -40.18% | -41.55% | -41.78% | 453% | 145% |

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

The following tables show **VTM 18.0** performance compared to **VTM 17.0** using SDR CTC:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **All Intra Main10** |  |  |
|  |  |  | **Over VTM-17.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | 0.00% | 0.00% | 0.00% | 93% | 98% |
| Class A2 | 0.00% | 0.00% | 0.00% | 97% | 98% |
| Class B | 0.00% | 0.00% | 0.00% | 96% | 100% |
| Class C | 0.00% | 0.00% | 0.00% | 97% | 102% |
| Class E | 0.00% | 0.00% | 0.00% | 95% | 99% |
| **Overall** | 0.00% | 0.00% | 0.00% | 96% | 100% |
| Class D | 0.00% | 0.00% | 0.00% | 98% | 104% |
| Class F | 0.00% | 0.00% | 0.00% | 100% | 100% |
|  |  |  |  |  |  |
|  |  |  | **Random access Main10** |  |  |
|  |  |  | **Over VTM-17.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | 0.00% | 0.00% | 0.00% | 91% | 99% |
| Class A2 | 0.00% | 0.00% | 0.00% | 91% | 98% |
| Class B | 0.00% | 0.00% | 0.00% | 90% | 97% |
| Class C | 0.00% | 0.00% | 0.00% | 92% | 99% |
| Class E |  |  |  |  |  |
| **Overall** | 0.00% | 0.00% | 0.00% | 91% | 98% |
| Class D | 0.00% | 0.00% | 0.00% | 92% | 98% |
| Class F | 0.00% | 0.00% | 0.00% | 92% | 99% |
|  |  |  |  |  |  |
|  |  |  | **Low delay B Main10** |  |  |
|  |  |  | **Over VTM-17.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | 0.00% | 0.00% | 0.00% | 91% | 96% |
| Class C | 0.00% | 0.00% | 0.00% | 90% | 93% |
| Class E | 0.00% | 0.00% | 0.00% | 86% | 94% |
| **Overall** | 0.00% | 0.00% | 0.00% | 89% | 95% |
| Class D | 0.00% | 0.00% | 0.00% | 89% | 95% |
| Class F | 0.00% | 0.00% | 0.00% | 89% | 98% |
|  |  |  |  |  |  |
|  |  |  | **Low delay P Main10** |  |  |
|  |  |  | **Over VTM-17.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | 0.00% | 0.00% | 0.00% | 90% | 99% |
| Class C | 0.00% | 0.00% | 0.00% | 90% | 95% |
| Class E | 0.00% | 0.00% | 0.00% | 81% | 92% |
| **Overall** | 0.00% | 0.00% | 0.00% | 88% | 96% |
| Class D | 0.00% | 0.00% | 0.00% | 90% | 99% |
| Class F | 0.00% | 0.00% | 0.00% | 89% | 98% |

The following tables show **VTM 18.0** performance compared to **VTM 17.0** using HDR CTC:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access** | | | | | | | | | |
|  | **Over VTM17.0** | | | | | | | | | |
|  |  |  | **wPSNR** |  |  | **PSNR** |  |  |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | -4.93% | -0.05% | -0.01% | -6.26% | -3.15% | 0.00% | 2.32% | 5.56% | 92% | 95% |
| Class H2 |  |  |  |  |  | 0.00% | 0.00% | 0.00% | 89% | 93% |
| **Overall** | -4.93% | -0.05% | -0.01% | -6.26% | -3.15% | 0.00% | 1.48% | 3.54% | 91% | 94% |
|  |  |  |  |  |  |  |  |  |  |  |
|  | **All Intra** | | | | | | | | | |
|  | **Over VTM17.0** | | | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | -0.54% | 0.00% | 0.01% | -1.01% | -0.22% | 0.01% | 0.29% | 1.03% | 96% | 96% |
| Class H2 |  |  |  |  |  | 0.00% | 0.00% | 0.00% | 97% | 97% |
| **Overall** | -0.54% | 0.00% | 0.01% | -1.01% | -0.22% | 0.01% | 0.19% | 0.66% | 96% | 96% |

Full results are attached to this AHG report as Excel files.

3.2 Issues in VTM affecting conformance

The following issues in VTM master branch may affect conformance:

• Missing HLS features (see sections below)

There are no known issues in VTM that affect processing of current VVC v1 and v2 conformance bitstreams. It should be noted that VTM version 18.0 fails on some v2 conformance bitstreams and that the related issue has been fixed for the upcoming VTM version 18.1.

3.3 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.

4 HM related activities

HM 16.26 was tagged on Oct. 13, 2022. Changes include:

• JVET-Y0155: Fixes for motion-compensated temporal prefilter

• JVET-Y0105: An improved VVC rate control scheme

• JVET-Y0077: Block Importance Mapping

• Silence compiler warning

• Update copyright date to include 2022

• fix ticket #1516

• Fix gcc-11.3 compiling error/warning

• Fix memory leak

The following tables show HM 16.25 performance compared to HM 16.26:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **All Intra Main10** |  |  |
|  |  |  | **Over HM-16.25** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | 0.00% | 0.00% | 0.00% | 101% | 101% |
| Class A2 | 0.00% | 0.00% | 0.00% | 99% | 99% |
| Class B | 0.00% | 0.00% | 0.00% | 100% | 102% |
| Class C | 0.00% | 0.00% | 0.00% | 101% | 98% |
| Class E | 0.00% | 0.00% | 0.00% | 98% | 100% |
| **Overall** | 0.00% | 0.00% | 0.00% | 100% | 100% |
| Class D | 0.00% | 0.00% | 0.00% | 100% | 102% |
| Class F | 0.00% | 0.00% | 0.00% | 99% | 97% |
|  |  |  |  |  |  |
|  |  |  | **Random access Main10** |  |  |
|  |  |  | **Over HM-16.25** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | -0.03% | 0.12% | -0.09% | 100% | 100% |
| Class A2 | 0.07% | 0.19% | 0.14% | 99% | 101% |
| Class B | -0.03% | 0.02% | -0.05% | 99% | 101% |
| Class C | -0.03% | 0.04% | -0.10% | 99% | 102% |
| Class E |  |  |  |  |  |
| **Overall** | -0.01% | 0.08% | -0.03% | 99% | 101% |
| Class D | -0.07% | 0.03% | -0.25% | 99% | 104% |
| Class F | 0.00% | 0.00% | 0.00% | 100% | 100% |
|  |  |  |  |  |  |
|  |  |  | **Low delay B Main10** |  |  |
|  |  |  | **Over HM-16.25** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -0.02% | -0.09% | 0.17% | 101% | 105% |
| Class C | 0.06% | 0.00% | -0.19% | 100% | 98% |
| Class E | -0.02% | -0.07% | -0.07% | 102% | 104% |
| **Overall** | 0.01% | -0.05% | -0.01% | 101% | 102% |
| Class D | -0.03% | -0.28% | 0.31% | 98% | 103% |
| Class F | 0.00% | 0.00% | 0.00% | 101% | 100% |
|  |  |  |  |  |  |
|  |  |  | **Low delay P Main10** |  |  |
|  |  |  | **Over HM-16.25** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | 0.02% | -0.16% | 0.33% | 101% | 105% |
| Class C | -0.02% | -0.10% | -0.05% | 101% | 102% |
| Class E | -0.02% | 0.29% | -0.06% | 100% | 100% |
| **Overall** | 0.00% | -0.03% | 0.11% | 101% | 103% |
| Class D | -0.01% | 0.34% | 0.34% | 95% | 97% |
| Class F | 0.00% | 0.00% | 0.00% | 101% | 99% |

The following MRs are pending [with status indicated]:

* JVET-X0048: implementation of film grain characteristics (FGC) SEI message [under review]
* Port the Y4M support [one issue remains]
* JVET-AA0130: VTM and HM common test conditions for high bit depth [testing]
* JCTVC-AD0021(JVET-T0056) SEI manifest & SEI prefix indication [issues raised]
* 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-16.26+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.

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:

* 38 tickets for “HM”, most of which are more than 5 years,
* 1 ticket for “HM RExt”, which was created during this reporting period,
* 7 tickets for “HM SCC”, all of which are at least 3 years old,

Help to address these tickets would be appreciated.

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

5 360Lib related activities

The latest 360Lib software (360Lib-13.3) can be found at https://vcgit.hhi.fraunhofer.de/jvet/360lib/-/tags/360Lib-13.3

The following table is for the projection formats comparison using VTM-18.0 according to 360-degree video CTC (JVET-U2012) compared to that using VTM-17.0 (VTM-17.0 as anchor).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **PERP: VTM-18.0 over VTM-17.0** | | | | | |
|  | **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-18.0 (PERP as anchor).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **GCMP Over PERP** | | | | | |
|  | **End-to-end WS-PSNR** | | | **End-to-end S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | -11.53% | -5.55% | -6.04% | -11.51% | -5.49% | -5.99% |
| Class S2 | -3.72% | 0.80% | 1.26% | -3.69% | 0.90% | 1.33% |
| **Overall** | -8.40% | -3.01% | -3.12% | -8.38% | -2.93% | -3.06% |

The following tables are for PERP and GCMP coding comparison between VTM-18.0 and HM-16.22 (HM as anchor), respectively.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **VTM-18.0 PERP Over HM-16.22 PERP (anchor)** | | | | | |
|  | **End-to-end WS-PSNR** | | | **End-to-end S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | -30.91% | -38.80% | -41.29% | -30.90% | -38.85% | -41.29% |
| Class S2 | -36.88% | -37.35% | -39.69% | -36.87% | -37.39% | -39.75% |
| **Overall** | -33.30% | -38.22% | -40.65% | -33.29% | -38.27% | -40.67% |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **VTM-18.0 GCMP Over HM-16.22 PCMP (anchor)** | | | | | |
|  | **End-to-end WS-PSNR** | | | **End-to-end S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | -35.61% | -40.88% | -42.79% | -35.57% | -40.82% | -42.74% |
| Class S2 | -39.84% | -39.56% | -41.51% | -39.85% | -39.56% | -41.55% |
| **Overall** | -37.30% | -40.35% | -42.28% | -37.28% | -40.31% | -42.27% |

6 SCM related activities

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

7 SHM related activities

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

8 HTM related activities

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

The next release will include the following changes:

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

9 HDRTools related activities

There had not been any updates of HDRTools.

New development is being added under the branch named 0.24-dev.

10 JM, JSVM, JMVM related activities

There had not been any updates to the JM, JSVM and JMVM software.

11 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.

12 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.

13 CTC alignment and merging

Following the merger for high-bit depth CTCs (JVET-AA0130), there are currently 6 related CTC documents, namely:

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

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.

It was reported that CTC documents may contain an error in the formulas describing PSNR computation. For example, JVET-Y2010 states:

“For 10-bit video, PSNR is calculated as 10\*(log10((255 << 2) / MSE), 8-bit content is converted to 10-bit input in the encoder by shifting 2 bits to the left and 10-bit PSNR calculation is used to report testing results.”

This should be corrected to:

“For 10-bit video, PSNR is calculated as 10\*(log10((255 << 2)2 / MSE), 8-bit content is converted to 10-bit input in the encoder by shifting 2 bits to the left and 10-bit PSNR calculation is used to report testing results.”

14 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

VTM 18 has various new implementations of SEI messages. Are all these sufficiently documented in the VTM description document JVET-Z2002 (VTM 17)? It was later confirmed that a new version of VTM document (VTM 18) shall be issued.

[JVET-AB0004](https://jvet-experts.org/doc_end_user/current_document.php?id=11997) JVET AHG report: Test material and visual assessment (AHG4) [V. Baroncini, T. Suzuki, M. Wien (co-chairs), S. Liu, G. Martin-Cocher, A. Segall, P. Topiwala, S. Wenger, J. Xu, Y. Ye (vice-chairs)]

2 Activities

2.1 Preparations for visual assessment of ECM vs. VTM

A total of three AhG meetings of AhGs4, 7, and 12 were held jointly with SC29/AG5 in preparation of a visual assessment comparing the compression efficienty of the ECM vs. the VTM. The meetings were held on 2022-08-16, 2022-09-20, and 2022-10-11. The reports of the three meetings are available in JVET-AB0041. Tests are proposed for UHD-RA, HD-RA, and HD-LD with rate-matched test points for ECM and VTM.

2.2 VVC Verification tests

An input contribution has been submitted proposing an initial draft of a verification test plan for VVC multilayer coding. The proposed draft is provided in document JVET-AB0123 and has a focus on scalability.

Furthermore, a proposal for a new verification test on the FGC SEI message has been registered as input to this meeting. The document number is JVET-AB0122.

2.3 Test sequences

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

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

3 Related contributions

|  |  |  |
| --- | --- | --- |
| [JVET-AB004](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0041-v3.zip)1 | AHG4, 7, 12: Report on AHG meetings on ECM performance evaluation preparation | [M. Wien (AHG chair)](mailto:wien@lfb.rwth-aachen.de) |
| [JVET-AB0122](https://jvet-experts.org/doc_end_user/current_document.php?id=12049) | Proposed FGC SEI message verification test draft plan | [P. de Lagrange (InterDigital)](mailto:philippe.delagrange@interdigital.com), [W. Husak (Dolby)](mailto:wjh@dolby.com) |
| [JVET-AB0123](https://jvet-experts.org/doc_end_user/current_document.php?id=12050) | Proposed multilayer VVC verification test draft plan | [P. de Lagrange (InterDigital)](mailto:philippe.delagrange@interdigital) |

Note: JVET-AB0041 is considered to be relevant for both, AHG4, 7 and 12.

4 Recommendations

The AHG recommends:

• To conduct the visual assessment as prepared and recommended in JVET-AB0041.

• To review and discuss the proposed verification test plan for VVC multilayer coding proposed in JVET-AB0123.

• To review and discuss the proposed verification test plan for the FGC SEI mssage proposed in JVET-AB0122.

• To collect volunteers to conduct further verification tests, including volunteers to encode.

• To continue to discuss and to update the non-finalized categories of the verification test plan, including those which have not been addressed yet.

• To review the set of available test sequences for the verification 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.

In the context of the AHG presentation, JVET-AB0041 was also presented, and the plans for the subjective comparison of ECM vs. VTM were confirmed.

It was reported that Tall buildings was the only sequences where the rate matching did not come to the suggested margin. It was suggested to keep it in the tests and decide later what can be concluded from the results.

Tests are planned to be conducted Sunday from 1400 and Monday after 1600. Volunteers shall contact Mathias, and then a selection will be done, preferably with good diversity between companies.

[JVET-AB0005](https://jvet-experts.org/doc_end_user/current_document.php?id=11998) JVET AHG report: Conformance testing (AHG5) [D. Rusanovskyy, I. Moccagatta (co-chairs), F. Bossen, K. Kawamura, T. Ikai, H.-J. Jhu, K. Sühring, Y. Yu (vice-chairs)]

2 Timeline

The progress on the Conformance testing specification is consistent with the preliminary timeline previously agreed, as follows:

• VVCv1 conformance:

o ISO/IEC FDIS 23090-15 issued from 2021-10 meeting, pending FDIS ballot

o H.266.1 V1 Consent 2022-01, last call to end 2022-04-28

• VVCv2 conformance:

o ISO/IEC 23090-15/Amd.1 CDAM: 2021-10

o ISO/IEC 23090-15/Amd.1 DAM: 2022-01

o ISO/IEC 23090-15/Amd.1 FDAM: 2022-07

o ISO/IEC 23090-15/Amd.1 AMD: 2023-01

o H.266.1 V2 Consent 2022-10

3 Status on bitstream submission

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

• conformance bitstreams for VVC:

o 104 bitstream categories have been identified

o At least one bitstream has been submitted in each identified category

o 283 total bitstreams have been provided, checked, and made available

o No changes between 27th and 28th meeting.

• conformance bitstreams for VVC operation range extensions:

o 57 bitstream categories have been identified

o 1 bitstream of 1 identified category has been re-generated

o 128 (was 127) bitstreams of 57 (was 56) identified categories have been cross-checked and uploaded.

4 Activities and Discussion

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

VVC activities:

As agreed during the 27th JVET meeting, the regeneration of HRD\_A\_Fujitsu\_3 to avoid CPB underflow has been reported in JVET-AA1004 in the “Errata items for VVC conformance” section. It has also been noted that, as per ISO request, a version of the regenerated package (HRD\_A\_Fujitsu\_4.zip) with removed company names from filenames and removed email addresses from text file descriptions will have to be generated.

There are not currently any known issues with the VVC conformance bitstream packages.

All provided bitstreams are decoded correctly using VTM18.0 w/o range extension support.

A compilation error on Linux when compiling VTM 18.0 / latest (SHA 89c013aa3ee3638cc5184a30b7db603ef326dd34) with range extension support was reported (https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1576). When using the debug version to decode VVC conformance streams mismatches were detected (https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1577 and https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1578 for operation range extension streams, respectively).

The compilation error was quickly fixed by MR https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware\_VTM/-/merge\_requests/2326. All provided bitstreams are decoded correctly using VTM18.0 and this MR with range extension support. The MR has been merged.

VVC operation range extensions activities:

One bitstream of 1 identified category (12b420SPvvc1\_A\_KDDI\_2.zip) has been regenerated following VTM16.0 changes.

All provided bitstreams are decoded correctly using VTM18.0 and the now merged MR https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware\_VTM/-/merge\_requests/2326.

The operation range extension streams will be copied in https://www.itu.int/wftp3/av-arch/jvet-site/bitstream\_exchange/VVCv2/under\_test/VTM-18.1 once VTM18.1 is tagged

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

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

5 Contributions

None.

6 Ftp site information

The procedure to exchange the bitstream (ftp cite, bitstream files, etc.) is specified in Sec 2 “Procedure” of JVET-R2008. 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

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)

7 Recommendations

The AHG recommends the following:

• Start the generation, cross-checking, and documentation of the conformance streams for the HEVC multiview profiles supporting extended bit depth (JVET-AA1011).

• Maintain and update the conformance bitstream database and contribute to report problems in JVET document 1004.

From the discussion about the AHG report:

It is noted that the V2 conformance is under ballot in ISO/IEC (closing November 15), and it is anticipated that corresponding comments would be made.

As no fixes to all potential problems noted above are available, it would be premature to submit the V2 conformance to ITU-T consent at this meeting.

[JVET-AB0006](https://jvet-experts.org/doc_end_user/current_document.php?id=11954) JVET AHG report: ECM software development (AHG6) [V. Seregin (chair), J. Chen, F. Le Léannec, K. Zhang (vice-chairs)]

2 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 fixes were integrated on top of ECM-5.0 and it is released as ECM-5.1:

Fix for valid number of affine candidates (MR 170)

Optimization for GDR software to lower runtime (MR 169)

Fix for reference picture reordering when DMVD if disabled (MR 168)

Fix for IBC macros with ARMC usage (MR 171)

Fix for GDR with IBC HMVP (MR 172)

Fix for pairwise candidates in GPM merge list (MR 173)

Fix for lintf under (MR 176)

Fix decoder crash when AML=0 and IBC=1 (MR 174)

Fix for storing and loading CABAC windows for temporal CABAC (MR 175)

Port the memory from VTM for EncDbOpt (MR 161)

Port y4m support from VTM (MR 181)

Fix for macro and config parameter difference for GPM split reordering (MR 180)

Fix initialization of adaptive QP map for BIM (MR 179)

Fix wrong L1 MVP in SMVD for MVD sign prediction (MR 178)

Encoder fix for IBC-TM when DMVD is disabled (MR 185)

Fix macro interaction for IBC BVD and IBC TM (MR 186)

Fix GPM intra encoder/decoder mismatch when sign prediction is disabled (MR 187)

Fix for MVP index bit count for SMVD when DMVD is disabled (MR 189)

Fix array out of bound for MVP cost when DMVD is disabled (MR 190)

The following adopted aspects were integrated into ECM-6.0:

JVET-AA0042: Longer RPR filters for luma and chroma (Test 2.10b) (MR 195)

JVET-AA0097: Fix for block-level OOB check (MR 191)

JVET-AA0146: Fixes on ECM for 360-degree video coding (MR 193)

JVET-AA0058\_Test2.7a: GPM adaptive blending (MR 194)

JVET-AA0093: EE2-2.6: Combination tests of Test 2.3, Test 2.4 and Test 2.5 (MR 197)

JVET-AA0107: RMVF affine merge candidates (MR 200)

JVET-AA0129: inter-hash RDO with OBMC-off (MR 206)

JVET-AA0124: AmvpMerge mode enabling check when RPR is enabled and DMVD is disabled (MR 162)

JVET-AA0057: CCCM (MR 192)

JVET-AA0128 test b: Increased number of CABAC contexts of affine merge index (MR 207)

JVET-AA0095: Longer luma filter shape and using samples before deblocking filter for ALF (MR 199)

JVET-AA0098: MTTdepth by Tid for class A random access with QP 22 (MR 201)

JVET-AA0126: GLM (MR 209)

JVET-AA0096: Motion compensated picture boundary padding (MR 202)

JVET-AA0133: Inter MTS optimization (MR 211)

JVET-AA0132: ECM software configuration parameters for template matching tools (MR 198)

JVET-AA0106: Adjust IBC reference area to 2x128 rows above the current CTU (MR 215)

JVET-AA0062: BVD for IBC with IBC flipping (MR 216)

CTC: Enable palette for SCC (MR 222)

Included fixes:

Initialize GOP list when GDR is enabled (MR 205)

Fix ref pic reordering overhead (MR 204)

Fix deblocking RDO chroma distortion calculation for HDR-PQ coding (MR 182)

Fix for chroma 400 format (MR 213)

Prevent empty merge candidate list generation for AmvpMerge (MR 184)

Fix for bilateral filter encoder (MR 210)

Set default values for BCW and HPEL interpolation filter (MR 225)

Fix crash and mismatch IBC with disabled AML (MR 224)

Fix creating coding structure for GDR (MR 227)

Fix conditional jumps on uninitialized variables (MR 233)

Fix uninitialized variables for SAO (MR 236)

Reduce compile time for VC (MR 237)

Fix MD5 error triggers with disabled AML (MR 239)

The following adopted aspects were integrated into VTM-11ecm6 anchor:

JVET-AA0098: MTTdepth by Tid for class A random access with QP 22 (MR 208)

JVET-AA0133: Enable inter MTS (MR 214)

CTC: Enable palette for SCC (MR 223)

Fix: deblocking RDO chroma distortion calculation for HDR-PQ coding (MR 212)

ECM-5.1 was tagged on July 25, 2022.

ECM-6.0 was tagged on September 4, 2022.

VTM-11ecm6 was tagged on September 4, 2022.

2.1 CTC Performance

In this section, ECM test results following ECM CTC configuration descried in JVET-Y2017 are summarized.

The below tables show ECM-5.1 performance over ECM-5.0 anchor.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | |
|  | **Over ECM-5.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | 0.00% | 0.00% | 0.00% | 104% | 104% |
| Class A2 | 0.00% | 0.00% | 0.00% | 103% | 101% |
| Class B | 0.00% | 0.00% | 0.00% | 100% | 96% |
| Class C | 0.00% | 0.00% | 0.00% | 100% | 98% |
| Class E | 0.00% | 0.00% | 0.00% | 101% | 96% |
| **Overall** | 0.00% | 0.00% | 0.00% | 101% | 98% |
| Class D | 0.00% | 0.00% | 0.00% | 101% | 97% |
| Class F | 0.00% | 0.00% | 0.00% | 99% | 99% |
| Class TGM | 0.00% | 0.00% | 0.00% | 98% | 99% |
|  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | |
|  | **Over ECM-5.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -0.01% | 0.10% | -0.07% | 102% | 103% |
| Class A2 | -0.03% | 0.02% | 0.08% | 101% | 101% |
| Class B | 0.00% | -0.10% | -0.08% | 99% | 99% |
| Class C | 0.00% | 0.04% | 0.02% | 98% | 99% |
| Class E |  |  |  |  |  |
| **Overall** | -0.01% | 0.00% | -0.02% | 100% | 100% |
| Class D | 0.00% | -0.11% | -0.12% | 98% | 98% |
| Class F | -0.01% | -0.06% | 0.02% | 98% | 99% |
| Class TGM | 0.06% | 0.06% | 0.05% | 98% | 99% |
|  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | |
|  | **Over ECM-5.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | 0.02% | -0.13% | -0.16% | 100% | 102% |
| Class C | 0.00% | -0.12% | -0.20% | 95% | 98% |
| Class E | -0.16% | 0.25% | 0.24% | 98% | 100% |
| **Overall** | -0.03% | -0.03% | -0.07% | 98% | 100% |
| Class D | 0.01% | 0.85% | -0.28% | 97% | 99% |
| Class F | -0.29% | 0.05% | -0.37% | 98% | 101% |
| Class TGM | -0.03% | -0.03% | 0.01% | 98% | 102% |
|  |  |  |  |  |  |
|  | **Low delay P Main 10** | | | | |
|  | **Over ECM-5.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -0.02% | -0.22% | -0.08% | 100% | 102% |
| Class C | -0.02% | 0.23% | -0.18% | 101% | 101% |
| Class E | -0.03% | -0.34% | -1.23% | 99% | 99% |
| **Overall** | -0.02% | -0.10% | -0.40% | 100% | 101% |
| Class D | 0.03% | -0.01% | -0.66% | 101% | 102% |
| Class F | 0.15% | -0.26% | -0.18% | 99% | 101% |
| Class TGM | 0.00% | 0.10% | 0.07% | 98% | 96% |

Next tables show ECM-6.0 performance over ECM-5.1 anchor.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | |
|  | **Over ECM-5.1** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -1.41% | -2.56% | -4.75% | 115% | 115% |
| Class A2 | -5.87% | -4.58% | -7.42% | 117% | 116% |
| Class B | -0.81% | -3.57% | -1.80% | 104% | 105% |
| Class C | -0.27% | -0.80% | -0.63% | 106% | 106% |
| Class E | -0.45% | -3.35% | -1.17% | 104% | 105% |
| **Overall** | -1.57% | -2.92% | -2.86% | 108% | 109% |
| Class D | -0.16% | -0.07% | -0.36% | 105% | 107% |
| Class F | -4.00% | -6.49% | -6.17% | 108% | 105% |
| Class TGM | -12.50% | -16.87% | -16.28% | 102% | 93% |
|  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | |
|  | **Over ECM-5.1** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -2.30% | -3.06% | -4.85% | 115% | 120% |
| Class A2 | -3.45% | -2.93% | -4.57% | 122% | 130% |
| Class B | -1.68% | -4.10% | -2.80% | 120% | 113% |
| Class C | -1.60% | -0.96% | -1.31% | 115% | 114% |
| Class E |  |  |  |  |  |
| **Overall** | -2.13% | -2.82% | -3.17% | 118% | 118% |
| Class D | -1.70% | -0.93% | -0.51% | 116% | 116% |
| Class F | -5.00% | -5.97% | -6.32% | 113% | 109% |
| Class TGM | -10.25% | -10.47% | -10.46% | 108% | 113% |
|  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | |
|  | **Over ECM-5.1** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -1.79% | -2.35% | -2.39% | 116% | 130% |
| Class C | -1.92% | -1.11% | -1.51% | 117% | 126% |
| Class E | -0.98% | -4.53% | -2.90% | 118% | 121% |
| **Overall** | -1.63% | -2.48% | -2.23% | 117% | 126% |
| Class D | -2.20% | -1.55% | -1.79% | 121% | 126% |
| Class F | -4.19% | -4.60% | -4.17% | 116% | 114% |
| Class TGM | -7.89% | -7.06% | -7.21% | 110% | 106% |
|  |  |  |  |  |  |
|  | **Low delay P Main 10** | | | | |
|  | **Over ECM-5.1** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -1.47% | -1.50% | -1.00% | 116% | 129% |
| Class C | -1.32% | -0.69% | -0.36% | 120% | 124% |
| Class E | -1.01% | -3.07% | -1.33% | 114% | 117% |
| **Overall** | -1.31% | -1.62% | -0.87% | 117% | 124% |
| Class D | -1.43% | -0.44% | -0.87% | 120% | 123% |
| Class F | -3.40% | -3.64% | -3.31% | 116% | 114% |
| Class TGM | -4.76% | -4.65% | -4.71% | 108% | 110% |

The below tables show ECM-6.0 performance comparing to VTM-11.0ecm6 anchor.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | |
|  | **Over VTM-11.0ecm6** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -8.15% | -16.36% | -23.24% | 490% | 338% |
| Class A2 | -11.93% | -19.18% | -22.12% | 468% | 311% |
| Class B | -6.68% | -18.88% | -17.18% | 413% | 285% |
| Class C | -7.36% | -11.49% | -11.91% | 414% | 277% |
| Class E | -8.32% | -16.40% | -15.40% | 388% | 300% |
| **Overall** | -8.22% | -16.45% | -17.55% | 430% | 298% |
| Class D | -6.07% | -9.36% | -9.16% | 413% | 298% |
| Class F | -17.83% | -24.91% | -24.43% | 336% | 299% |
| Class TGM | -30.28% | -35.62% | -34.91% | 303% | 304% |
|  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | |
|  | **Over VTM-11.0ecm6** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -18.76% | -21.33% | -28.13% | 526% | 652% |
| Class A2 | -20.59% | -26.86% | -29.77% | 544% | 884% |
| Class B | -16.88% | -26.97% | -24.74% | 449% | 678% |
| Class C | -18.75% | -21.90% | -21.83% | 427% | 692% |
| Class E |  |  |  |  |  |
| **Overall** | -18.50% | -24.47% | -25.65% | 475% | 713% |
| Class D | -19.78% | -23.23% | -23.42% | 426% | 747% |
| Class F | -21.21% | -27.21% | -27.56% | 384% | 442% |
| Class TGM | -29.30% | -33.62% | -34.08% | 417% | 356% |
|  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | |
|  | **Over VTM-11.0ecm6** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -15.49% | -29.59% | -28.03% | 390% | 595% |
| Class C | -16.89% | -24.94% | -25.01% | 396% | 580% |
| Class E | -14.04% | -22.57% | -22.67% | 390% | 390% |
| **Overall** | -15.59% | -26.29% | -25.68% | 392% | 531% |
| Class D | -19.04% | -26.06% | -26.25% | 392% | 669% |
| Class F | -19.92% | -27.91% | -26.96% | 384% | 391% |
| Class TGM | -28.62% | -34.50% | -34.38% | 401% | 324% |
|  |  |  |  |  |  |
|  | **Low delay P Main 10** | | | | |
|  | **Over VTM-11.0ecm6** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -15.02% | -29.61% | -27.54% | 371% | 580% |
| Class C | -16.58% | -24.65% | -23.84% | 343% | 571% |
| Class E | -14.04% | -21.98% | -22.82% | 363% | 392% |
| **Overall** | -15.30% | -26.05% | -25.13% | 359% | 523% |
| Class D | -19.39% | -26.08% | -26.11% | 333% | 621% |
| Class F | -19.18% | -27.86% | -26.86% | 383% | 391% |
| Class TGM | -25.55% | -32.36% | -32.30% | 423% | 327% |

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

3 ECM repository access

To access ECM software and EE2 repositories, a GitLab account shall be converted to internal. To convert an account to internal, the request should be send to VTM software coordinators as detailed in VTM Wiki https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware\_VTM/-/wikis/VVC-Software-Development-Workflow. Internal account provides read access, ability to submit merge requests, and issues.

Alternatively, ECM can be accessed using VCEG or MPEG shared accounts.

4 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.

It was decided that the ECM branch (not the EE branches) shall be made publicly available.

[JVET-AB0007](https://jvet-experts.org/doc_end_user/current_document.php?id=11999) JVET AHG report: Low latency and constrained complexity (AHG7) [A. Duenas, T. Poirier, S. Liu (co-chairs), L. Wang, J. Xu (vice-chairs)]

2 Related Contributions

JVET-AB0171 AHG7: Asymmetric Deblocking at Virtual Boundaries [S. Hong, L. Wang, K. Panusopone (Nokia)]

This contribution proposes asymmetric deblocking at virtual boundaries, where deblocking is still disabled for pixels on the refreshed area sides of virtual boundaries, but will be performed for pixels on the non-refreshed area sides of virtual boundaries. In addition, refreshed-area pixels of virtual boundaries are used to compensate the output of deblocking of corresponding non-refreshed area pixels.

3 Recommendations

The AHG recommends reviewing input contributions and:

• to study the impact of the asymmetric filtering across virtual boundaries and the option of enabling GDR for low latency and controlled complexity.

[JVET-AB0008](https://jvet-experts.org/doc_end_user/current_document.php?id=12000) JVET AHG report: High bit depth, high bit rate, and high frame rate coding (AHG8) [A. Browne, T. Ikai (co-chairs), D. Rusanovskyy, X. Xiu, Y. Yu (vice-chairs)]

2 Activities

Previously, the AHG has used the main JVET reflector, jvet@lists.rwth-aachen.de, with [AHG8] in message headers, however no correspondence marked as AHG8 was sent between the 27th and 28th meetings.

The major area of work of the AHG in this meeting cycle was JVET-AA2018, the combined VTM/HM CTC for high bit depths.

3 Contributions

There are no contributions relevant to the study of high bit depth, high bit rate or high frame rate coding have been registered for the 28th meeting.

4 Recommendations

The AHG recommends the following:

• To continue high bit depth, high bit rate, and high frame rate studies to support conformance activities for the next meeting cycle.

[JVET-AB0009](https://jvet-experts.org/doc_end_user/current_document.php?id=12001) JVET AHG report: SEI message studies (AHG9) [S. McCarthy, Y.-K. Wang (co-chairs), T. Chujoh, S. Deshpande, C. Fogg, P. de Lagrange, G. J. Sullivan, A. Tourapis, S. Wenger (vice-chairs)]

2 Related contributions

A total of 24 contributions are identified relating to the mandates of AHG9. One contribution relates to more than one mandate. The number of contributions relating to each mandate is as follows

• 20 contributions relate to the mandate to study the SEI messages in VSEI, VVC, HEVC, and AVC;

o 13 contributions relate to the neural-network post-filter SEI messages, including 1 summary of comments and proposals on NNPF SEI messages.

o 2 contributions relate to the SEI processing order SEI message

o 3 contributions relate to the DRAP and ERAP SEI messages

o 1 contribution relates to the film grain characteristics SEI message

o 1 contribution relates to the picture timing and DU information SEI messages

• 2 contribution relates to the mandate to collect software and showcase information for SEI messages;

• 3 contributions relate to the mandate to identify potential needs for additional SEI messages.

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

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

2.1.1 Neural-network post filter characteristics and activation SEI messages

2.1.1.1 Summary of comments and proposals on NNPF SEI messages

JVET-AB0193 AHG9: A summary of proposals on NNPF SEI messages [Y.-K. Wang (Bytedance)]

2.1.1.2 Proposals on NNPF SEI messages

JVET-AB0046 AHG9: On StrengthControlVal of the NNPFC SEI message [M. M. Hannuksela, M. `Santamaria, F. Cricri (Nokia)]

JVET-AB0047 AHG9: nnpfc\_mode\_idc related changes to the NNPFC SEI message [M. M. Hannuksela, F. Cricri, M. Santamaria (Nokia)]

JVET-AB0049 AHG9: Miscellaneous aspects of the two neural-network post-filtering SEI messages [Y.-K. Wang, Y. Li, C. Lin, J. Li, K. Zhang, L. Zhang (Bytedance)]

JVET-AB0050 AHG9: Activation of a neural-network post-processing filter for multiple pictures [Y.-K. Wang, K. Zhang, L. Zhang, C. Lin, J. Li, Y. Li (Bytedance)]

JVET-AB0058 AHG9: Frame Rate Upsampling Information in Neural-network Post-filter Characteristics SEI Message [S. Deshpande, A. Sidiya (Sharp)]

JVET-AB0059 AHG9: Comments on Neural-network Post-filter Characteristics SEI Message [S. Deshpande (Sharp)]

JVET-AB0060 AHG9: On activation of the neural-network post-filter characteristics SEI message [T. Chujoh, Y. Yasugi, T. Ikai (Sharp)]

JVET-AB0074 AHG9: Auxiliary input for neural-network post-processing filter [Y. Li, J. Li, C. Lin, K. Zhang, L. Zhang, Y.-K. Wang (Bytedance)]

JVET-AB0075 AHG9: Separate processing of chroma components for neural-network post-processing filter [Y. Li, J. Li, C. Lin, K. Zhang, L. Zhang, Y.-K. Wang (Bytedance)]

JVET-AB0134 AHG9: On NN post-filter activation SEI message [M. Pettersson, R. Sjöberg, J. Ström (Ericsson)]

JVET-AB0135 AHG9: On complexity metrics for NN post-filter characteristics SEI message [M. Pettersson, R. Sjöberg, J. Ström (Ericsson)]

JVET-AB0152 AHG9: Regional on/off control and selection of NNPFs [J. Li, C. Lin, K. Zhang, L. Zhang, Y.-K Wang, Y. Li (Bytedance)]

2.1.2 SEI processing order SEI message

JVET-AB0051 AHG9: On the SEI processing order SEI message [Y.-K. Wang (Bytedance), Hendry (LGE)]

JVET-AB0069 AHG9: On the SEI processing order SEI message [Y. He, M. Coban, M. Karczewicz (Qualcomm)]

2.1.3 DRAP and EDRAP SEI messages

JVET-AB0055 AHG9: On leading pictures design in DRAP SEI Message [Hendry, S. Kim (LGE)]

JVET-AB0056 AHG9: On leading pictures design in EDRAP SEI Message [Hendry, S. Kim (LGE)]

(JVET-AB0056 also relates to the mandate to identify potential needs for additional SEI messages)

K. Suehring, B. Bross (Fraunhofer HHI), Y.-K. Wang (Bytedance)JVET-AB0057 AHG9: On the associated IRAP for DRAP and EDRAP pictures [Hendry, S. Kim (LGE)]

2.1.4 Film grain characteristics SEI message

JVET-AB0122 Proposed FGC SEI message verification test draft plan [P. de Lagrange (InterDigital), W. Husak (Dolby)]

2.1.5 Picture timing and DU information SEI messages

JVET-AB0223 AHG2: Text improvement for Timing / DU information SEI message in HEVC and VVC [K. Suehring, B. Bross (Fraunhofer HHI), Y.-K. Wang (Bytedance)]

2.2 Software and showcase information for SEI messages

JVET-AB0048 EE1-1.1: Content-adaptive post-filter with SADL inference and signalling of NN post-filter characteristics and activation SEI messages [M. Santamaria, R. Yang, F. Cricri, J. Lainema, H. Zhang, R. G. Youvalari, M. M. Hannuksela (Nokia)]

JVET-AB0072 VTM Encoder Implementation for Green-MPEG SEI Messaging [Christian Herglotz, Matthias Kränzler, André Kaup (FAU)]

2.3 Identify potential needs for additional SEI messages

JVET-AB0056 AHG9: On leading pictures design in EDRAP SEI Message [Hendry, S. Kim (LGE)]

(JVET-AB0056 also relates to the mandate to study the SEI messages in VSEI, VVC, HEVC and AVC)

JVET-AB0070 AHG9: On inclusion of post-filter hint SEI message into VSEI [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

JVET-AB0096 AHG9: Resolution Change Information SEI message [V. Drugeon, K. Abe, T. Toma (Panasonic)]

3 Activities

The regular JVET e-mail reflector was used for discussions (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. There was one email to the JVET reflector summarizing contributions related to NNPF SEI messages.

4 Recommendations

The AHG recommends to:

• Review all related contributions; and

• Continue SEI messages studies.

[JVET-AB0010](https://jvet-experts.org/doc_end_user/current_document.php?id=12002) JVET AHG report: Encoding algorithm optimization (AHG10) [P. de Lagrange, A. Duenas, R. Sjöberg, A. Tourapis (AHG chairs)]

2 Related contributions

A total of 6 contributions, not including cross-checks, are identified relating to AHG10, and summarized in the following sections.

2.1 Spatial scalability

2.1.1 JVET-AB0045 – AHG10: Study of VVC spatial scalability performance

This contribution reports about VVC test results in spatial scalability mode with two layers, following work reported previously in JVET-X0202, JVET-Y0047 and JVET-Y0048.

For 2x scaling ratio, performance impact of layer bitrate allocation (through QP offset) was investigated. Performance using optimal balance was measured using objective metrics, and visually with a formal assessment, using 6 sequences, most of them not in the CTCs. Gains compared to single-layer coding are reported; one some cases significant visual gains are reported on the full MOS range, including the higher end.

Results of preliminary tests using 3x scaling ratio are reported in addition.

2.1.2 JVET-AB0210 – AHG3/10: VTM multilayer profile encoder fixes

This contribution reports about recent fixes after VTM-18.0 impacting multilayer coding performance: two fixes reduce bitrate related to reference picture list signaling, that was present in both SPS and slice header. Another one is related to APS id space management, allocating reserved space for base layer. Overall coding gains of about 0.2 % (base layer) to 0.3% (enhancement layer) are reported on classes A and B.

2.2 Adaptive resolution

2.2.1 JVET-AB0080 – AHG10: GOP-based RPR encoder control

Following the work reported in JVET-W0082, and JVET-Y0068, this contribution proposes an encoder algorithm, implemented on top of VTM-18.0, to decide coding a GOP at reduced resolution (either 4/5, 2/3, or 1/2) based on the QP and upscaled PSNR of the first picture of the GOP (which is tested at all candidate resolutions). Pictures are coded at reduced resolution with a fixed QP offset depending on the scaling ratio to align the bitrates. Gains are reported compared to full-resolution coding using the CTCs (0.44% luma gain in RA, with around 1.3% chroma loss).

Compared to previous proposal, the 4/5 resolution is added and decision for reduced-resolution coding is less aggressive.

2.2.2 JVET-AB0081 – AHG10: Increased length of filters used for upscaling reconstructed pictures

This contribution proposes to use 12-tap interpolation filters instead of 8-tap to upscale the reconstructed picture when coding at reduced resolution. These longer filters are different from the ECM ones. Performance gain in RPR CTC LDB is around 1%. Visual quality is reported to be similar.

2.2.3 JVET-AB0082 – AHG12: Fixes for RPR in ECM

This contribution reports about fixes of RPR filters usage at encoder side in ECM (from 8-tap to 12-tap), for ME and output picture upscale, that have a significant impact on coding performance (up to 2% in luma using RPR CTC LDB).

2.2.4 JVET-AB0102 – AHG11/EE1-related: Updates on RPR encoder and filters

This contribution proposes to use a scheme similar to the combination of JVET-AB0080 and JVET-AB0081 as a secondary anchor for super-resolution experiments in EE1.

3 Recommendation

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

[JVET-AB0011](https://jvet-experts.org/doc_end_user/current_document.php?id=12003) JVET AHG report: Neural network-based video coding (AHG11) [E. Alshina, S. Liu, A. Segall (co chairs), F. Galpin, Y. Li, H. Wang, L. Wang, Z. Wang, M. Wien, P. Wu (vice chairs)]

2.1 Common Test Conditions

2.1.1 Document

The AHG released revised common test conditions as decided at the 27th meeting. The final version was uploaded as document JVET-AA2016 on August 6, 2022.

2.1.2 Anchor Software

Software for the revised common conditions was made available on the Git repository used for the AHG activity. A full version of the anchor software was tagged as VTM-11\_nnvc-2.0 on August 4, 2022. The software is available at https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/VVCSoftware\_VTM/-/tree/VTM-11.0\_nnvc-2.0. Additionally, a patch file capturing the changes to the CTC was provided on August 5, 2022 and available at https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/nnvc-ctc/-/tree/master/Software%20Patches

2.1.3 Anchor Encoding

Anchors for the NN-based video coding activity were released on September 5, 2022. The anchors were also made available on the Git repository used for the AHG activity: https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/nnvc-ctc/-/tree/master.

2.1.4 Additional Comments

The major change in the common conditions was to enable the EncDbOpt for the RA, LDB and LDP configurations. This was done to follow the group decision to enable EncDbOpt, and the NNVC CTC follows the same approach as what is done in the VTM and ECM CTCs.

Following the release of the CTC, the AhG further discussed if EncDbOpt should have been enabled for AI. It is unclear from the notes at the last meeting if this was intended, and the AhG decided to not re-issue the CTC. Having said that, the AhG would recommend that the group discuss if EncDbOpt should be enabled for the AI configuration during the 28th meeting. This could include the NNVC and/or other CTCs. Additionally, different configurations of the candidate software (discussed in Section 2.4) take different approaches to enabling the EncDbOpt parameter. So guidance from the group would be helpful for those configurations as well.

2.2 EE Coordination

The AHG finalized, conducted and discussed the EE on NN based video coding. The final version of the EE description was uploaded to the document repository on August 6, 2022.

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

JVET-AB0023 EE1: Summary of Exploration Experiments on Neural Network-based Video Coding E. Alshina, F. Galpin, Y. Li, M. Santamaria, H. Wang, L. Wang, Z. Xie

2.3 Study and Maintain SADL

The SADL library was updated during the AHG period. Updates were provided on October 13, 2022. The software is available at https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/sadl.

Change logs include:

Commit cd3f79eb6b8d8c2d5b6fd0715c14db7fd5d9c7b7

Changes:

- add conv2d 5x5 (support for one case for auto encoder)

- increase maximum total size of tensor

- better output for info on model

- clean python code

2.4 Create Common Software

The AhG created a common software base according to the decisions in the 27th meeting. The software is tentatively called the Neural Compression Software (NCS).

A first release candidate for the NCS was announced on August 24, 2022 and available at https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/VVCSoftware\_VTM/-/tree/NCS-rc1.0. It included the following features:

• Common API for data dumping (VTM side) and loading (python side)

• Common API for inference (VTM side)

• Common part for SPS etc.

• Two neural network-based filter sets:

o set 0 (corresponding to AA0088)

o set 1 (corresponding to AA0111)

• Training scripts for each set, including the stages:

o Data dumping

o Dataset generation

o Training (from scratch or from already generated models)

o Conversion to SADL format

An official candidate was announced on September 5, 2022, and available at https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/VVCSoftware\_VTM/-/tree/NCS-1.0.

2.4.1 Anchor Encoding

The coding performance of the common software as input to the meeting in document JVET-AB0176. As reported, the performance of NCS-1.0 NN-based filter set #0 (int16 precision) performance compared to the NNVC-2.0 anchor is

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -9.15% | -13.89% | -18.71% | 153% | 73018% |
| Class A2 | -9.37% | -19.11% | -14.46% | 149% | 70078% |
| Class B | -8.33% | -18.68% | -20.42% | 156% | 70892% |
| Class C | -8.37% | -20.30% | -20.35% | 143% | 60876% |
| Class E |  |  |  |  |  |
| **Overall** | -8.71% | -18.24% | -18.87% | 151% | 68316% |
| Class D | -9.58% | -19.42% | -20.42% | 139% | 58850% |
| Class F | -3.72% | -11.90% | -10.92% | 197% | 32349% |
|  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -7.32% | -17.59% | -20.15% | 144% | 69084% |
| Class C | -7.86% | -20.18% | -20.40% | 131% | 55098% |
| Class E | -8.43% | -17.14% | -16.50% | 213% | 54715% |
| **Overall** | -7.78% | -18.74% | -20.26% | 154% | 60438% |
| Class D | -9.21% | -18.73% | -19.41% | 132% | 51121% |
| Class F | -4.17% | -10.03% | -8.79% | 185% | 32820% |
|  |  |  |  |  |  |
|  | **All Intra Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -6.12% | -14.26% | -17.26% | 211% | 50641% |
| Class A2 | -5.73% | -15.81% | -12.98% | 159% | 41067% |
| Class B | -5.93% | -15.54% | -17.46% | 149% | 38489% |
| Class C | -6.40% | -16.42% | -18.16% | 133% | 27723% |
| Class E | -8.83% | -15.25% | -16.12% | 167% | 44744% |
| **Overall** | -6.52% | -15.52% | -16.61% | 159% | 38826% |
| Class D | -6.41% | -15.30% | -18.56% | 126% | 25806% |
| Class F | -4.02% | -11.81% | -11.71% | 128% | 32540% |

The performance of the NCS-1.0 NN-based filter set #0 (float precision) over the NNVC-2.0 anchor is:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -9.21% | -14.66% | -18.35% | 170% | 97724% |
| Class A2 | -9.37% | -19.29% | -14.32% | 163% | 93013% |
| Class B | -8.31% | -19.02% | -20.33% | 172% | 93877% |
| Class C | -8.34% | -20.27% | -20.09% | 159% | 80518% |
| Class E |  |  |  |  |  |
| **Overall** | -8.71% | -18.53% | -18.67% | 166% | 90671% |
| Class D | -9.48% | -19.52% | -20.41% | 162% | 77552% |
| Class F | -3.73% | -11.83% | -11.00% | 229% | 42242% |
|  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -7.31% | -18.18% | -20.28% | 161% | 91967% |
| Class C | -7.85% | -19.93% | -20.16% | 142% | 73330% |
| Class E | -8.48% | -16.46% | -16.42% | 250% | 72758% |
| **Overall** | -7.78% | -18.96% | -20.23% | 172% | 80428% |
| Class D | -9.22% | -18.82% | -19.35% | 148% | 66622% |
| Class F | -4.23% | -9.99% | -8.98% | 211% | 42845% |
|  |  |  |  |  |  |
|  | **All Intra Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -6.13% | -14.24% | -17.26% | 251% | 67748% |
| Class A2 | -5.75% | -15.93% | -12.99% | 176% | 54763% |
| Class B | -5.93% | -15.56% | -17.45% | 166% | 51137% |
| Class C | -6.38% | -16.47% | -18.12% | 148% | 35566% |
| Class E | -8.81% | -15.27% | -16.32% | 180% | 58517% |
| **Overall** | -6.51% | -15.55% | -16.64% | 177% | 51141% |
| Class D | -6.40% | -15.48% | -18.58% | 141% | 32964% |
| Class F | -4.04% | -11.89% | -11.74% | 137% | 44279% |

The performance of the NCS-1.0 NN-based filter set #1 (int16 precision) over the NNVC-2.0 anchor is:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -8.92% | -15.94% | -18.51% | 210% | 52791% |
| Class A2 | -10.15% | -20.20% | -17.29% | 200% | 50531% |
| Class B | -9.01% | -23.08% | -21.77% | 209% | 51646% |
| Class C | -9.82% | -21.78% | -22.31% | 195% | 50009% |
| Class E |  |  |  |  |  |
| **Overall** | -9.44% | -20.73% | -20.37% | 203% | 51205% |
| Class D | -11.43% | -23.49% | -23.88% | 196% | 47037% |
| Class F | -4.69% | -11.70% | -10.66% | 298% | 25719% |
|  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -7.80% | -16.00% | -14.02% | 196% | 51463% |
| Class C | -9.14% | -15.17% | -14.89% | 171% | 45228% |
| Class E | -8.63% | -17.25% | -17.03% | 344% | 43518% |
| **Overall** | -8.45% | -15.63% | -14.40% | 216% | 47271% |
| Class D | -10.78% | -18.72% | -17.99% | 174% | 39565% |
| Class F | -4.97% | -8.98% | -6.08% | 272% | 23963% |
|  |  |  |  |  |  |
|  | **All Intra Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -6.45% | -17.96% | -19.88% | 212% | 41577% |
| Class A2 | -6.50% | -20.46% | -17.82% | 164% | 33736% |
| Class B | -6.52% | -21.87% | -21.59% | 157% | 31884% |
| Class C | -7.40% | -18.86% | -21.76% | 153% | 23335% |
| Class E | -9.87% | -20.85% | -20.66% | 171% | 36317% |
| **Overall** | -7.26% | -20.14% | -20.56% | 168% | 32075% |
| Class D | -7.32% | -19.50% | -21.39% | 153% | 20201% |
| Class F | -4.06% | -11.76% | -10.93% | 134% | 21676% |

The performance of the NCS-1.0 NN-based filter set #1 (float precision) over the NNVC-2.0 anchor is:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -8.86% | -16.12% | -18.83% | 257% | 80229% |
| Class A2 | -10.16% | -20.28% | -17.34% | 246% | 77940% |
| Class B | -8.99% | -22.95% | -21.83% | 255% | 78170% |
| Class C | -9.83% | -21.70% | -22.38% | 226% | 71605% |
| Class E |  |  |  |  |  |
| **Overall** | -9.42% | -20.72% | -20.48% | 245% | 76716% |
| Class D | -11.45% | -23.40% | -23.80% | 231% | 70983% |
| Class F | -4.64% | -11.65% | -10.59% | 363% | 37408% |
|  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -7.75% | -15.97% | -13.64% | 244% | 76605% |
| Class C | -9.14% | -15.24% | -15.19% | 201% | 68109% |
| Class E | -8.52% | -17.06% | -17.80% | 458% | 65923% |
| **Overall** | -8.41% | -15.65% | -14.33% | 268% | 70947% |
| Class D | -10.76% | -18.20% | -17.94% | 200% | 58824% |
| Class F | -4.97% | -9.04% | -5.93% | 351% | 35888% |
|  |  |  |  |  |  |
|  | **All Intra Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -6.42% | -17.97% | -19.74% | 247% | 64929% |
| Class A2 | -6.48% | -20.48% | -17.81% | 182% | 52976% |
| Class B | -6.49% | -21.90% | -21.61% | 174% | 49418% |
| Class C | -7.42% | -18.84% | -21.71% | 156% | 33740% |
| Class E | -9.87% | -20.81% | -20.61% | 188% | 56011% |
| **Overall** | -7.25% | -20.15% | -20.52% | 184% | 49081% |
| Class D | -7.35% | -19.43% | -21.33% | 149% | 32772% |
| Class F | -4.08% | -11.73% | -10.89% | 141% | 33401% |

***Technical Evaluation***

The AHG made meaningful progress on the mandate to evaluate and quantify potential NN based video coding technologies. A summary of AHG11 related non-EE1 contributions provided as input to the 28th meeting is provided below:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Title** | **Common Test Conditions** | **Results** | | | **Training Data** | |
| **RA** | **LDB** | **AI** | **CTC** | **Additional** |
| **Loop Filter** | | | | | | | |
| [JVET-AB0108](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12035) | AHG11: ALF-SPLIT for NCS-1.0 | No | No | No | Yes | Unknown | - |
| [JVET-AB0136](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12063) | AHG11: Complexity Reduction on Neural-Network Loop Filter | No | Yes | No | No | BVI-DVC |  |
| [JVET-AB0158](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12085) | [AHG11] On chroma order adjustment in NNLF | No | Yes | Yes | No | Unknown |  |
| [JVET-AB0159](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12086) | [AHG11] On adjustment of residual for NNLF | No | Yes | Yes | No | Unknown |  |
| [JVET-AB0160](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12087) | [AHG11] Combination of chroma order adjustment and residual adjustment for NNLF | No | Yes | Yes | No | Unknown |  |
| **Post Filtering** | | | | | | | |
| [JVET-AB0109](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12036) | AHG11: Content-adaptive NN post-filter with new QP normalisation | No | No | No | No | - | JVET-CTC |
| [*JVET-AB0046*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11961) | *AHG9: On StrengthControlVal of the NNPFC SEI message* | No | No | No | No | - | - |
| [*JVET-AB0047*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11962) | *AHG9: nnpfc\_mode\_idc related changes to the NNPFC SEI message* | No | No | No | No | - | - |
| [*JVET-AB0049*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11964) | *AHG9: Miscellaneous aspects of the two neural-network post-filtering SEI messages* | No | No | No | No | - | - |
| [*JVET-AB0050*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11965) | *AHG9: Activation of a neural-network post-processing filter for multiple pictures* | No | No | No | No | - | - |
| [*JVET-AB0058*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11973) | *AHG9: Frame Rate Upsampling Information in Neural-network Post-filter Characteristics SEI Message* | No | No | No | No | - | - |
| [*JVET-AB0059*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11974) | *AHG9: Comments on Neural-network Post-filter Characteristics SEI Message* | No | No | No | No | - | - |
| [*JVET-AB0060*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11975) | *AHG9: On activation of the neural-network post-filter characteristics SEI message* | No | No | No | No | - | - |
| [*JVET-AB0074*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11989) | *AHG9: Auxiliary input for neural-network post-processing filter* | No | No | No | No | - | - |
| [*JVET-AB0075*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11990) | *AHG9: Separate processing of chroma components for neural-network post-processing filter* | No | No | No | No | - | - |
| [*JVET-AB0134*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12061) | *AHG9: On NN post-filter activation SEI message* | No | No | No | No | - | - |
| [*JVET-AB0135*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12062) | *AHG9: On complexity metrics for NN post-filter characteristics SEI message* | No | No | No | No | - | - |
| [*JVET-AB0152*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12079) | *AHG9: Regional on/off control and selection of NNPFs* | No | No | No | No | - | - |
| [*JVET-AB0193*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12120) | *AHG9: A summary of proposals on NNPF SEI messages* | No | No | No | No | - | - |
| **Super-Resolution** | | | | | | | |
| [JVET-AB0107](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12034) | Non-EE1: CNN-based super resolution with luma-only rescaling | No | No | No | No | BVI-DVC | DIV2K |
| [JVET-AB0101](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12028) | AHG11: Lightweight CNN filter for RPR-based SR with Wavelet Decomposition | No | No | No | No | BVI-DVC | DIV2K |
| [JVET-AB0102](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12029) | AHG11/EE1-related: Updates on RPR encoder and filters | Yes | Yes | No | Yes | Unknown | - |
| **Inter-Prediction** | | | | | | | |
| [JVET-AB0114](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12041) | AHG11: Deep Reference Frame Generation for Inter Prediction Enhancement | Yes | Yes | Yes | No | Unknown | - |
| [JVET-AB0121](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12048) | AHG11: Assistant Reference Picture Method for NNVC | No | No | No | No | Unknown |  |
| **Intra Prediction** | | | | | | | |
| [JVET-AB0149](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12076) | Non-EE1: neural network-based intra prediction with learned mapping to VVC intra prediction modes | No | Yes | No | Yes | BVI-DVC, TVD, UVG | DIV2K, ILSVCR2012 |

1. **Input contributions**

There are 60 input contriubtions related to the AHG mandates. Twenty-seven of the contributions are part of the EE activity, while the remaining 33 contributions are related to AHG11 but not part of the EE. The list of input contributions is provided below.

***EE and Related Input Contributions***

|  |  |  |
| --- | --- | --- |
| **Reporting** | | |
| [JVET-AB0023](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12151) | EE1: Summary of Exploration Experiments on Neural Network-based Video Coding | [E. Alshina](mailto:elena.alshina@huawei.com), [F. Galpin](mailto:Franck.Galpin@InterDigital.com), [Y. Li](mailto:yue.li@bytedance.com%20%20), [M. Santamaria](mailto:maria.santamaria_gomez@nokia.com), [H.Wang](mailto:hongtaow@qti.qualcomm.com), [L.Wang](mailto:liqiangwang@tencent.com%20), [Z.Xie](mailto:xiezhihuang@oppo.com) |
| **EE Technology** | | |
| [JVET-AB0048](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11963) | EE1-1.1: Content-adaptive post-filter with SADL inference and signalling of NN post-filter characteristics and activation SEI messages | [M. Santamaria](mailto:maria.santamaria_gomez@nokia.com), R. Yang, F. Cricri, J. Lainema, H. Zhang, R. G. Youvalari, M. M. Hannuksela (Nokia) |
| [JVET-AB0052](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11967) | EE1-1.5: One luma model with IPB and/or skip for filtering intra and inter luma slices | [D. Liu](mailto:du.liu@ericsson.com), [J. Ström](mailto:jacob.strom@ericsson.com), [M. Damghanian](mailto:mitra.damghanian@ericsson.com), [P. Wennersten](mailto:per.wennersten@ericsson.com), [K. Andersson (Ericsson)](mailto:kenneth.r.andersson@ericsson.com) |
| [JVET-AB0053](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11968) | EE1-1.2: NN intra model without attention and partitioning strength | [J. Ström](mailto:jacob.strom@ericsson.com), [D. Liu](mailto:du.liu@ericsson.com), [K. Andersson](mailto:kenneth.r.andersson@ericsson.com), [P. Wennersten](mailto:per.wennersten@ericsson.com), [M. Damghanian](mailto:mitra.damghanian@ericsson.com), [R. Yu (Ericsson)](mailto:ruoyang.yu@ericsson.com) |
| [JVET-AB0054](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11969) | EE1-1.3: CNN Based In-Loop Filter with WCDANN | [H. Zhang](mailto:13227706628@163.com), [C. Jung (Xidian Univ.)](mailto:zhengzk@xidian.edu.cn), [D. Zou](mailto:zoudan@oppo.com), [M. Li (OPPO)](mailto:myron.li@oppo.com) |
| [JVET-AB0068](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11983) | EE1-1.6: RDO Considering Deep In-Loop Filtering | [J. Li](mailto:lijunru@bytedance.com), [Y.Li](mailto:yue.li@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com) |
| [JVET-AB0073](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11988) | EE1-1.4: Deep In-Loop Filter with Additional Input Information | [Y. Li](mailto:yue.li@bytedance.com), [K. Zhang](mailto:zhangkai.vide@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com) |
| [JVET-AB0076](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11991) | EE1-2.1: RPR-Based Super-Resolution Guided by Partition Information | [Q. Han](mailto:hanqihui2013@163.com), [C. Jung (Xidian Univ.)](mailto:zhengzk@xidian.edu.cn), [Y. Liu](mailto:serena@oppo.com), [M. Li (OPPO)](mailto:myron.li@oppo.com) |
| [JVET-AB0077](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11992) | EE1-2.2: CNN Filter for Super-Resolution with RPR functionality in VVC | [S. Huang](mailto:shimin_huang2022@163.com), [C. Jung (Xidian Univ.)](mailto:zhengzk@xidian.edu.cn), [Y. Liu](mailto:serena@oppo.com), [M. Li (OPPO)](mailto:myron.li@oppo.com) |
| [JVET-AB0083](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12010) | EE1-1.8: More refinements on NN based in-loop filter with a single model | [L. Wang](mailto:liqiangwang@tencent.com), [X. Xu](mailto:xiaozhongxu@tencent.com), [S. Liu (Tencent)](mailto:shanl@tencent.com), [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-AB0084](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12011) | EE1-2.3：A CNN-based Super Resolution Method with GOP Level Adaptive Resolution | [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), [J. Nam](mailto:junghak.nam@lge.com), S. Yoo, J. Lim, S. Kim (LGE) |
| [JVET-AB0164](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12091) | EE1-1.7: Capacity Ablation of CNN-based in-loop filtering | [S. Eadie](mailto:seadie@qti.qualcomm.com), [H. Wang](mailto:hongtaow@qti.qualcomm.com), [M. Coban](mailto:mcoban@qti.qualcomm.com), [M. Karczewicz (Qualcomm)](mailto:martak@qti.qualcomm.com) |
| **EE Related** | | |
| [JVET-AB0090](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12017) | EE1-1.3 related: Lightweight and Efficient CNN In-loop Filter | [H. Zhang](mailto:13227706628@163.com), [C. Jung (Xidian Univ.)](mailto:zhengzk@xidian.edu.cn), [Y. Liu](mailto:serena@oppo.com), [M. Li (OPPO)](mailto:myron.li@oppo.com) |
| [JVET-AB0093](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12020) | EE1-2.2 related: Lightweight CNN Filter for Super-Resolution with RPR functionality in VVC | [S. Huang](mailto:shimin_huang2022@163.com), [C. Jung (Xidian Univ.)](mailto:zhengzk@xidian.edu.cn), [Y. Liu](mailto:serena@oppo.com), [M. Li (OPPO)](mailto:myron.li@oppo.com) |
| [JVET-AB0098](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12025) | EE1-2.3 related: GOP Level Adaptive Resampling with CNN-based Super Resolution | [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-AB0141](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12068) | EE1-related: QP-based loss function design for NN-based in-loop filter | [C. Zhou](mailto:chuan.zhou@vivo.com), [Z. Lv](mailto:zhuoyi.lv@vivo.com), [J. Zhang (vivo)](mailto:jirong.zhang@vivo.com), W. Chen, J. Guo, B. Ai (BJTU) |
| [JVET-AB0146](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12073) | EE1-1.8-related: encoder-only optimization for NN based in-loop filter with a single model | [L. Wang](mailto:liqiangwang@tencent.com), [X. Xu](mailto:xiaozhongxu@tencent.com), [S. Liu (Tencent)](mailto:shanl@tencent.com) |
| [JVET-AB0147](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12074) | EE1-1.8-related: using additional models for higher temporal layers | [L. Wang](mailto:liqiangwang@tencent.com), [X. Xu](mailto:xiaozhongxu@tencent.com), [S. Liu (Tencent)](mailto:shanl@tencent.com) |
| [JVET-AB0179](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12106) | EE1-related: Deep In-Loop Filter with Wide Activation and Large Receptive Field | [Y. Li](mailto:yue.li@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com) |
| **Cross Checks** | | |
| [JVET-AB0063](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11978) | Cross-check of JVET-AB0053 (EE1-1.2: NN intra model without attention and partitioning strength) | [M. Santamaria](mailto:maria.santamaria_gomez@nokia.com), F. Cricri (Nokia) |
| [JVET-AB0064](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11979) | Cross-check of JVET-AB0054 (EE1-1.3: CNN Based In-Loop Filter with WCDANN) | [M. Santamaria](mailto:maria.santamaria_gomez@nokia.com), F. Cricri (Nokia) |
| [JVET-AB0088](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12015) | Crosscheck of JVET-AB0083 (EE1-1.8: More refinements on NN based in-loop filter with a single model) | [D. Liu (Ericsson)](mailto:du.liu@ericsson.com) |
| [JVET-AB0089](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12016) | Crosscheck of JVET-AB0052 (EE1-1.5: One luma model with IPB and/or skip for filtering intra and inter luma slices) | [L. Wang (Tencent)](mailto:liqiangwang@tencent.com) |
| [JVET-AB0097](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12024) | Crosscheck of JVET-AB0084 (EE1-2.3：A CNN-based Super Resolution Method with GOP Level Adaptive Resolution) | [D. Liu (Ericsson)](mailto:du.liu@ericsson.com) |
| [JVET-AB0105](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12032) | Crosscheck of JVET-AB0076 (EE1-2.1: RPR-Based Super-Resolution Guided by Partition Information) | [R. Chang (Tencent)](mailto:renjiechang@tencent.com) |
| [JVET-AB0106](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12033) | Crosscheck of JVET-AB0077 (EE1-2.2: CNN Filter for Super-Resolution with RPR functionality in VVC) | [R. Chang (Tencent)](mailto:renjiechang@tencent.com) |
| [JVET-AB0137](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12064) | Crosscheck of JVET-AB0068 (EE1-1.6: RDO Considering Deep In-Loop Filtering) | [J. Ström (Ericsson)](mailto:jacob.strom@ericsson.com) |

***Non-EE Input Contributions***

|  |  |  |
| --- | --- | --- |
| **Reporting** | | |
| [JVET-AB0011](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12003) | JVET AHG report: Neural network-based video coding (AHG11) | E. Alshina, S. Liu, A. Segall (co chairs), F. Galpin, Y. Li, H. Wang, L. Wang, Z. Wang, M. Wien, P. Wu (vice chairs) |
| **Loop Filtering** | | |
| [JVET-AB0108](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12035) | AHG11: ALF-SPLIT for NCS-1.0 | [W.Zou](mailto:wjzou@xidian.edu.cn), [Y.Zhou](mailto:yi.zhou@stu.xidian.edu.cn), [C.M.Gu (Xidian University)](mailto:cmgu@stu.xidian.edu.cn), [C.Huang](mailto:huang.cheng5@zte.com.cn), [Y.X.Bai](mailto:bai.yaxian@zte.com.cn), [Y.J. Zhang(ZTE Corporation)](mailto:zhang.yuanjian@zte.com.cn) |
| [JVET-AB0136](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12063) | AHG11: Complexity Reduction on Neural-Network Loop Filter | [J. N. Shingala](mailto:jay.shingala@ittiam.com), S. Kadaramandalgi, A. Shyam (Ittiam), T. Shao, A. Arora, [P. Yin](mailto:pyin@dolby.com), Sean McCarthy (Dolby) |
| [JVET-AB0158](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12085) | [AHG11] On chroma order adjustment in 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-AB0159](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12086) | [AHG11] On adjustment of residual 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-AB0160](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12087) | [AHG11] Combination of chroma order adjustment and residual adjustment 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) |
| **Post Filtering** | | |
| [JVET-AB0109](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12036) | AHG11: Content-adaptive NN post-filter with new QP normalisation | [M. Santamaria](mailto:maria.santamaria_gomez@nokia.com), F. Cricri, M. M. Hannuksela (Nokia) |
| [*JVET-AB0046*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11961) | *AHG9: On StrengthControlVal of the NNPFC SEI message* | [*M. M. Hannuksela*](mailto:miska.hannuksela@nokia.com)*, M. Santamaria, F. Cricri (Nokia)* |
| [*JVET-AB0047*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11962) | *AHG9: nnpfc\_mode\_idc related changes to the NNPFC SEI message* | [*M. M. Hannuksela*](mailto:miska.hannuksela@nokia.com)*, F. Cricri, M. Santamaria (Nokia)* |
| [*JVET-AB0049*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11964) | *AHG9: Miscellaneous aspects of the two neural-network post-filtering SEI messages* | [*Y.-K. Wang*](mailto:yekui.wang@bytedance.com)*,* [*Y. Li*](mailto:yue.li@bytedance.com)*,* [*C. Lin*](mailto:linchaoyi.cy@bytedance.com)*,* [*J. Li*](mailto:lijunru@bytedance.com)*,* [*K. Zhang*](mailto:zhangkai.video@bytedance.com)*,* [*L. Zhang (Bytedance)*](mailto:lizhang.idm@bytedance.com) |
| [*JVET-AB0050*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11965) | *AHG9: Activation of a neural-network post-processing filter for multiple pictures* | [*Y.-K. Wang*](mailto:yekui.wang@bytedance.com)*,* [*K. Zhang*](mailto:zhangkai.video@bytedance.com)*,* [*L. Zhang*](mailto:lizhang.idm@bytedance.com)*,* [*C. Lin*](mailto:linchaoyi.cy@bytedance.com)*,* [*J. Li*](mailto:lijunru@bytedance.com)*,* [*Y. Li (Bytedance)*](mailto:yue.li@bytedance.com) |
| [*JVET-AB0058*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11973) | *AHG9: Frame Rate Upsampling Information in Neural-network Post-filter Characteristics SEI Message* | [*S. Deshpande*](mailto:sdeshpande@sharplabs.com)*, A. Sidiya (Sharp)* |
| [*JVET-AB0059*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11974) | *AHG9: Comments on Neural-network Post-filter Characteristics SEI Message* | [*S. Deshpande (Sharp)*](mailto:sdeshpande@sharplabs.com) |
| [*JVET-AB0060*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11975) | *AHG9: On activation of the neural-network post-filter characteristics SEI message* | [*T. Chujoh*](mailto:chujoh.takeshi@sharp.co.jp)*, Y. Yasugi, T. Ikai (Sharp)* |
| [*JVET-AB0074*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11989) | *AHG9: Auxiliary input for neural-network post-processing filter* | [*Y. Li*](mailto:yue.li@bytedance.com)*,* [*J. Li*](mailto:lijunru@bytedance.com)*,* [*C. Lin*](mailto:linchaoyi.cy@bytedance.com)*,* [*K. Zhang*](mailto:zhangkai.vide@bytedance.com)*,* [*L. Zhang*](mailto:lizhang.idm@bytedance.com)*,* [*Y.-K. Wang (Bytedance)*](mailto:yekui.wang@bytedance.com) |
| [*JVET-AB0075*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=11990) | *AHG9: Separate processing of chroma components for neural-network post-processing filter* | [*Y. Li*](mailto:yue.li@bytedance.com)*,* [*J. Li*](mailto:lijunru@bytedance.com)*,* [*C. Lin*](mailto:linchaoyi.cy@bytedance.com)*,* [*K. Zhang*](mailto:zhangkai.vide@bytedance.com)*,* [*L. Zhang*](mailto:lizhang.idm@bytedance.com)*,* [*Y.-K. Wang (Bytedance)*](mailto:yekui.wang@bytedance.com) |
| [*JVET-AB0134*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12061) | *AHG9: On NN post-filter activation SEI message* | [*M. Pettersson*](mailto:martin.m.pettersson@ericsson.com)*, R. Sjöberg, J. Ström (Ericsson)* |
| [*JVET-AB0135*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12062) | *AHG9: On complexity metrics for NN post-filter characteristics SEI message* | [*M. Pettersson*](mailto:martin.m.pettersson@ericsson.com)*, R. Sjöberg, J. Ström (Ericsson)* |
| [*JVET-AB0152*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12079) | *AHG9: Regional on/off control and selection of NNPFs* | [*J. Li*](mailto:lijunru@bytedance.com)*,* [*C. Lin*](mailto:linchaoyi.cy@bytedance.com)*,* [*K. Zhang*](mailto:zhangkai.video@bytedance.com)*,* [*L. Zhang*](mailto:lizhang.idm@bytedance.com)*,* [*Y.-K Wang*](mailto:yekui.wang@bytedance.com)*,* [*Y. Li (Bytedance)*](mailto:yue.li@bytedance.com) |
| [*JVET-AB0193*](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12120) | *AHG9: A summary of proposals on NNPF SEI messages* | [*Y.-K. Wang (Bytedance)*](mailto:yekui.wang@bytedance.com) |
| **Super Resolution** | | |
| [JVET-AB0107](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12034) | Non-EE1: CNN-based super resolution with luma-only rescaling | [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-AB0101](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12028) | AHG11: Lightweight CNN filter for RPR-based SR with Wavelet Decomposition | [H. Lan](mailto:lanhui_xidian@163.com), [C. Jung (Xidian Univ.)](mailto:zhengzk@xidian.edu.cn), [Y. Liu](mailto:serena@oppo.com), [M. Li (OPPO)](mailto:myron.li@oppo.com) |
| [JVET-AB0102](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12029) | AHG11/EE1-related: Updates on RPR encoder and filters | [J. Nam](mailto:junghak.nam@lge.com), [S. Yoo](mailto:sunmi.yoo@lge.com), [J. Lim](mailto:jaehyun.lim@lge.com), [S. Kim (LGE)](mailto:seunghwan3.kim@lge.com) |
| **Intra Prediction** | | |
| [JVET-AB0149](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12076) | Non-EE1: neural network-based intra prediction with learned mapping to VVC intra prediction modes | [T.Dumas](mailto:thierry.dumas@interdigital.com), F.Galpin, P.Bordes (InterDigital) |
| **Inter Prediction** | | |
| [JVET-AB0114](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12041) | AHG11: Deep Reference Frame Generation for Inter Prediction Enhancement | [J. Jia](mailto:jiajh2021@whu.edu.cn), [Y. Zhang](mailto:yuantongzhang@whu.edu.cn), [H. Zhu](mailto:zhuhanlyx@whu.edu.cn), [Z. Chen (Wuhan Univ.)](mailto:zzchen@whu.edu.cn), [Z. Liu](mailto:zizhengliu@tencent.com), [L. Wang](mailto:liqiangwang@tencent.com), [X. Xu](mailto:xiaozhongxu@tencent.com), [S. Liu (Tencent)](mailto:shanl@tencent.com) |
| [JVET-AB0121](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12048) | AHG11: Assistant Reference Picture Method for NNVC | [C. M. Gu](mailto:cmgu@stu.xidian.edu.cn), [W. Zou](mailto:wjzou@xidian.edu.cn), [Y. Zhou](mailto:yi.zhou@stu.xidian.edu.cn), [J. W. Fan (Xidian Univ.)](mailto:jw.fan@stu.xidian.edu.cn), [Y. X. Bai](mailto:bai.yaxian@zte.com.cn), [C. Huang](mailto:huang.cheng5@zte.com.cn), [Y. J. Zhang (ZTE)](mailto:zhang.yuanjian@zte.com.cn) |
| **E2E Methods** | | |
| [JVET-AB0125](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12052) | AHG11 - CompressAI models integration using SADL | [F. Galpin](mailto:franck.galpin@interdigital.com), F.Levebvre, F.Racape (InterDigital) |
| **Software** | | |
| [JVET-AB0126](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12053) | AhG11 - SADL update | [F.Galpin](mailto:franck.galpin@interdigital.com), T.Dumas, P.Bordes, E.Francois (InterDigital) |
| [JVET-AB0176](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12103) | NCS-1.0 status report | [Y. Li (Bytedance)](mailto:yue.li@bytedance.com), [F. Galpin (InterDigital)](mailto:franck.galpin@interdigital.com), [H. Wang (Qualcomm)](mailto:hongtaow@qti.qualcomm.com), [L. Wang (Tencent)](mailto:liqiangwang@tencent.com) |
| [JVET-AB0183](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12110) | Preliminary draft of algorithm description for Neural Compression Software (NCS-1) | [Y. Li](mailto:yue.li@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com), [H. Wang](mailto:hongtaow@qti.qualcomm.com), [S. Eadie](mailto:seadie@qti.qualcomm.com), [M. Coban](mailto:mcoban@qti.qualcomm.com), [M. Karczewicz (Qualcomm)](mailto:martak@qti.qualcomm.com), [L. Wang](mailto:liqiangwang@tencent.com), [X. Xu](mailto:xiaozhongxu@tencent.com), [S. Liu (Tencent)](mailto:shanl@tencent.com), [F. Galpin (InterDigital)](mailto:Franck.galpin@interdigital.com), [J. Ström (Ericsson)](mailto:jacob.strom@ericsson.com) |
| **Cross Checks** | | |
| [JVET-AB0204](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12131) | Crosscheck of JVET-AB0158 ([AHG11] On chroma order adjustment in NNLF) | [L. Wang (Tencent)](mailto:liqiangwang@tencent.com) |
| [JVET-AB0205](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12132) | Crosscheck of JVET-AB0159 ([AHG11] On adjustment of residual for NNLF) | [L. Wang (Tencent)](mailto:liqiangwang@tencent.com) |
| [JVET-AB0206](file:////Users/shanliu-sl/Documents/contribution/jvet28ab/current_document.php%3fid=12133) | Crosscheck of JVET-AB0160 ([AHG11] Combination of chroma order adjustment and residual adjustment for NNLF) | [L. Wang (Tencent)](mailto:liqiangwang@tencent.com) |

1. **Recommendations**

The AHG recommends:

* Review all input contributions.
* Discuss and align on the configuration of EncDbOpt for the AI configuration.
* Discuss if the new common software is acceptable to the group and update CTC as appropriate.
* Continue investigating neural network-based video coding tools, including coding performance and complexity.

It was commented that enabling the DBF encoder optimization was found to be not beneficial in AI in VTM. Therefore, it should also not be enabled for AI CTC.

It was decided, that starting from the current meeting, a separate AHG should be established for the development and maintenance of the NN software packages (“NCS”, SADL).

Software copyright status was confirmed later by A. Segall.

[JVET-AB0012](https://jvet-experts.org/doc_end_user/current_document.php?id=12004) JVET AHG report: Enhanced compression beyond VVC capability (AHG12) [M. Karczewicz, Y. Ye, L. Zhang (co-chairs), B. Bross, G. Li, X. Li, K. Naser, H. Yang (vice-chairs)]

Presented Fri 21 Oct. 1410-1415

1. **Activities**

The primary activity of the AHG was the “Exploration experiment on enhanced compression beyond VVC capability” (JVET-AA2024). The combined improvements of the ECM-6.0 over VTM-11.0ecm anchorfor AI, RA and LB configurations are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | All Intra Main10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -8.15% | -16.36% | -23.24% | 490% | 338% |
| Class A2 | -11.93% | -19.18% | -22.12% | 468% | 311% |
| Class B | -6.68% | -18.88% | -17.18% | 413% | 285% |
| Class C | -7.36% | -11.49% | -11.91% | 414% | 277% |
| Class E | -8.32% | -16.40% | -15.40% | 388% | 300% |
| Overall | **-8.22%** | **-16.45%** | **-17.55%** | **430%** | **298%** |
| Class D | -6.07% | -9.36% | -9.16% | 413% | 298% |
| Class F | -17.83% | -24.91% | -24.43% | 336% | 299% |
| Class TGM | -30.28% | -35.62% | -34.91% | 303% | 304% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Random Access Main 10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -18.76% | -21.33% | -28.13% | 526% | 652% |
| Class A2 | -20.59% | -26.86% | -29.77% | 544% | 884% |
| Class B | -16.88% | -26.97% | -24.74% | 449% | 678% |
| Class C | -18.75% | -21.90% | -21.83% | 427% | 692% |
| Class E |  |  |  |  |  |
| Overall | **-18.50%** | **-24.47%** | **-25.65%** | **475%** | **713%** |
| Class D | -19.78% | -23.23% | -23.42% | 426% | 747% |
| Class F | -21.21% | -27.21% | -27.56% | 384% | 442% |
| Class TGM | -29.30% | -33.62% | -34.08% | 417% | 356% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Low Delay B Main 10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -15.49% | -29.59% | -28.03% | 390% | 595% |
| Class C | -16.89% | -24.94% | -25.01% | 396% | 580% |
| Class E | -14.04% | -22.57% | -22.67% | 390% | 390% |
| Overall | **-15.59%** | **-26.29%** | **-25.68%** | **392%** | **531%** |
| Class D | -19.04% | -26.06% | -26.25% | 392% | 669% |
| Class F | -19.92% | -27.91% | -26.96% | 384% | 391% |
| Class TGM | -28.62% | -34.50% | -34.38% | 401% | 324% |

The rate reduction for natural sequences over VTM for RA configuration for {Y, U, V} increased from {-16.80%, -22.11%, -23.07%} to {-18.50%, -24.47%, -25.65%}. For SCC sequences (class TGM) the rate reduction for RA configuration increased from {-24.41%, -28.17%, -28.50%} to {-29.30%, -33.62%, -34.08%}.

Jointly with AG5 and JVET AHG4 and AHG7, AHG12 organized teleconferences to prepare a subjective assessment of the ECM performance at the 28th JVET meeting.

1. **Contributions**

In addition to 26 EE2 contributions, 45 (comparing to 38 last meeting) related contributions were received which can be subdivided as follows:

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

JVET-AB0181, "Non-EE2: Using prediction samples or residual samples for adaptive loop filter", C. Ma, X. Xiu, C.-W. Kuo, W. Chen, H.-J. Jhu, N. Yan, X. Wang (Kwai)

JVET-AB0185, "Non-EE2: ALF with Diversified Extended Taps", W. Yin, K. Zhang, L. Zhang”, (Bytedance)

JVET-AB0186, "EE2-related: Modification of extended offline-filter taps for ALF", I. Jumakulyyev, N. Hu, V. Seregin, M. Karczewicz (Qualcomm)

***Intra (16)***

JVET-AB0065, "Non-EE2: Adaptive reference region DIMD", Z. Fan, Y. Yasugi, T. Ikai (Sharp)

JVET-AB0099, "Non-EE2: CCCM with Multi-shape filters", C. Fang, S. Peng, D. Jiang, J. Lin, X. Zhang (Dahua)

JVET-AB0103, "EE2-1.16 related: Modifications of picture-level geometry transform", J. Choi, S. Yoo, J. Lim, S. Kim (LGE)

JVET-AB0104, "EE2-related: On directional planar prediction", S. Yoo, J. Choi, J. Nam, M. Hong, J. Lim, S. Kim (LGE)

JVET-AB0110, "EE2-1.15-related: Improvements on planar horizontal and planar vertical mode", K. Kim, D. Kim, J.-H. Son, J.-S. Kwak (WILUS)

JVET-AB0111, "Non-EE2: Directional intra prediction fusion with PMPM list", G. Moon, D. Park, Y.-U. Yoon, J.-G. Kim (KAU)

JVET-AB0116, "AHG12 - Location-dependent Decoder-side Intra Mode Derivation", S. Blasi, J. Lainema (Nokia)

JVET-AB0117, "AHG12 - Template-based Intra Mode Derivation with Directional blending", S. Blasi, J. Lainema, I. Zupancic, D. Bugdayci Sansli (Nokia)

JVET-AB0119, "Non-EE2: Gradient and location based convolutional cross-component model (GL-CCCM) for intra prediction", R. G. Youvalari, P. Astola, J. Lainema (Nokia)

JVET-AB0128, "EE2-related: CCCM template selection", P. Bordes, K. Naser, E. Francois, F. Galpin (InterDigital)

JVET-AB0138, "EE2-related: MRL candidate list reordering", Yujin Lee, Bumyoon Kim, Byeungwoo Jeon (SKKU)

JVET-AB0139, "EE2-related: On Chroma Fusion improvement", C. Zhou, Z. Lv, J. Zhang (vivo)

JVET-AB0142, "Non-EE2: optimizing the use of available decoded reference samples", T. Dumas, K. Reuzé, Y. Chen, K. Naser (Interdigital)

JVET-AB0161, "EE2-1.16 related: Encoder optimization for picture-level geometry transform", W. Jia, K. Zhang, Y. Wang, T. Fu, Y. Li, L. Zhang (Bytedance)

JVET-AB0162, "EE2-ralated: On horizontal and vertical planar modes", X. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba)

JVET-AB0174, "AHG12: Division-free operation and dynamic range reduction for convolutional cross-component model (CCCM)", A. Aminlou, J. Lainema, R. G. Youvalari, P. Astola (Nokia)

***Inter (9)***

JVET-AB0145, "EE2-2.6-related: On Decoder-side Affine Model Refinement (DAMR)", J. Chen, R.-L. Liao, X. Li, Y. Ye (Alibaba)

JVET-AB0151, "EE2-2.1 related: ARMC merge candidate list reordering for AMVP-merge mode for low-delay pictures", K. Cui, C. S. Coban, Z. Zhang, V. Seregin, H. Huang, M. Karczewicz (Qualcomm), H. Jang (LGE)

JVET-AB0166, "Non-EE2: Unified pruning of affine merge candidates derivation", Z. Deng, K. Zhang, L. Zhang (Bytedance)

JVET-AB0168, "Non-EE2: Pixel based affine motion compensation", Z. Zhang, H. Huang, Y. Zhang, P. Garus, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AB0177, "EE2-related: Sub-block processing for affine DMVR", H. Huang, Y. Zhang, Z. Zhang, C.-C. Chen, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AB0178, "EE2-related: Control-point motion vector refinement for affine DMVR", H. Huang, C.-C. Chen, Y. Zhang, Z. Zhang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AB0182, "Non-EE2: Bi-predictive local illumination compensation", X. Xiu, N. Yan, H.-J. Jhu, W. Chen, C.-W. Kuo, C. Ma, X. Wang (Kwai)

JVET-AB0189, "AHG12: On bit length control of regression based affine merge candidate derivation", Y. Zhang, H. Huang, V. Seregin, C.-C. Chen, M. Karczewicz (Qualcomm)

JVET-AB0192, "Non-EE2: Extended partitioning mode for the inter/intra prediction", Y. Kidani, H. Kato, K. Kawamura (KDDI)

***RPR (2)***

JVET-AB0082, "AHG12: Fixes for RPR", K.Andersson, R. Yu, (Ericsson)

JVET-AB0133, "AHG12: Inter-RPL and 1-byte NAL unit headers", R. Sjöberg, M. Pettersson, J. Ström (Ericsson)

***Transform Coding (3)***

JVET-AB0100, "Non-EE2: Separable KLT for intra coding", M. Koo, J. Zhao, J. Lim, S. Kim (LGE)

JVET-AB0115, "EE2-1.14 related: Modifications of MTS and LFNST for IntraTMP coded block", D. KIM, K. KIM, J. Son, J. S KWAK(WILUS)

JVET-AB0175, "Non-EE2: Non-Separable Primary Transform for Intra Coding", P. Garus, M. Coban, B. Ray, V. Seregin, M. Karczewicz (Qualcomm)

***Screen Content Coding (11)***

JVET-AB0062, "EE2-related: Modifications of EE2-3.2 and EE2-3.3", W. Lim, D. Kim, J. Kim, S.-C. Lim, J. S. Choi (ETRI), K. Naser, T. Dumas, T. Poirier, F. Galpin, A. Robert (InterDigital)

JVET-AB0094, "Non-EE2: Direct block vector (DBV) mode for chroma prediction", J.-Y. Huo, X. Hao, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.), J. Ren, M. Li (OPPO)

JVET-AB0095, "Non-EE2: Block Vector Difference Sign Prediction (BVDSP) for IBC blocks", J.-Y. Huo, X. Hao, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.), J. Ren, M. Li (OPPO)

JVET-AB0144, "EE2 related: Extension of test EE2-3.3", F. Le Leannec, P. Andrivon, M. Radosavljević, M. Blestel (Xiaomi)

JVET-AB0170, "Non-EE2: Block Vector Difference Prediction for IBC blocks", A. Filippov, V. Rufitskiy (Ofinno)

JVET-AB0173, "AHG12: BVP candidates clustering and BVD sign derivation for Reconstruction-Reordered IBC mode", Damian Ruiz Coll, Vikas Warudkar, Jung-Kyung Lee, (Ofinno)

JVET-AB0180, "Non-EE2: CCCM using non-downsampled luma samples", H.-J. Jhu, C.-W. Kuo, X. Xiu, W. Chen, N. Yan, C. Ma, X. Wang (Kwai)

JVET-AB0187, "Non-EE2: No luma subsampling for CCCM", V. Seregin, Y.-J. Chang, B. Ray, M. Karczewicz (Qualcomm)

JVET-AB0188, "Non-EE2: Extensions of intra block copy", Y. Wang, K. Zhang, L. Zhang, N. Zhang (Bytedance)

JVET-AB0190, "Non-EE2: Combination of JVET-AB0094 and JVET-AB0095 for screen content", J.-Y. Huo, X. Hao, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.), J. Ren, M. Li (OPPO)

JVET-AB0191, "Non-EE2: Combined intra block copy and intra mode", C. Ma, X. Xiu, W. Chen, J.-H. Jhu, C.-W. Kuo, N. Yan, X. Wang (Kwai)

***Other (1)***

JVET-AB0066, "ECM-6 intra performance evaluation on non-CTC dataset", Y. Yasugi, T. Ikai (Sharp)

1. Recommendations

The AHG recommends to:

* To review all the related contributions.
* Conduct a subjective assessment of the ECM performance.

[JVET-AB0013](https://jvet-experts.org/doc_end_user/current_document.php?id=12005) JVET AHG report: Film grain technologies (AHG13) [W. Husak, M. Radosavljević, W. Wan (co-chairs), D. Grois, Y. He, P. de Lagrange, A. Segall, A. Tourapis (vice-chairs)]

2 Discussion

The group focused its efforts on updating the technical report during the intermeeting period. The editors met on a biweekly cadence throughout the period. The editorial process consisted of converging on a structure, adding examples of frequency filtered and autoregressive models and including available tools.

The report was reorganized for readability and flow. Major sections include:

• An overview of film grain

• Film grain synthesis

• Film grain analysis

• Metadata applicable to film grain

Two methods are also discussed: 1) frequency filtered and 2) auto regressive. The frequency filtered method has its origins from the method codified in SMPTE RDD-5 including updates suggested to JVET. Two autoregressive methods are discussed: 1) ISO/IEC/ITU and 2) AFGS1 from AV1.

A multipart annex is included that discusses implementations of film grain technologies. These implementations include authoring, editing and translation tools.

3 Related contributions

Three contributions related to AHG13 were identified as of 10/19/2022:

• One was the AHG report:

o JVET-AB0013 JVET AHG report: Film grain technologies (AHG13)

• Two were related to the technical report and explore film grain testing:

o JVET-AB0042 Film grain synthesis technology for video applications (Draft 3)

o JVET-AB0122 Proposed FGC SEI message verification test draft plan

3.1 Contributions

3.1.1 JVET-AA0042 Film grain synthesis technology for video applications (Draft 3)

This draft technical report provides guidance on the use of film grain synthesis technology for video applications. Such technology may be used in conjunction with metadata signalling mechanisms, such as the supplemental enhancement information messages available in several video coding standards. The purpose of this document is to provide a publicly referenceable overview of the end-to-end processing steps for film grain synthesis, which may include content analysis, noise/film grain removal and film grain model parameter estimation, parameter encoding, encapsulation, and decoding, and film grain synthesis and blending for consumer distribution applications.

3.1.2 JVET-AA0122 Proposed FGC SEI message verification test draft plan

A draft test plan is proposed to assess the visual impact of film grain synthesis, evaluating the potential reduction of bitrate on video content that originally contains grain, maintaining the artistic intent while video coding removes grain. The proposed test compares VVC Main 10 profile with and without film grain synthesis, making use of the FGC SEI message.

The AHG recommends:

• the related input contributions are reviewed;

• the draft TR text be edited during the JVET meeting period;

• testing of FGC be discussed; and

• continue the study of film grain technologies in JVET.

It is suggested to defer issuing the PDTR in January. Issue draft 3 (and WG 5 WD) from the current meeting

# Project development (25)

## Deployment and advertisement of standards (2)

Contributions in this area were discussed at 0910–0920 on Tuesday 25 October 2022 (chaired by JRO).

[JVET-AB0020](https://jvet-experts.org/doc_end_user/current_document.php?id=11959) Deployment status of the HEVC standard [G. J. Sullivan (Microsoft)]

In September 2022, support for HEVC decoding was enabled in the Google Chrome browser (build 105) when HEVC decoding is supported by the device hardware

[JVET-AB0021](https://jvet-experts.org/doc_end_user/current_document.php?id=11960) Deployment status of the VVC standard [G. J. Sullivan (Microsoft)]

Mediatek’s Pentatonic 700 chipset for 4K televisions announced in August 2022 targets 4K televisions and supports frame rates up to 144 Hz

Realtek announced its RTD1319D SoC for set-top boxes in August 2022, featuring VVC support with 4K resolution

Spin Digital announced a real-time VVC encoder designed for live 4K 60 fps and 8K 30 fps UHD streaming and broadcasting in August 2022 ‎[68].The encoder was CPU-based and had been optimized for Intel Xeon Scalable / Ice Lake processors. It included SIMD processing using Intel AVX-512 and Intel Deep Learning Boost instructions, supported advanced multithreading for systems with tens of CPU cores, and had been integrated into a live streaming framework. In October 2022, Spin Digital provided current information to JVET describing their real-time encoder ‎[69]. The encoder was reported to achieve 17.82%, 17.34%, and 18.70% BD-rate savings based on PSNR, MS-SSIM and VMAF, respectively, with an encoding time ratio of 1.71, compared to Spin Digital’s HEVC real-time encoder and was reportedly able to process 4Kp60 and 8Kp30 HDR 10-bit video in real-time using a single server with a dual-socket CPU architecture with 76 CPU cores. A more favourable tradeoff of fidelity with complexity was reported relative to various other encoders (see ‎[69], Figures 3–6 and 11). This encoder had been integrated into a framework for live applications that included SDI capture and HTTP (HLS and DASH) and TS over IP (RTP, SRT, RIST) streaming capabilities. The encoder was able to reduce the company’s recommended bit rates for UHD live video from the 16–17 Mbps needed by the company’s HEVC encoder to 13–14 Mbps for the VVC encoder. The improved capability had reportedly been verified by subjective testing, with 8 out of 11 one-minute test sequences rated “excellent” at 14 Mbps and none rated below “good” quality.

For recent implementations (Spin Digital, VVEnc/VVDec), see also section 4.9.

## Text development and errata reporting (2)

Contributions in this area were discussed at 1155–1210 on Thursday 20 October 2022 (chaired by JRO).

[JVET-AB0223](https://jvet-experts.org/doc_end_user/current_document.php?id=12150) AHG2: Text improvement for Timing / DU information SEI message in HEVC and VVC [K. Sühring, B. Bross (Fraunhofer HHI), Y.-K. Wang (Bytedance)] [late]

This contribution reports an issue in Picture Timing and DU information SEI message in HEVC and VVC and proposes a text fix.

Problem ***#1***

In VVC, Picture timing SEI semantics of pt\_dpb\_output\_delay contains the following text:

For pictures that are not output by the "bumping" process because they precede, in decoding order, a CVSS AU that has sh\_no\_output\_of\_prior\_pics\_flag equal to 1 or inferred to be equal to 1, the output times derived from pt\_dpb\_output\_delay shall be increasing with increasing value of PicOrderCntVal relative to all pictures within the same CVS.

In VVC sh\_no\_output\_of\_prior\_pics\_flag does not have any inference. The highlighted part seems to be a leftover from AVC, where the flag could be inferred by the decoder:

When the IDR picture is not the first IDR picture in the bitstream and the value of PicWidthInMbs, FrameHeightInMbs, or max\_dec\_frame\_buffering derived from the active sequence parameter set is different from the value of PicWidthInMbs, FrameHeightInMbs, or max\_dec\_frame\_buffering derived from the sequence parameter set active for the preceding picture, no\_output\_of\_prior\_pics\_flag equal to 1 may (but should not) be inferred by the decoder, regardless of the actual value of no\_output\_of\_prior\_pics\_flag.

VVC instead derives the variable NoOutputOfPriorPicsFlag in Annex C for the same purpose:

When the current picture is the first picture of a CVSS AU that is not AU 0, the following ordered steps are applied:

1. The variable NoOutputOfPriorPicsFlag is derived for the decoder under test as follows:

– If the value of PicWidthMaxInSamplesY, PicHeightMaxInSamplesY, MaxChromaFormat, MaxBitDepthMinus8, or dpb\_max\_dec\_pic\_buffering\_minus1[ Htid ] derived for the current AU is different from the value of PicWidthMaxInSamplesY, PicHeightMaxInSamplesY, MaxChromaFormat, MaxBitDepthMinus8, or dpb\_max\_dec\_pic\_buffering\_minus1[ Htid ], respectively, derived for the preceding AU in decoding order, NoOutputOfPriorPicsFlag may (but should not) be set equal to 1 by the decoder under test, regardless of the value of sh\_no\_output\_of\_prior\_pics\_flag of the current AU.

NOTE – Although setting NoOutputOfPriorPicsFlag equal to sh\_no\_output\_of\_prior\_pics\_flag of the current AU is preferred under these conditions, the decoder under test is allowed to set NoOutputOfPriorPicsFlag equal to 1 in this case.

– Otherwise, NoOutputOfPriorPicsFlag is set equal to sh\_no\_output\_of\_prior\_pics\_flag of the current AU.

2. The value of NoOutputOfPriorPicsFlag derived for the decoder under test is applied for the HRD, such that when the value of NoOutputOfPriorPicsFlag is equal to 1, all picture storage buffers in the DPB are emptied without output of the pictures they contain, and the DPB fullness is set equal to 0.

Thus the semantics should be changed to refer to NoOutputOfPriorPicsFlag instead.

The phrase “sh\_no\_output\_of\_prior\_pics\_flag equal to 1 or inferred to be equal to 1” occurs in total three times in the standard text:

* semantics of pt\_dpb\_output\_delay (Picture timing SEI message semantics)
* semantics of pt\_dpb\_output\_du\_delay (Picture timing SEI message semantics)
* semantics of dui\_dpb\_output\_du\_delay (DU information SEI message semantics)

All three references should be changed to “NoOutputOfPriorPicsFlag equal to 1”.

The same problem exists in in HEVC with slightly different syntax element names. The phrase “no\_output\_of\_prior\_pics\_flag equal to 1 or inferred to be equal to 1” occurs in

* semantics of pic\_dpb\_output\_delay (Picture timing SEI message semantics)
* semantics of pic\_dpb\_output\_du\_delay (Picture timing SEI message semantics)
* semantics of pic\_spt\_dpb\_output\_du\_delay (Decoding unit information SEI message semantics)

All three references should be changed to “NoOutputOfPriorPicsFlag equal to 1”.

Problem ***#2***

In VVC the

* semantics of pt\_dpb\_output\_du\_delay (Picture timing SEI message semantics)
* semantics of dui\_dpb\_output\_du\_delay (DU information SEI message semantics)

refer to “a CLVSS picture that has sh\_no\_output\_of\_prior\_pics\_flag equal to 1 or inferred to be equal to 1”. The part “a CLVSS picture” should be replaced by “a CVSS AU” as in the semantics of pt\_dpb\_output\_delay.

Problem ***#3***

In VVC a minor text improvement for the semantics of sh\_no\_output\_of\_prior\_pics\_flag is suggested. The word value should be removed as follows:

**sh\_no\_output\_of\_prior\_pics\_flag** affects the output of previously-decoded pictures in the DPB after the decoding of a picture in a CVSS AU that is not the first AU in the bitstream as specified in Annex ‎C.

It is a requirement of bitstream conformance that the value of sh\_no\_output\_of\_prior\_pics\_flag shall be the same for all slices in an AU that have sh\_no\_output\_of\_prior\_pics\_flag present in the SHs.

When all slices in an AU have sh\_no\_output\_of\_prior\_pics\_flag present in the SHs, the value of sh\_no\_output\_of\_prior\_pics\_flag in the SHs is also referred to as the value sh\_no\_output\_of\_prior\_pics\_flag of the AU.

Decision(ed. Alignment): Include in JVET-AB1004 bug report doc.

[JVET-AB0172](https://jvet-experts.org/doc_end_user/current_document.php?id=12099) Proposed ITP-PQ-C2 codepoint [W. Husak (Dolby)] [late]

SMPTE is in the process of standardizing a color representation used by streaming platforms to deliver HDR content to consumers. The representation is a variation of the ITP color format that was developed for print applications. ITP has a distinct characteristic where the hue is perceptually linear across a wide range of luminance values. The hue linearity from ITP coupled with the perceptual luminance linearity from the PQ (BT.2100) transfer function results in ITP-PQ-C2 being an ideal color volume for image processing - especially when expanding or contracting a color volume.

The following is proposed:

1. Add appropriate equations in the semantics under **matrix\_coeffs.**
2. Add an additional code point in Table E.5.

Action is requested for Rec. ITU-T H.265 | ISO/IEC 23008-2 and VUI parameter sets in Rec. ITU-T H.273 | ISO/IEC 23091-2.

Include in JVET-AB1008 (for AVC, HEVC and CICP)

W. Husak volunteers to draft liaison response to SMPTE on M61406.

## Test conditions (2)

Contributions in this area were discussed at 0920–0945 on Tuesday 25 October 2022 (chaired by JRO).

[JVET-AB0066](https://jvet-experts.org/doc_end_user/current_document.php?id=11981) ECM-6 intra performance evaluation on non-CTC dataset [Y. Yasugi, T. Ikai (Sharp)]

This contribution provides the results of a validation test of ECM-6.0 using DIV2K and Flikr2K datasets to test the performance of ECM on images of various subjects. Experimental results show -6.91%/-17.03%/-17.30% and -6.58%/-14.81%/-15.20% (All-intra, Y/U/V) on DIV2K and Flickr2K, respectively under all-intra condition. Experimental results confirmed that the performance is on per with that on Class-B CTC contents.

It was found that ECM performed especially well for images with lots of geometric structures, and less well for natural structures.

[JVET-AB0271](https://jvet-experts.org/doc_end_user/current_document.php?id=12200) AHG3: report of latest HM performance on HDR content [E. François, T. Lu, S. Iwamura, A. Segall] [late]

This contribution reports performance results of HM16.26 for HDR CTCs content.

The following changes related to HM or HDR content have been made since JVET-U meeting.

1. Addition of MCTF feature in HM, and activation in HM CTCs.
2. Changes of the GOP structure for Random Access, to conform the maximum DPB size specified for HEVC.
3. New version of the class H2 (HDR-HLG) clips, to conform type 2 chroma alignment; the old version has type 0 chroma alignment.

A first task was to reproduce the JVET-U0007 results. Regarding class H1, Dolby and Interdigital were able to get mutual matching results, using HM16.18HBD, but those results slightly differ from the ones reported in JVET-U0007, as depicted in Table 1. Regarding class H2 (old version), NHK and Interdigital were able to get mutual matching results, using HM16.18, but those results also slightly differ from the ones reported in JVET-U0007, as depicted in Table 1.

Table 1. HM16.18 Oct. 2022 over JVET-U0007.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access** | | | | | | | |
|  | **Over HM16.18 JVET-U** | | | | | | | |
|  |  |  | **wPSNR** |  |  | **PSNR** |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V |
| Class H1 | 0.01% | 0.02% | 0.00% | 0.01% | -0.13% | 0.00% | -0.02% | -0.12% |
| Class H2 |  |  |  |  |  | 0.00% | 0.01% | -0.05% |
| **Overall** | 0.01% | 0.02% | 0.00% | 0.01% | -0.13% | 0.00% | -0.01% | -0.10% |
|  |  |  |  |  |  |  |  |  |
|  | **All Intra** | | | | | | | |
|  | **Over HM16.18 JVET-U** | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V |
| Class H1 | -0.02% | -0.02% | -0.01% | -0.02% | -0.01% | -0.01% | -0.01% | 0.00% |
| Class H2 |  |  |  |  |  | 0.00% | 0.00% | 0.00% |
| **Overall** | -0.02% | -0.02% | -0.01% | -0.02% | -0.01% | -0.01% | -0.01% | 0.00% |

In Table 2, test is HM16.26HBD (using configuration files attached to HM16.26), anchor is HM16.18HBD (using configuration files attached to HM16.18). Gains are observed in RA for all metrics, mostly due to the activation of MCTF.

In Table 3, test is HM16.26HBD, anchor is HM16.26. Minor differences are observed.

Corresponding xls files are attached to the contribution.

Table 2. HM16.26HBD over HM16.18HBD.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access** | | | | | | | |
|  | **Over HM16.18 HBD** | | | | | | | |
|  |  |  | **wPSNR** |  |  | **PSNR** |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V |
| Class H1 | -3.65% | -3.41% | -3.15% | -1.64% | 1.37% | -3.20% | -1.78% | 0.83% |
| Class H2 new |  |  |  |  |  | -4.42% | -3.72% | -3.30% |
| **Overall** | -3.65% | -3.41% | -3.15% | -1.64% | 1.37% | -3.65% | -2.49% | -0.67% |
|  |  |  |  |  |  |  |  |  |
|  | **All Intra** | | | | | | | |
|  | **Over HM16.18 HBD** | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V |
| Class H1 | 0.00% | 0.00% | 0.00% | -0.07% | -0.04% | 0.00% | -0.06% | -0.07% |
| Class H2 new |  |  |  |  |  | 0.00% | 0.00% | 0.00% |
| **Overall** | 0.00% | 0.00% | 0.00% | -0.07% | -0.04% | 0.00% | -0.04% | -0.04% |

Table 3. HM16.26HBD over HM16.26.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access** | | | | | | | |
|  | **Over HM16.26** | | | | | | | |
|  |  |  | **wPSNR** |  |  | **PSNR** |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V |
| Class H1 | 0.05% | 0.01% | 0.03% | 0.14% | -0.11% | 0.03% | 0.17% | -0.15% |
| Class H2 new |  |  |  |  |  | 0.01% | -0.10% | -0.02% |
| **Overall** | 0.05% | 0.01% | 0.03% | 0.14% | -0.11% | 0.02% | 0.07% | -0.10% |
|  |  |  |  |  |  |  |  |  |
|  | **All Intra** | | | | | | | |
|  | **Over HM16.26** | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V |
| Class H1 | -0.10% | 0.04% | 0.02% | -0.31% | -0.22% | 0.02% | -0.35% | -0.20% |
| Class H2 new |  |  |  |  |  | 0.00% | -0.04% | -0.09% |
| **Overall** | -0.10% | 0.04% | 0.02% | -0.31% | -0.22% | 0.01% | -0.24% | -0.16% |

A mismatch between parallel and sequential coding is observed with the HM16.18HBD and HM16.26HBD versions. This requires further investigation.

It is suggested:

• to add to the HM package a folder comprising the HDR per-sequence configuration files (see attached file “per-sequence-HDR.zip”).

• to use HM HBD version for both H1 and H2 classes.

Decision(CTC): Adopt (config files for HDR in HM)

## Subjective quality testing and verification testing (4)

Contributions in this area were discussed in a joint meeting with AG 5 at 1400–1530 on Tuesday 25 October 2022 (chaired by JRO).

[JVET-AB0041](https://jvet-experts.org/doc_end_user/current_document.php?id=11953) AHG4, 7, 12: Report on AHG meetings on ECM performance evaluation preparation [M. Wien (RWTH)]

[JVET-AB0122](https://jvet-experts.org/doc_end_user/current_document.php?id=12049) Proposed FGC SEI message verification test draft plan [P. de Lagrange (InterDigital), W. Husak (Dolby)]

A draft test plan is proposed to assess the visual impact of film grain synthesis, evaluating the potential reduction of bitrate on video content that originally contains grain, maintaining the artistic intent while video coding removes grain. The proposed test compares VVC Main 10 profile with and without film grain synthesis, making use of the FGC SEI message.

Film grain synthesis can be used to recreate a grain that visually match the original, that has been lost due to compression, thus restoring the original artistic intent. It can also be used to mask coding defects to some extent.

This feature has been available for a long time, but is currently driving more attention since recent advances in video compression have come with a significant reduction of noise, making film grain preservation more difficult at broadcast or streaming bitrates.

It is proposed to compare visually the result of VVC video compression with and without FGC SEI message and film grain synthesis, to evaluate how much rate can be saved using this feature, on video content that is relevant for this purpose (e.g. movie content, that actually contains grain).

Further work is necessary to define sequences, rate points, quality of the codings of the sequences with and without film grain.

It was also suggested to investigate how close the synthesized film grain comes to the original one in terms of subjective perception. A methodology to assess this would need to be identified, and this would likely only be tested at (almost) transparent quality of the video.

Better call it “subjective quality test plan”

It should be based on the available software implementation of the SEI message; if other implementations are proposed to JVET before the test is executed, those might also be considered, depending on available ressources.

Action item to further work out these issues in AHG discussions.

[JVET-AB0123](https://jvet-experts.org/doc_end_user/current_document.php?id=12050) Proposed multilayer VVC verification test draft plan [P. de Lagrange, G. Marquant, C. Salmon-Legagneur, F. Urban (InterDigital), M. Wien (RWTH)]

This document contains a proposed draft plan for the video verification test to be conducted to verify the coding performance of the VVC Multilayer Main 10 profile in several configurations: spatial scalability with optimized coding performance, quality ladder involving both spatial and quality scalability, and stereo coding. The three configurations involve output layer sets with maximum two layers.

It is suggested that a formal subjective evaluation be conducted comparing the VVC Multilayer Main 10 profile to VVC single layer and simulcast using the VVC Main 10 profile. SHVC as an additional anchor is to be discussed.

The suggested tests include three categories, covering different aspects of multilayer coding:

* Spatial scalability coding efficiency, using UHD enhancement layer on top of 1.5x or 2x downscaled base layer with a QP offset that optimizes coding efficiency. This addresses use cases where spatial scalability feature itself (= multiple resolutions) is of interest, with quality comparable to or better than single layer. In this category, enhancement layer quality is compared to single layer.
* Quality ladder, using a range of resolutions and bit rates. This exercises both spatial and quality scalability, with resolution ranges from 960x540 to UHD with bitrates from 1 to 10x, and quality points grouped in pairs of layers which can have 1.33x, 1.5x or 2x resolution ratios. The use case addressed is local cache storage space for streaming applications. In this category, each quality point (which can be either base or enhancement layer) is compared with single layer.
* Dual-view (stereo). The anchor is simulcast.

It was suggested to start with the first case.

JVET-AB0045 relates to that case. Follow-up discussion was performed in joint BoG with AG 5.

[JVET-AB0270](https://jvet-experts.org/doc_end_user/current_document.php?id=12199) Report on subjective performance evaluation of the ECM [M. Wien, V. Baroncini (Test coordinators)]

Initial results presented (powerpoint slides) Tuesday 25 Oct. 1400

Subjective improvement of ECM was found in a relatively consistent way, though confidence intervals slightly overlapping in the cases of expert viewing in Mainz. Very similar results obtained with non-expert viewers in Rome, but confidence intervals smaller, gain of ECM even more consistent than from the expert viewing.

Request a joint meeting with parent bodies to discuss about future perspectives of video coding.

## Test material (0)

This section is kept as a template for future use.

## Quality assessment (0)

This section is kept as a template for future use.

## Conformance test development (1)

Contributions in this area were discussed at 1235–1250 on Wednesday 25 October 2022 (chaired by JRO).

[JVET-AB0085](https://jvet-experts.org/doc_end_user/current_document.php?id=12012) AHG5: On conformance test of multilayer coding [S. Iwamura, S. Nemoto, A. Ichigaya (NHK)] [late]

This contribution proposes addition of new conformance tests of multiple profile/tier/level (PTL) support in multilayer coding. In VVC specifications, when the number of layers to be encoded is greater than 1, multiple PTLs can be signalled in VPS. For each PTL, profile, tier and general constraints information is optionally signalled based on vps\_pt\_present\_flag. However, the current implementation of VTM-18.0 only allows one profile, tier, and general constraints information to be signalled in the VPS. Due to this limitation, VTM-18.0 does not work properly in the case of multiple PTL signalling for different OLSs such as applying Main 10 profile to the base layer and Multilayer Main 10 profile to the enhancement layer.

A patch file for the software fix is provided in the contribution. Note that this fix only affects VTM software and VVC specification is not changed.

In addition to the software fix, addition of two new conformance tests is proposed for the following use cases:

* Different profiles applying to different OLSs
* The same profile applying to multiple independent layers

The sample bitstreams of the proposed conformance tests are also provided in this contribution.

It was noted that for the decoder bug fix there is already a pending action from the side of the software coordinators.

The corresponding encoder change is also necessary

Decision(SW): Adopt the corresponding encoder change from JVET-AB0085.

For the suggested action about additional conformance bitstreams, it is too late trying to include them into the ongoing v2 amendment.

It was planned to issue a JVET output document “Additional conformance bitstreams for VVC multilayer configurations”, which can have a long editing period.

## Software development (6)

Contributions in this area were discussed at 0945–XXXX on Tuesday 25 October 2022 (chaired by JRO).

[JVET-AB0072](https://jvet-experts.org/doc_end_user/current_document.php?id=11987) VTM Encoder Implementation for Green-MPEG SEI Messaging [C. Herglotz, M. Kränzler, A. Kaup (FAU)]

This document proposes a reference encoder implementation for generation and signaling of Green MPEG complexity metrics for decoder power reduction and quality metrics for quality recovery after low power encoding, which are both defined for VVC SEI messages. The implementation complies with the issued FDIS ISO/IEC 23001-11 3rd edition Energy-Efficient Media Consumption (Green Metadata). The implementation is available for the VTM-17.2 encoder implementation. It is proposed to include the code for a future version of VTM.

Was there any impact on encoder run time? It was answered that there is some effect, but less than 10% run time increase. It was requested that the functionality would need to be disabled by a macro, such that there is no impact on run time if the SEI message is not used.

Decision(SW): Adopt JVET-AB0072 (optional, not CTC)

[JVET-AB0176](https://jvet-experts.org/doc_end_user/current_document.php?id=12103) NCS-1.0 status report [Y. Li (Bytedance), F. Galpin (InterDigital), H. Wang (Qualcomm), L. Wang (Tencent)]

This report summarizes the activities on Neural Compression Software development that has taken place between the 27th and 28th JVET meetings.

NCS software was located at <https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/VVCSoftware_VTM/-/tree/NCS-1.0>.

NCS software is based on VTM-11.0 with enabled MCTF including the update from JVET-V0056, and enabled EncDbOpt for LD and RA (same as in NNVC-2.0).

The following adopted aspects were integrated into NCS-1.0:

JVET-AA0088: Neural network based in-loop filter with a single model (configured to be Filter set #0)

JVET-AA0111: Deep In-Loop Filter With Fixed Point Implementation (configured to be Filter set #1)

JVET-AA0086: Small ad-hoc deep-learning library (SADL) update

The additional SPS syntax is integrated into NCS software as agreed to enable or disable the NN-based filters, or switch between Filter set #0 and Filter set #1. The choice of the filter set is configurable directly from the configuration file or the command line.

Pretrained models are available in the repository and used to compute the anchors.

Training scripts and documentation have been updated to recompute the models.

NCS-1.0 was tagged on September 5, 2022.

In this section, NCS-1.0 test results following NCS CTC configuration descried in JVET-AA2016 are summarized.

The following tables show NCS-1.0 NN-based filter set #0 (int16 precision) performance over NNVC-2.0 (<https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/VVCSoftware_VTM/-/tree/VTM-11.0_nnvc-2.0>) anchor.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -9.15% | -13.89% | -18.71% | 153% | 73018% |
| Class A2 | -9.37% | -19.11% | -14.46% | 149% | 70078% |
| Class B | -8.33% | -18.68% | -20.42% | 156% | 70892% |
| Class C | -8.37% | -20.30% | -20.35% | 143% | 60876% |
| Class E |  |  |  |  |  |
| **Overall** | -8.71% | -18.24% | -18.87% | 151% | 68316% |
| Class D | -9.58% | -19.42% | -20.42% | 139% | 58850% |
| Class F | -3.72% | -11.90% | -10.92% | 197% | 32349% |
|  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -7.32% | -17.59% | -20.15% | 144% | 69084% |
| Class C | -7.86% | -20.18% | -20.40% | 131% | 55098% |
| Class E | -8.43% | -17.14% | -16.50% | 213% | 54715% |
| **Overall** | -7.78% | -18.74% | -20.26% | 154% | 60438% |
| Class D | -9.21% | -18.73% | -19.41% | 132% | 51121% |
| Class F | -4.17% | -10.03% | -8.79% | 185% | 32820% |
|  |  |  |  |  |  |
|  | **All Intra Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -6.12% | -14.26% | -17.26% | 211% | 50641% |
| Class A2 | -5.73% | -15.81% | -12.98% | 159% | 41067% |
| Class B | -5.93% | -15.54% | -17.46% | 149% | 38489% |
| Class C | -6.40% | -16.42% | -18.16% | 133% | 27723% |
| Class E | -8.83% | -15.25% | -16.12% | 167% | 44744% |
| **Overall** | -6.52% | -15.52% | -16.61% | 159% | 38826% |
| Class D | -6.41% | -15.30% | -18.56% | 126% | 25806% |
| Class F | -4.02% | -11.81% | -11.71% | 128% | 32540% |

The following tables show NCS-1.0 NN-based filter set #0 (float precision) performance over NNVC-2.0 anchor.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -9.21% | -14.66% | -18.35% | 170% | 97724% |
| Class A2 | -9.37% | -19.29% | -14.32% | 163% | 93013% |
| Class B | -8.31% | -19.02% | -20.33% | 172% | 93877% |
| Class C | -8.34% | -20.27% | -20.09% | 159% | 80518% |
| Class E |  |  |  |  |  |
| **Overall** | -8.71% | -18.53% | -18.67% | 166% | 90671% |
| Class D | -9.48% | -19.52% | -20.41% | 162% | 77552% |
| Class F | -3.73% | -11.83% | -11.00% | 229% | 42242% |
|  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -7.31% | -18.18% | -20.28% | 161% | 91967% |
| Class C | -7.85% | -19.93% | -20.16% | 142% | 73330% |
| Class E | -8.48% | -16.46% | -16.42% | 250% | 72758% |
| **Overall** | -7.78% | -18.96% | -20.23% | 172% | 80428% |
| Class D | -9.22% | -18.82% | -19.35% | 148% | 66622% |
| Class F | -4.23% | -9.99% | -8.98% | 211% | 42845% |
|  |  |  |  |  |  |
|  | **All Intra Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -6.13% | -14.24% | -17.26% | 251% | 67748% |
| Class A2 | -5.75% | -15.93% | -12.99% | 176% | 54763% |
| Class B | -5.93% | -15.56% | -17.45% | 166% | 51137% |
| Class C | -6.38% | -16.47% | -18.12% | 148% | 35566% |
| Class E | -8.81% | -15.27% | -16.32% | 180% | 58517% |
| **Overall** | -6.51% | -15.55% | -16.64% | 177% | 51141% |
| Class D | -6.40% | -15.48% | -18.58% | 141% | 32964% |
| Class F | -4.04% | -11.89% | -11.74% | 137% | 44279% |

The following tables show NCS-1.0 NN-based filter set #1 (int16 precision) performance over NNVC-2.0 anchor.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -8.92% | -15.94% | -18.51% | 210% | 52791% |
| Class A2 | -10.15% | -20.20% | -17.29% | 200% | 50531% |
| Class B | -9.01% | -23.08% | -21.77% | 209% | 51646% |
| Class C | -9.82% | -21.78% | -22.31% | 195% | 50009% |
| Class E |  |  |  |  |  |
| **Overall** | -9.44% | -20.73% | -20.37% | 203% | 51205% |
| Class D | -11.43% | -23.49% | -23.88% | 196% | 47037% |
| Class F | -4.69% | -11.70% | -10.66% | 298% | 25719% |
|  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -7.80% | -16.00% | -14.02% | 196% | 51463% |
| Class C | -9.14% | -15.17% | -14.89% | 171% | 45228% |
| Class E | -8.63% | -17.25% | -17.03% | 344% | 43518% |
| **Overall** | -8.45% | -15.63% | -14.40% | 216% | 47271% |
| Class D | -10.78% | -18.72% | -17.99% | 174% | 39565% |
| Class F | -4.97% | -8.98% | -6.08% | 272% | 23963% |
|  |  |  |  |  |  |
|  | **All Intra Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -6.45% | -17.96% | -19.88% | 212% | 41577% |
| Class A2 | -6.50% | -20.46% | -17.82% | 164% | 33736% |
| Class B | -6.52% | -21.87% | -21.59% | 157% | 31884% |
| Class C | -7.40% | -18.86% | -21.76% | 153% | 23335% |
| Class E | -9.87% | -20.85% | -20.66% | 171% | 36317% |
| **Overall** | -7.26% | -20.14% | -20.56% | 168% | 32075% |
| Class D | -7.32% | -19.50% | -21.39% | 153% | 20201% |
| Class F | -4.06% | -11.76% | -10.93% | 134% | 21676% |

The following tables show NCS-1.0 NN-based filter set #1 (float precision) performance over NNVC-2.0 anchor.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -8.86% | -16.12% | -18.83% | 257% | 80229% |
| Class A2 | -10.16% | -20.28% | -17.34% | 246% | 77940% |
| Class B | -8.99% | -22.95% | -21.83% | 255% | 78170% |
| Class C | -9.83% | -21.70% | -22.38% | 226% | 71605% |
| Class E |  |  |  |  |  |
| **Overall** | -9.42% | -20.72% | -20.48% | 245% | 76716% |
| Class D | -11.45% | -23.40% | -23.80% | 231% | 70983% |
| Class F | -4.64% | -11.65% | -10.59% | 363% | 37408% |
|  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -7.75% | -15.97% | -13.64% | 244% | 76605% |
| Class C | -9.14% | -15.24% | -15.19% | 201% | 68109% |
| Class E | -8.52% | -17.06% | -17.80% | 458% | 65923% |
| **Overall** | -8.41% | -15.65% | -14.33% | 268% | 70947% |
| Class D | -10.76% | -18.20% | -17.94% | 200% | 58824% |
| Class F | -4.97% | -9.04% | -5.93% | 351% | 35888% |
|  |  |  |  |  |  |
|  | **All Intra Main10** | | | | |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -6.42% | -17.97% | -19.74% | 247% | 64929% |
| Class A2 | -6.48% | -20.48% | -17.81% | 182% | 52976% |
| Class B | -6.49% | -21.90% | -21.61% | 174% | 49418% |
| Class C | -7.42% | -18.84% | -21.71% | 156% | 33740% |
| Class E | -9.87% | -20.81% | -20.61% | 188% | 56011% |
| **Overall** | -7.25% | -20.15% | -20.52% | 184% | 49081% |
| Class D | -7.35% | -19.43% | -21.33% | 149% | 32772% |
| Class F | -4.08% | -11.73% | -10.89% | 141% | 33401% |

It was suggested to integrate all NN related developments into only one software package, which has also a configuration option that disables all NN tools (which would be the equivalent of what is currently called NNVC2.0). The next version would be called NNVC3.0, which is NNVC2.0+NCS1.0+SADL+any additional adoptions from this meeting.

It was reported that the usual copyright statement was already used in most parts of NCS1.0, except for some of the training scripts. These have been added in the most recent versions.

Any part of the NNVC3.0 shall have the copyright statement.

The definition of an anchor (or anchors) is a matter of the EE. The option with all NN based tools switched off will probably be mandatory as an anchor for all parts of the EE, but certain parts may define additional anchors.

It was asked if as a part of the software development, the coordinators would run a performance test of the different adopted tools for each meeting cycle (similar as done above for the two filter sets). It was commented that this might be duplicating some of the investigations in the EE. Currently, with only two or three sets of filters in the software, this would not be a big effort, but it will be needed in future developments, which amount of detailed checking of adoptions is practical. Generally, it would be assumed that tools to be adopted would have been sufficiently checked in the EE before they are adopted for the software, but some checks might be necessary if in the future several tools are integrated whose interaction has not been investigated before.

[JVET-AB0183](https://jvet-experts.org/doc_end_user/current_document.php?id=12110) Preliminary draft of algorithm description for Neural Compression Software (NCS-1) [Y. Li, K. Zhang, L. Zhang (Bytedance), H. Wang, S. Eadie, M. Coban, M. Karczewicz (Qualcomm), L. Wang, X. Xu, S. Liu (Tencent), F. Galpin (InterDigital), J. Ström (Ericsson)]

In JVET-AA meeting, it was decided to establish a common software for better evaluate neural network-based video coding technologies. Considering that an algorithm description output document will facilitate the study and development of the common software, proponents of the coding tools in NCS-1 provide a preliminary draft of the suggested algorithm description output document in this contribution.

This shall become a new output document JVET-AB2019. This should also include a short reference to VTM which was used as basis of NNVC software, not only algorithm description, but also description of training scripts, references where NNVC and SADL can be found, reference to SADL software manual, etc.

Editors: Y. Li (Bytedance), H. Wang (Qualcomm), L. Wang (Tencent), F. Galpin (InterDigital), J. Ström (Ericsson)

[JVET-AB0086](https://jvet-experts.org/doc_end_user/current_document.php?id=12013) Optimized VVC encoder with multilayer coding capability [S. Iwamura, S. Nemoto, A. Ichigaya (NHK)] [late]

This contribution reports an optimized VVC encoder with multilayer coding capability, which is developed based on VVenC-1.4.0 released from HHI in GitHub. The VVenC-multilayer software is developed to accelerate the investigation of the optimum rate/quality settings for base and enhancement layers in spatial scalability scenario. To balance quality and encoder runtime, 5 presets (faster, fast, medium, slow, and slower) are defined for encoder configurations as is the case with VVenC-1.4.0.

The simulation is conducted using 4K test sequences distributed by ITE (Institute of Image Information and Television Engineers). In the simulation, 2K sequence for the input of base layer is generated by SHVC down sampling filter. The BD-rate of {Y, Cb, Cr} components calculated from PSNR values of the enhancement layer and total bitrate, and the encoder speedup of single thread coding over VTM-16.0 are reported as follows:

Faster BD-rate: {40.3%, 32.8%, 43.1%}, Encoder speedup: 226.7x

Fast BD-rate: {29.0%, 25.4%, 29.5%}, Encoder speedup: 114.2x

Medium BD-rate: {10.5%, 10.4%, 11.4%}, Encoder speedup: 35.1x

Slow BD-rate: {4.9%, 4.0%, 3.9%}, Encoder speedup: 11.3x

Slower BD-rate: {0.1%, 0.3%, -0.1%}, Encoder speedup: 1.2x

The output multilayer bitstream conforms VVC specifications and it is confirmed that the bitstream can be decoded by VTM-16.0. The public release of the developed software and the pull request to the original repository of VVenC are currently being considered.

Very interesting contribution, but no action necessary for JVET. Inclusion in VVEnc would be welcome from the viewpoint of JVET, to emphasize the multilayer capabilities of VVC. Could also be used in a verification test on scalable functionality.

[JVET-AB0210](https://jvet-experts.org/doc_end_user/current_document.php?id=12137) AHG3/10: VTM multilayer profile encoder fixes [F. Urban, P. de Lagrange, G. Marquant, C. Salmon-Legagneur (InterDigital)] [late]

This document presents simulation results of multilayer VTM encoder with two layers spatial scalability. Latest VTM code from master branch after VTM-18.0 is used to include recent merge requests for multilayer profile. Overall, a bdrate improvement of (Y; U; V) -0.3%; -0.4%; -0.3% is achieved, with same encoding / decoding time.

Contribution for information. The changes have already been integrated in VTM.

[JVET-AB0228](https://jvet-experts.org/doc_end_user/current_document.php?id=12157) Lambda-QP relationship fix for slice-level multi-QP optimization [J. Liao, L. Li, D. Liu, H. Li, F. Wu (USTC)] [late]

This contribution proposes to fix the problems of slice-level multi-QP optimization in VTM software. The main modifications we propose to the slice-level multi-QP optimization of the original version in VTM are as follows:

1. We propose to modify the QP-lambda relationship in slice-level multi-QP RDO.
2. We propose to fix the slice-level lambda while traversing the QP within a predefined range.
3. We fix the conflict between slice-level multi-QP optimization and CU-level mode reuse.

The test results show that the fix brings significant performance improvement compared with the original implementation as follows:

AI: {xx%, xx%, xx%}; RA: {xx%, xx%, xx%}; LD: {xx%, xx%, xx%}; LDB: {xx%, xx%, xx%}

It was commented that the proposed changes 1 and 2 appear appropriate, but 1. should be implemented such that it can be switched by macro. It was further commented that change 3 is not necessarily a conflict. What is the impact? It was also asked why (from the partial results presented) in some classes losses occur.

Further study recommended.

## Implementation studies and complexity analysis (2)

Contributions in this area were discussed at 1345–1400 on Tuesday 25 October 2022 and at 1045–1100 on Wednesday 26 October 2022 (chaired by JRO).

[JVET-AB0043](https://jvet-experts.org/doc_end_user/current_document.php?id=11956) A VVC/H.266 real-time software encoder for UHD live video applications [S. Sanz-Rodriguez, M. Alvarez-Mesa, C. C. Chi (Spin Digital)]

This contribution presents a VVC real-time software encoder for Ultra HD (4K and 8K) live streaming and broadcasting. The encoder was compared to Spin Digital’s HEVC real-time encoder and open-source implementations of H.264, HEVC, AV1, and VVC in terms of compression efficiency, complexity, and encoding speed. According to the experimental results, the encoder achieves 17.82% and 18.70% BD-rate savings based on PSNR and VMAF, respectively, compared to Spin Digital’s HEVC real-time encoder.

The VVC encoder is able to process 4Kp60 and 8Kp30 HDR 10-bit videos in real-time using a single server with a state-of-the-art dual-socket CPU architecture. In the near future, through the use of advanced encoding algorithms and next-generation CPU architectures, real-time 8Kp60 encoding and higher compression efficiency are expected to be achieved.

The VVC encoder has been integrated into a framework for live applications that includes SDI capture and HTTP (HLS and DASH) and TS over IP (RTP, SRT, RIST) streaming capabilities. A complete VVC live encoding, streaming, playback workflow has been validated for 4Kp60 and 8Kp30 HDR 10-bit video.

This contribution aims to present information about the capabilities of VVC for UHD (4K and 8K) live encoding, and the trade-off between computational complexity and bitrate savings in the context of live encoding.

The encoder is approx. 4-5 times faster than VVEnc 1.5.0 “faster” mode (but more than 15% worse in compression).

Multi-threaded encoder, no tiles used.

[JVET-AB0044](https://jvet-experts.org/doc_end_user/current_document.php?id=11957) Update on open, optimized VVC implementations VVenC and VVdeC [A. Wieckowski, J. Brandenburg, C. Bartnik, V. George, J. Güther, G. Hege, C. Helmrich, A. Henkel, T. Hinz, C. Lehmann, C. Stoffers, B. Bross, H. Schwarz, D. Marpe, T. Schierl (HHI)]

This document provides updated information on features, coding efficiency and runtime for version 1.6.1 of the open VVC software encoder VVenC released in September 2022 and version 1.6.0 of the open VVC software decoder VVdeC released in September 2022. In addition, an example implementation of a web-based player that enables to use VVdeC for VVC playback in a web browser has been made available on GitHub.

A 3rd party effort has been started in October 2022 to add VVC encoding and decoding support to FFmpeg with VVenC and VVdeC. The patches are currently under review.

Main changes for VVenC v1.6.1 since version 1.5.0 include:

* Faster presets: ~15% speedup for faster, fast, medium, ~5% for slow (1.6.0) and ~18% for slower (1.6.1)
* Minor efficiency improvements (1.6.0)
* Major RC stability improvements, especially for 1pRC, high rates (1.6.0) and low rates (1.6.1)
* Bugfix to multi-threaded encoding with tiles (1.6.1)
* Added auto-logic for correct tickspersec setting (1.6.1)
* Made apputils an internal header only lib, fixing linking problems (1.6.1)
* Optimized non-RD-optimized quantization (1.6.1)
* Various cleanups and improvements

Without QP adaptation for subjective optimization and 8 threads the following PSNR-based YUV BD-rates compared to HM-16.24 (GOP16+MCTF) as well as speedup factors compared to HM-16.24 and VTM-17.2 (GOP32+MCTF) are reported for different presets and JVET class B (HD), class A (UHD) as well as both (HD4K) test sequences:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Preset | HD |  |  | UHD |  |  | HD4K |  |  |
| PSNRYUV BD-rate vs. HM‑16.24 | Speedup vs. HM‑16.24 | Speedup vs. VTM-17.2 | PSNRYUV BD-rate vs. HM‑16.24 | Speedup vs. HM‑16.24 | Speedup vs. VTM-17.2 | PSNRYUV BD-rate vs. HM‑16.24 | Speedup vs. HM‑16.24 | Speedup vs. VTM-17.2 |
| FASTER | -9,8% | 210x | 1400x | -14,3% | 250x | 1500x | -12,3% | 230x | 1500x |
| FAST | -22,4% | 110x | 760x | -24,9% | 140x | 860x | -23,8% | 130x | 810x |
| MEDIUM | -31,9% | 32x | 210x | -34,9% | 43x | 270x | -33,5% | 38x | 240x |
| SLOW | -35,7% | 9,8x | 65x | -38,3% | 14x | 87x | -37,1% | 12x | 77x |
| SLOWER | -39,0% | 2,4x | 16x | -41,6% | 3,4x | 22x | -40,4% | 2,9x | 19x |

With QP adaptation for subjective optimization and 8 threads, the following MS-SSIM-based YUV BD-rates compared to HM-16.24 (GOP16+MCTF) as well as speedup factors compared to HM-16.24 and VTM-17.2 (GOP32+MCTF) are reported for different presets:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Preset | HD |  |  | UHD |  |  | HD4K |  |  |
| MS-SSIMYUV BD-rate vs.  HM‑16.24 | Speedup vs.  HM‑16.24 | Speedup vs.  VTM-17.2 | MS-SSIMYUV BD-rate vs. HM‑16.24 | Speedup vs.  HM-16.24 | Speedup vs. VTM-17.2 | MS-SSIMYUV BD-rate vs.  HM‑16.24 | Speedup vs.  HM-16.24 | Speedup vs. VTM-17.2 |
| FASTER | -19,2% | 210x | 1400x | -18,0% | 240x | 1700x | -18,6% | 230x | 1600x |
| FAST | -30,8% | 110x | 790x | -29,5% | 130x | 940x | -30,1% | 120x | 870x |
| MEDIUM | -37,3% | 31x | 220x | -38,3% | 42x | 290x | -37,8% | 37x | 260x |
| SLOW | -40,5% | 9,8x | 69x | -41,6% | 14x | 96x | -41,1% | 12x | 83x |
| SLOWER | -43,1% | 2,4x | 17x | -44,6% | 3,4x | 24x | -43,9% | 2,9x | 20x |

Main changes for VVdeC v1.6.0 since version 1.5.0 include various fixes and improvements:

* Added an allocator API to reduce copying to application buffers
* Fixed conformance (VPS handling, SbTMVP)
* Improved build (reduced exported symbols, fixed Xcode build)
* Minor speedups and improvements

Also Android and DASH player support.

It was reported that the “faster” preset allows live encoding of HD@60fps on a sufficiently powerful computer.

## AHG7: Low latency and constrained complexity (1)

Contributions in this area were discussed at 1100–1115 on Wednesday 26 October 2022 (chaired by JRO).

[JVET-AB0171](https://jvet-experts.org/doc_end_user/current_document.php?id=12098) AHG7: Asymmetric Deblocking at Virtual Boundaries [S. Hong, L. Wang, K. Panusopone (Nokia)]

In the current design of VVC and ECM, deblocking is disabled at virtual boundaries of GDR/recovering pictures. Virtual boundaries are therefore likely visible. This contribution proposes asymmetric deblocking at virtual boundaries, where deblocking is still disabled for pixels on the refreshed area sides of virtual boundaries, but will be performed for pixels on the non-refreshed area sides of virtual boundaries. In addition, refreshed-area pixels of virtual boundaries are used to compensate the output of deblocking of corresponding non-refreshed area pixels.

Simulations were conducted under LDB configuration on ECM 6.0. The results demonstrate that

1. Objectively, BD rates of asymmetric deblocking at virtual boundaries over no deblocking at virtual boundaries are
   * Class B:
   * Class C: 0.01% for Y, 0.08% for U and 0.04% for V
   * Class E: 0.04% for Y, 0.14% for U and -0.44% for V
   * Class D: -0.04% for Y, -0.25% for U and -0.33% for V
   * Class F:
2. Subjectively, asymmetric deblocking improves the quality around virtual boundaries.

Examples from still screen captures were shown where visual improvements were visible.

It was requested to remove the note “confidential” on the slides. It was suggested to investigate if th differences would still be visible in video.

Though not overly important for the ECM exploration (GDR is not used in CTC), this appears useful for the low latency AHG study.

Decision(SW): Adopt JVET-AB0171. Enable/disable by macro.

[JVET-AB0259](https://jvet-experts.org/doc_end_user/current_document.php?id=12188) Crosscheck of JVET-AB0171 (AHG7: Asymmetric Deblocking at Virtual Boundaries) [T. Poirier (Interdigital)] [late]

[JVET-AB0272](https://jvet-experts.org/doc_end_user/current_document.php?id=12201) Crosscheck of JVET-AB0171 (AHG 7: Asymmetric Deblocking at Virtual Boundaries) [Jack Enhorn (Ericsson)] [late]

## AHG10: Encoding algorithm optimization (3)

Contributions in this area were discussed at 1115–1215 on Wednesday 25 October 2022 (chaired by JRO).

[JVET-AB0045](https://jvet-experts.org/doc_end_user/current_document.php?id=11958) AHG10: Study of VVC spatial scalability performance [P. de Lagrange, G. Marquant, C. Salmon-Legagneur, F. Urban (InterDigital)]

This contribution reports the results VVC tests in spatial scalability mode, following work reported previously in JVET-X0202, JVET-Y0047 and JVET-Y0048.

Multiple profiles have been specified in VVC, including multilayer profiles for 4:2:0 and 4:4:4 content that give access to spatial scalability using a single decoder instance. The VVC 4:2:0 multilayer profile was tested in a spatially scalable configuration with two layers and a spatial resolution ratio of 2x and 3x both vertically and horizontally. The tests were performed using various bitrate allocations between the base layer and the enhancement layer, and the best tradeoff leading to optimal enhancement layer compression performance was investigated. Both objective and subjective evaluations have been carried out and show that with VVC, spatial scalability using dual-layer coding can be on par or outperform single-layer coding in specific conditions. It is noted that visual comparison is more often in favor of dual-layer coding than PSNR.

It was reported that in terms of visual benefit, subjective improvement over single-layer coding was also found at higher rate points (different from the PSNR based comparison).

It was reported that the enhancement layer provides quality improvement compared to the case of simple post-decoder baselayer upscaling. It was commented that it would be interesting if this could be shown in the verification test (i.e. comparison against adaptive resolution coding).

It was asked if it was tried to achieve luma/chroma balance. This was not investigated so far. It would be relevant to have a similar balance between luma and chroma for the single and dual layer cases.

Further discussion in BoG at 1600.

[JVET-AB0080](https://jvet-experts.org/doc_end_user/current_document.php?id=12007) AHG10: GOP-based RPR encoder control [K. Andersson, R. Yu, J. Ström, P. Wennersten, W. Ahmad (Ericsson)]

The VVC reference software VTM currently lacks support for selecting when to use Reference Picture Resampling (RPR). This contribution proposes a GOP-based selection method that decides when it is likely better to encode pictures in reduced resolution with RPR and when it is likely better to encode them in the source resolution. The selection is based on QP and picture self-similarity after re-scaling. The self-similarity test is only conducted on the first source picture in display order within each GOP. If the re-scaled picture is determined to have sufficient similarity with the source picture, all pictures in the GOP are encoded at reduced resolution, otherwise they are encoded in the source resolution.

The BD-rate performance (Y/Cb/Cr) compared to CTC with VTM-18.0 as anchor is as follows:

RA (luma/Cb/Cr): -0.44%/1.30%/1.26%

LDB (luma/Cb/Cr): -0.03%/0.23%/0.17%

AI (luma/Cb/Cr): -0.30%/0.62% /0.46%

The method is especially beneficial for higher resolutions and higher QPs.

It is asserted that the method, when it is selected, often can improve subjective quality compared to encoding in source resolution.

It is requested to include the method in VTM.

In v2 data for anchor was corrected. Results for the UHD set in JVET-AB0041 are also provided.

Is parallel coding still possible in RA? Yes.

RPR is used in the loop (reference of other resolution is used in case of switching)

Decision(SW): Adopt JVET-AB0080. Also include in JVET-AB2002

[JVET-AB0238](https://jvet-experts.org/doc_end_user/current_document.php?id=12167) Crosscheck of JVET-AB0080 (AHG10: GOP-based RPR encoder control) [J. Nam (LGE)] [late]

[JVET-AB0276](https://jvet-experts.org/doc_end_user/current_document.php?id=12205) Crosscheck of JVET-AB0080 (AHG10: GOP-based RPR encoder control) [C. Bartnik (HHI)] [late]

[JVET-AB0081](https://jvet-experts.org/doc_end_user/current_document.php?id=12008) AHG10: Increased length of filters used for upscaling reconstructed pictures [K. Andersson, R. Yu, L. Litwic (Ericsson)]

RPR enable usage of a reference pictures with different resolution than the current picture for inter prediction. In VTM 8-tap luma and 4-tap chroma MC filters are used for upscaling luma respectively chroma when RPR is used but also for upscaling the reconstructed picture to full resolution (enabled by UpscaledOutput=2).

This contribution proposes to increase the filter length from 8-tap to 12-tap for luma and 4-tap to 6-tap for chroma for the filters used for upscaling the reconstructed picture to full resolution.

Compared to VTM-18.0 for RPR CTC LDB (with switch of resolution every 0.5s):

2x: Y/U/V: -0.80%/-2.18%/-1.89%

1.5x: Y/U/V: -0.97%/-1.51%/-1.32%

A visual inspection revealed similar visual quality.

It is suggested to include the filters to the next release of VTM and use them for upscaling the reconstructed picture to full resolution.

Filters are different from the 12-tap filters (for luma) and 6-tap filters (for chroma) used in ECM. Proponents claim that their filters would be “less aggressive” than ECM filters, i.e. less high frequency. Current ECM filters would have higher objective gain, though.

This would be a non-normative element in VTM. It was suggested to make it switchable, i.e. retain the option of using the existing RPR filters for post-decoder upsampling. Another option would be

It was asked if the filters could also be used with the adaptive resolution coding of JVET-AB0080. It was commented by the proponents that this might not make much difference, as the gain of adaptive resolution is mainly coming at low rate/quality point.

Decision(SW): Adopt JVET-AB0081, both VTM and ECM, and allow in both packages to switch between RPR 8/4-tap filters, current ECM 12/6-tap MC filters, and the 12/6-tap filters from JVET-AB0081 by configuration option (12/6 means luma/chroma tap numbers). Also add description in JVET-AB2002 and JVET-AB2025.

[JVET-AB0277](https://jvet-experts.org/doc_end_user/current_document.php?id=12206) Crosscheck of JVET-AB0081 (AHG10: Increased length of filters used for upscaling reconstructed pictures) [C. Bartnik (HHI)] [late]

## Profile/tier/level specification (0)

This section is kept as a template for future use.

## Proposed modification of system interface (0)

This section is kept as a template for future use.

## Use cases of standards related to specific applications (3)

Contributions in this area were discussed at 1220–1325 on Wednesday 25 October 2022 (chaired by JRO).

[JVET-AB0087](https://jvet-experts.org/doc_end_user/current_document.php?id=12014) Multilayer coding use cases for broadcasting and streaming applications [S. Nemoto, S. Iwamura, A. Ichigaya (NHK)] [late]

This contribution provides information of a number of use cases of the broadcasting applications using multilayer coding techniques. In broadcasting applications, video distribution with different format (e.g. resolution and/or quality) for various type of devices are demanded since displaying capability and processing power may differ among devices. For such demand, it is suitable to distribute multiple-resolution videos using spatial scalability. Meanwhile, there has been increasing another demand for personalized video experiences in recent years. Picture-in-picture effect for additional information (e.g. sign language, multi-view videos) and open caption are often used in Japanese TV programs, however, it is not possible for the users to turn on/off those “sub-contents” in conventional broadcasting system. Using multilayer coding, main content is encoded as base layer and sub-contents are coded as enhancement layers. By doing so, it is possible to select sub-contents at the TV set side based on the user’s demand.

We believe that multilayer coding technology has a wide range of applicability, therefore developing use cases of multilayer coding would be beneficial not only for broadcasting applications but also streaming applications.

One possible application that is mentioned in the presentation for 3-layer representation would be transmission of HD/4K in a broadcast system, and 8K enhancement layer over broadband.

Content scalability (via scaling window) is also mentioned.

All fiunctionality required is supported by VVC.

[JVET-AB0274](https://jvet-experts.org/doc_end_user/current_document.php?id=12203) Encoder-only algorithms in Alibaba and City University of Hong Kong’s response to the Video Coding for Machines CfP [S. Wang, B. Li, Z. Wang, Y. Ye (Alibaba), S. Wang (City University of Hong Kong)] [late]

In the video coding for machines (VCM) CfP response from Alibaba and City University of Hong Kong, submitted as MPEG document m60737, the joint adaptive spatial temporal (JAST) system is proposed. The JAST system includes both encoder-only optimization algorithms and machine-oriented pre- and post-processing modules to achieve high compression efficiency gains for machine-analysis tasks. Due to the adaptive nature of the proposed system, some of the results reported in m60737 were obtained using just encoder-only algorithms. These results are reported in this contribution: specifically, average coding efficiency gain of -53.68% is obtained for object detection task for the FLIR dataset. This contribution also reports additional results for other datasets and other machine-analysis tasks using the encoder-only algorithms in JAST.

This is the encoder-only version of a proposal in the CfP. Currently, only results for one category (object detection); results for other categories were announced for an update of the contribution.

The method uses constant QP, but may zero out some content.

It was noted that the VVC anchor has very low MAP at the lowest rate point (<15%). It was reported that the uncompressed video would approximately be the same as the highest rate point (approx. >40%)

[JVET-AB0275](https://jvet-experts.org/doc_end_user/current_document.php?id=12204) Information about Ericsson’s response to the CfP for Video Coding for Machines [C. Hollmann, J. Ström, P. Wennersten, L. Litwic, R. Sjöberg, M. Damghanian (Ericsson)] [late]

This contribution presents Ericsson’s response to the Call for Proposals for Video Coding for Machines, which was issued by SC 29/WG 2 in April 2022. The proposed method only changes the VTM encoder, i.e., the bitstream is VVC compliant and the decoder unchanged. The proposed method is reported to work by setting QP offsets on a CTU basis without any additionally signaled metadata and no postprocessing was used to obtain the results. A BD-rate gain of around −23% is reported for the two video datasets that were used in the evaluation of the CfP. It is also reported that the method can be implemented on other standards such as HEVC as well.

Going forward, it is proposed to establish a new ad-hoc group to study non-normative encoder changes that benefit video compression for machine vision use cases.

Setup an AHG (C. Hollmann, S. Liu, S. Wang, M. Zhou)

Intent to develop a TR – Tentative name: Optimization of emcoders and receiving systems for machine analysis of coded video content.

# Low-level tool technology proposals

## AHG8: High bit rate and high bit depth coding for VVC (0)

Remove this section and discontinue AHG8?

## AHG11: Neural network-based video coding (31)

### Summary, BoG reports, and information documents

Contributions in this area were discussed at 1100–1300 on Friday 21 October 2022, and 0900-XXXX on Saturday 22 October 2022 (chaired by JRO).

[JVET-AB0023](https://jvet-experts.org/doc_end_user/current_document.php?id=12151) EE1: Summary of Exploration Experiments on Neural Network-based Video Coding [E. Alshina, F. Galpin, Y. Li, M. Santamaria, H. Wang, L. Wang, Z. Xie (EE coordinators)]

**Abstract**

This document summarizes Exploration Experiment 1 (EE1) tests performed between the JVET-AA and JVET-AB meetings to evaluate **Neural Network-based Video Coding (**NNVC) technologies, analyze their performance, and analyze their complexity aspects. In NN-based in-loop filter category BD-rate gain over AhG11 NNVC anchor goes up to **10.5%** (depending on complexity). By content adaptive post-filter gain of **5%** over AhG11 anchor was demonstrated. In Super Resolution category **~7**% BD-rate average for 4K content gain over VTM (with enabled adaptive coded picture resolution) can be achieved.

1. **Introduction**

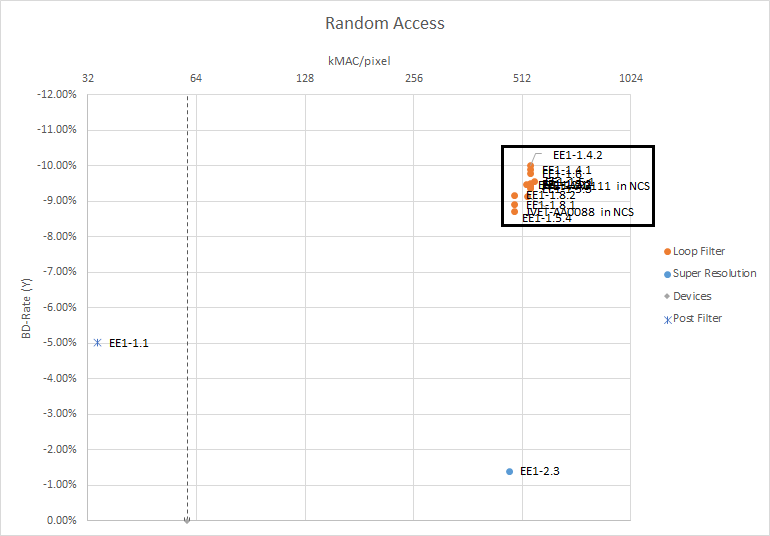
Group continues evaluation of new promising NN-based video coding technologies, answering questions and addressing suggestions from JVET members made during presentation NN-based technologies at JVET-AA meeting. Additionally, a common reference software is now available as NCS-1.0.

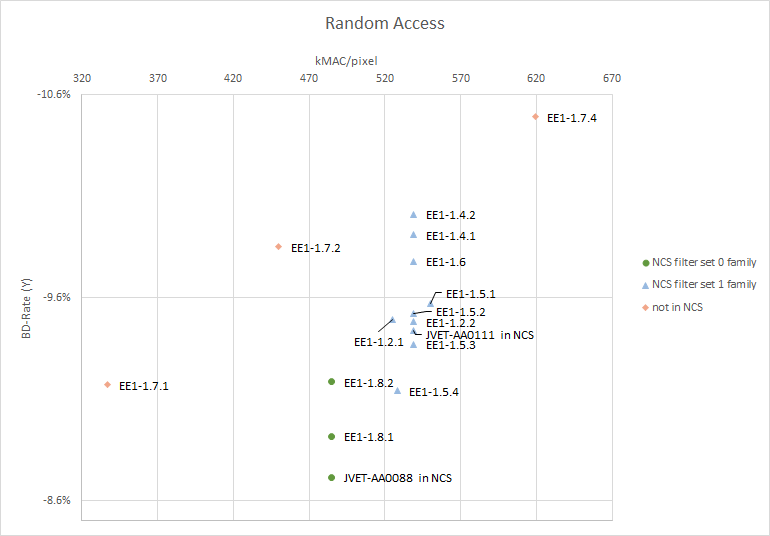
Tests were planned to be conducted in three categories: enhancement filters (in-loop or post), super-resolution and E2E NN-based video coding. Unfortunately, the test in E2E NN-based video coding category has been withdrawn.

EE1 test were actively cross-checked in this EE1 round. Cross-check status is reported in section 5 of this document.

Tests were performed on top of VTM-NNVC v2.0 / NCS 1.0.

BD-rate gain over NNVC anchor in Random Access configuration vs computational complexity (in kMAC/pxl) and memory size for Model parameters (in MB) for typical representatives of tests in this EE1 are shown on Fig. 1 and Fig. 2 respectively.



Fig. 1. Gain vs computational complexity. (a) Y-BD-rate gain in RA configuration vs kMAC/pxl, (b) in-loop filters in mode details.

Note that results for super resolution are comparing against plain VTM (not using RPR). Some of the SR proposals (EE1-2.1 and EE1-2.2) are not included, as they are either out of the complexity ranges shown, or do provide results only on 4K sequences.

1. **Exploration experiments on Enhancement filters**

***NN-filter architecture based on NCS-1.0 filter set #1******(JVET-AA0111)***

Tests in this sub-category are based on architecture JVET-AA0111 (one of two NN-based filters in NCS). Filter is referred below as **NCS-1.0 filter set #1**. **NCS-1.0 filter set #1** processes Luma separately, two chrominance components together. Additional inputs to the network: partitioning, prediction, QP map, boundary strength (BS). In total four models are used. The difference between the input samples and the NN filtered samples (residues) are scaled by the scaling factors before being added to input samples. The number of residual blocks is N=8. Worst case 682 kMAC/pixel (block-basis) and total number of parameters 6.24M

**JVET-AB0053 EE1-1.2 “EE1-1.2: NN intra model without attention and partitioning strength”**

Test was done using code of JVET-AA0111 (identical to **NCS-1.0 filter set #1**). Model was converted from pytorch to SADL (still float point), minor change in scaling values samples data and QP to float. Training is slightly different from **NCS-1.0 filter set #1** (longer).

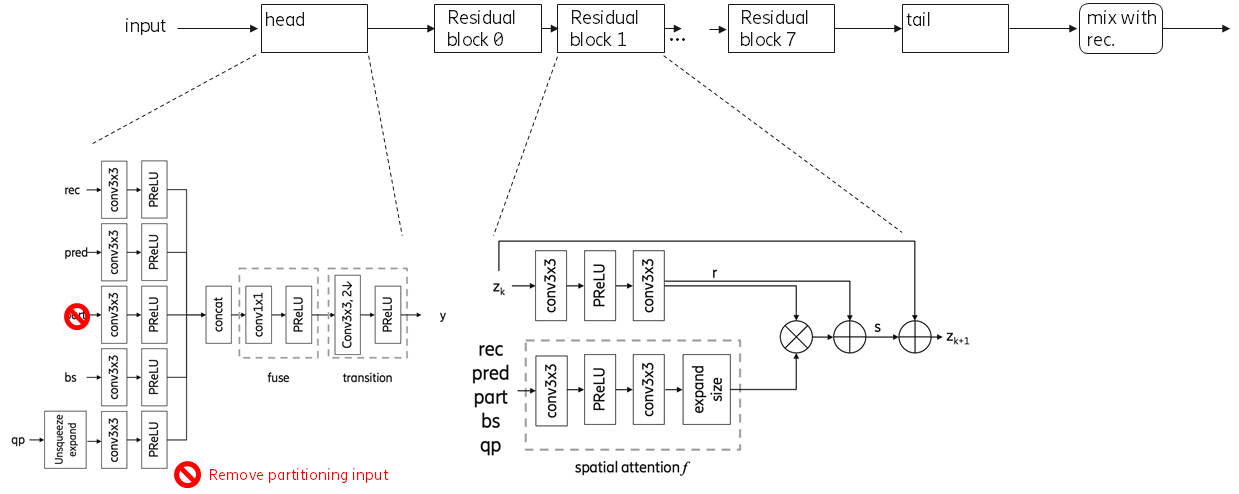


Fig. 3. Removal partitioning information from JVET-AA0111 filter design.

Sub-test performed:

EE1-1.2.2

Architecture identical to **NCS-1.0 filter set #1**, but training is by Ericsson: 0.1 % better in RA, 0.3% better in “All intra”, 0.1% better in LDB configuration. Training strategy by Ericsson works better than training from original proponent for **NCS-1.0 filter set #1**.

EE1-1.2.1

Removal partitioning information from NN-filter design (Fig. 3). Effect: negligible performance deviation (±0.02%), saving 18 kMAC/pxl (~3% of total filter complexity), 14% shorter training time.

Training process was successfully cross-checked for this test.

**JVET-AB0073 EE1-1.4 “EE1-1.4: Deep In-Loop Filter with Additional Input Information”**

In this test, it is proposed to feed collocated blocks from reference frames into the CNN-based in-loop filtering network. Another aspect is the input samples of CNN can be flipped and the output samples of CNN can be flipped back. The proposed CNN model is implemented on top of NNVC common software (NCS) using SADL int16 precision.

In EE1-1.4.1 and EE1-1.4.2, the proposed multiple frames-based CNN model is applied on two and three highest temporal layers, respectively. As for the sample flipping process, there is no difference between EE1-1.4.1 and EE1-1.4.2. BD-rate changes of {Y, Cb, Cr} on top of NCS-1.0 (filter set #1) and NNVC-2.0 are reportedly summarized as below:

EE1-1.4.1:

Compared with NCS-1.0 (filter set #1), RA: {-0.53%, 0.08%, 0.05%}, EncT 90%, DecT 92%

Compared with NNVC-2.0, RA: {-9.91%, -20.66%, -20.33%}, EncT 183%, DecT 47043%

EE1-1.4.2:

Compared with NCS-1.0 (filter set #1), RA: {-0.63%, 0.07%, 0.04%}, EncT 91%, DecT 91%

Compared with NNVC-2.0, RA: {-10.01%, -20.67%, -20.33%}, EncT 185%, DecT 46796%



Fig. 4. Collocated blocks from reference frame are fed to the **NCS-1.0 filter set #1** network.

The worst case kMAC/pixel remains unchanged while total number of parameters increases 6.24M🡪 7.8M), because one additional model is introduced (as shown in Fig. 4).

It is noted that the decoder run time (CPU, SADL) with inclusion of this NN based loop filter is longer than plain VTM encoding run time.

**JVET-AB0052 EE1-1.5 “EE1-1.5: One luma model with IPB and/or skip for filtering intra and inter luma slices”**

This test studies modification of **NCS-1.0 filter set #1** (which uses four models).In this test only three models are used.This modification is tested in EE1-1.5.4 (~0.3% loss is observed in RA and LDB). To compensate the drop, the following modification is tested: the tested models use block type information IPB (intra/uni-predicted/bi-predicted) and/or block skip information (bypassed or not) as additional input(s) on top of the inter luma model from

* **NCS-1.0 filter set #1** – natural anchor for this test
  + Worst case 682 kMAC/pixel (block-basis) and num. of para. 6.24M
* Test 1.5.1: Test the IPB**+skip** model used for intra luma and inter luma slices (as proposed in JVET-AA0090)
  + Worst case 696 kMAC/pix and num. of para. 4.70M
  + Vs. **NCS-1.0 filter set #1** BDR-Y: -0.17% RA, 0.13% LDB, and -0.13% AI.
* Test 1.5.2: Test the **IPB** model used for intra luma and inter luma slices.
  + Worst case 682 kMAC/pix and num. of para. 4.68M
  + Vs. **NCS-1.0 filter set #1** BDR-Y: -0.11% RA, 0.29% LDB, and -0.16% AI.
* Test 1.5.3: Test the **skip mode**l used for intra luma and inter luma slices.
  + Worst case 682 kMAC/pix and num. of para. 4.68M
  + Vs. **NCS-1.0 filter set #1** BDR-Y: 0.04% RA, 0.20% LDB, and -0.12% AI.
* Test 1.5.4: Test the original inter luma model from Vs. **NCS-1.0 filter set #1** used for intra luma and inter luma slices.
  + Worst case 668 kMAC/pix and num. of para. 4.67M
  + Vs. **NCS-1.0 filter set #1**, BDR-Y: 0.30% RA, 0.33% LDB, and -0.09% AI

**JVET-AB0068 EE1-1.6 “EE1-1.6: RDO Considering Deep In-Loop Filtering”**

Only encoder change for **NCS-1.0 filter set #1** (which has block level on/off decision for NN-based filter). Filtering consideration is added to RDO process in this test. Encoding run time increases ~10% in motion compensation scenarios and ~30% for all intra configuration.

Y-BD-rate vs. **NCS-1.0 filter set #1: BDR-Y**: -0.38% RA (Enc. 108%), -0.33% AI (Enc. 129%).

***NN-filter architecture based on NCS-1.0 filter set #0******(JVET-AA0088)***

Tests in this sub-category are based on architectureJVET-AA00088 (one of two NN-based filters in NCS). Filter is referred below as **NCS-1.0 filter set #0**.

**JVET-AB0083 EE1 1.8 “EE1-1.8: More refinements on NN based in-loop filter with a single model”**

Position in processing pipeline: reconstructed image after LMCS is fed into the network. Then, the output of NN filter is processed by ALF and CCALF. Content adaptation: weighted sum of CNN filter and “conventional” (de-blocking and SAO) filter output (there possible weights 1, 4/3 or 1/2), residual scaling (same as ***NCS-1.0 filter set #1***) and QP adjustment (which was the goal of the test in this EE1 round).

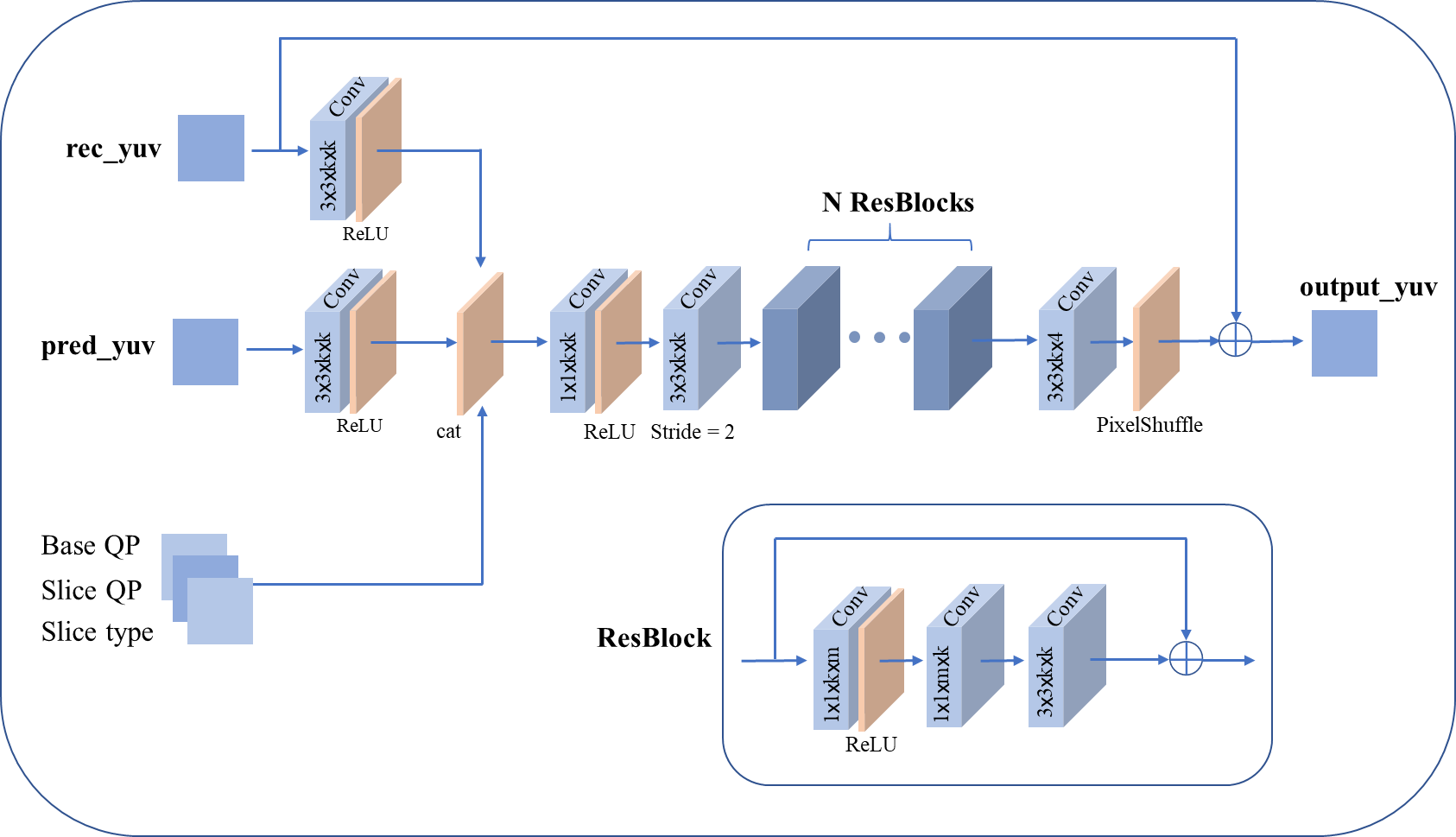


Fig. 5 Illustration of NN-based in-loop filter in test EE1-1.8 (**NCS-1.0 filter set #0** family).

Filter architecture is shown on Fig. 5. Three color components coded together, Base QP, Slice QP and slice type are extra input to the NN-module, One down-sampling CONV, N=32 Residual Blocks (RB), PixelShuffle. In Residual Block number of channels increases from 64 to 160 and reduced after activation layer. Scaling factor (adjusts the strength of the filter) is signalled in slide header.

NN-model is quantized to 16 bits integer, only one model is used. Implementation is block-based, requires 615kMAC/pxl. If implemented at picture level, complexity is 485 kMAC/pxl.

Two sub-tests have been conducted:

Test 1.8.1: the QP adjustment method as it was proposed in JVET-AA0089.

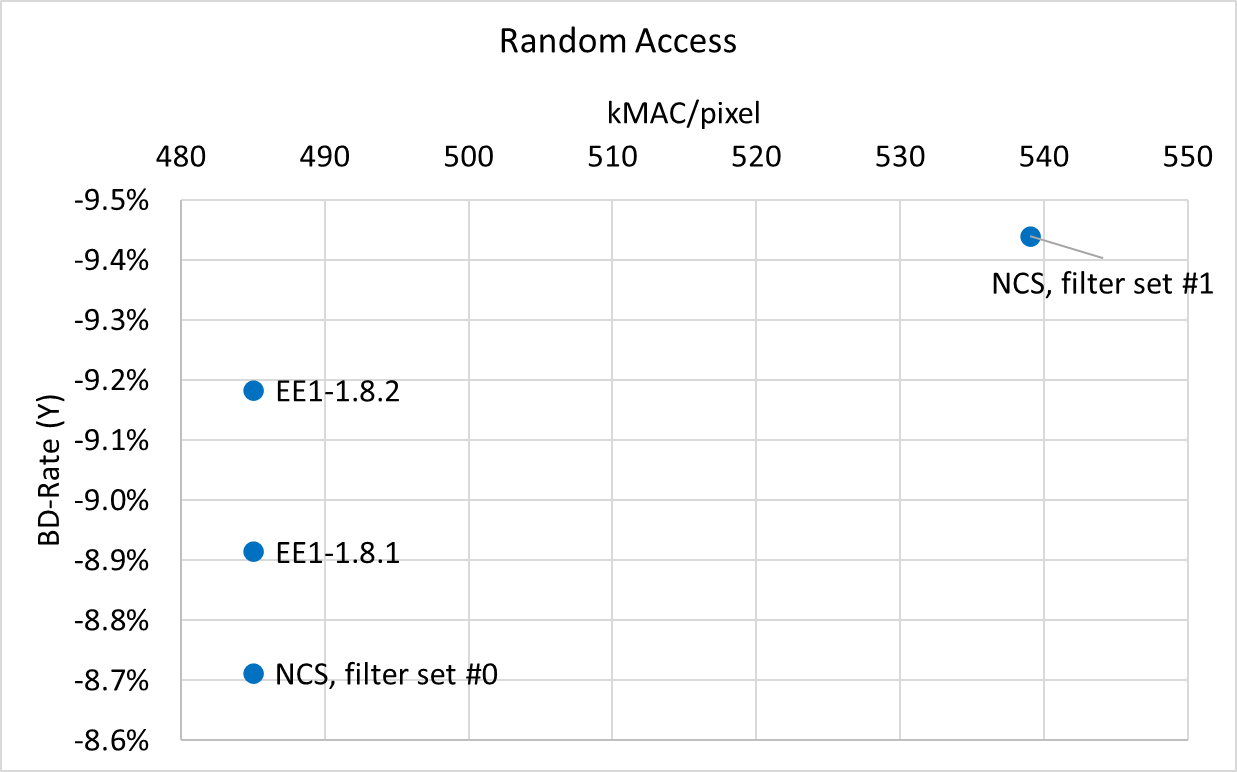
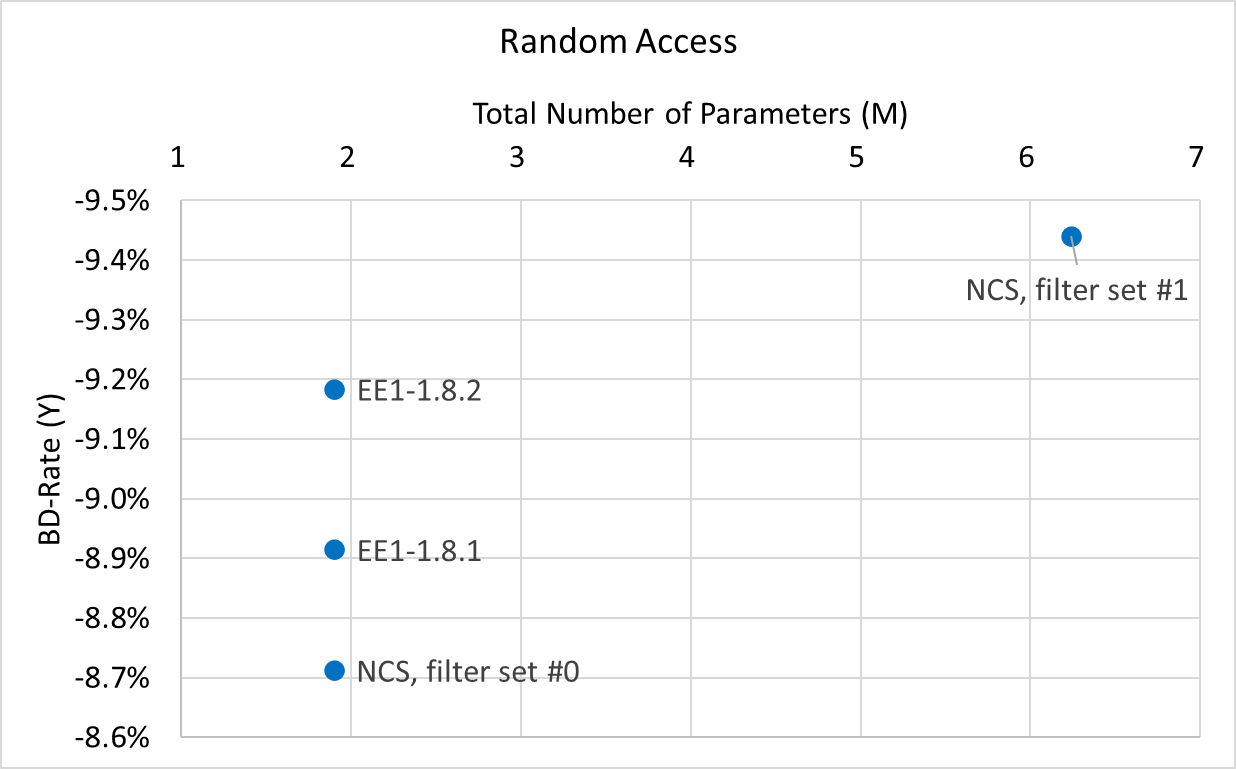
Test 1.8.2: fine tuning of the model from **NCS-1.0 filter set #0** based on Test 1.8.1.

In JVET-AA0089, a QP-adjustment method is proposed to add a small offset to the BaseQP. The offset candidates can be -5 and 5. Then, multiple results are inferred by the common one model but with the different inputs {BaseQP, BaseQP - 5, BaseQP + 5} into the BaseQP interface of the network, where the selected index is signaled at the slice-level.

On top of **NCS-1.0 filter set #0**, the QP adjustment method shows 0.21% (8.71% to 8.92%) luma gain (RA) with 11% (139% to 150%) runtime increment at encoder.

Based on Test 1.8.1, fine tuning improves performance by 0.26% (8.92% to 9.18%) luma gain (RA) / 0.1% luma gain (All Intra).

The corresponding complexity/ performance plots are shown below (Figs. 6). It seems that EE1-1.8.2 achieves a better trade-off, especially for the trade-off between total number parameters/ performance.

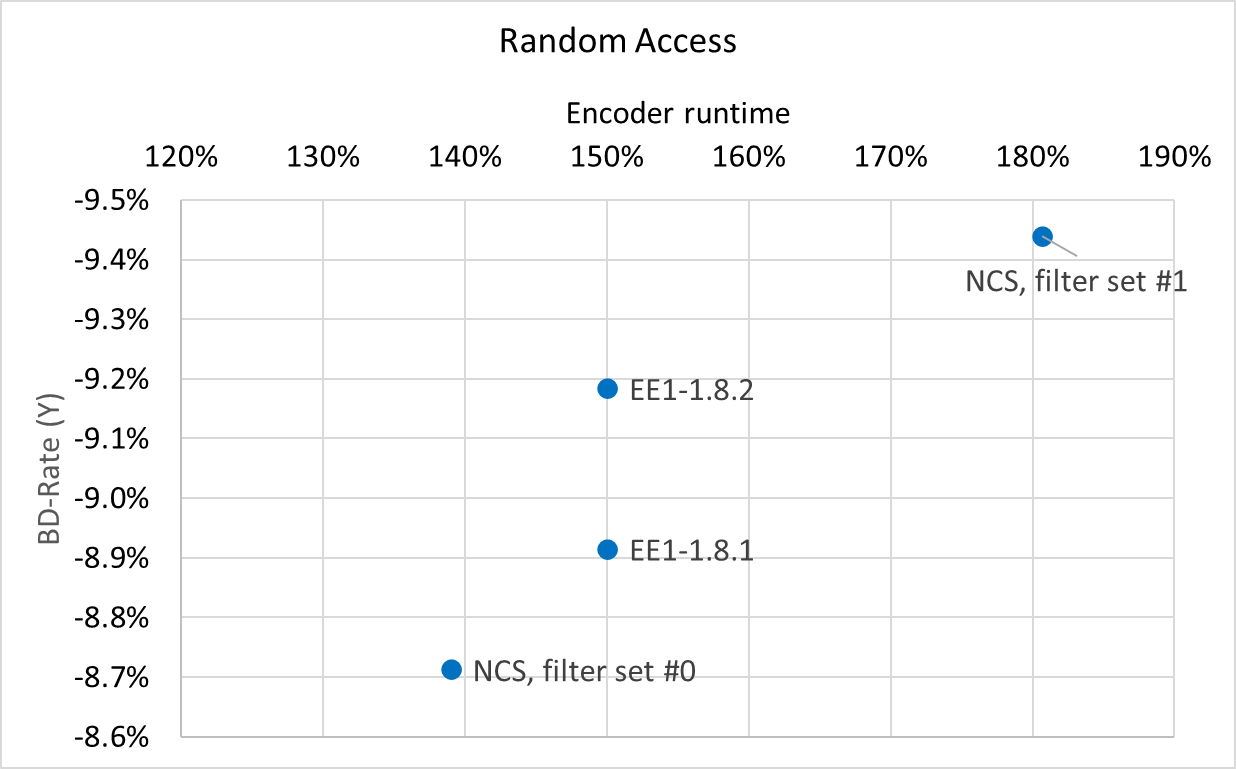


Fig. 6 Complexity performance analysis of NCS filter set 0 family tests.

Note that performance increase comes with no increase in number of parameters and kMAC/pix, but higher encoder run time.

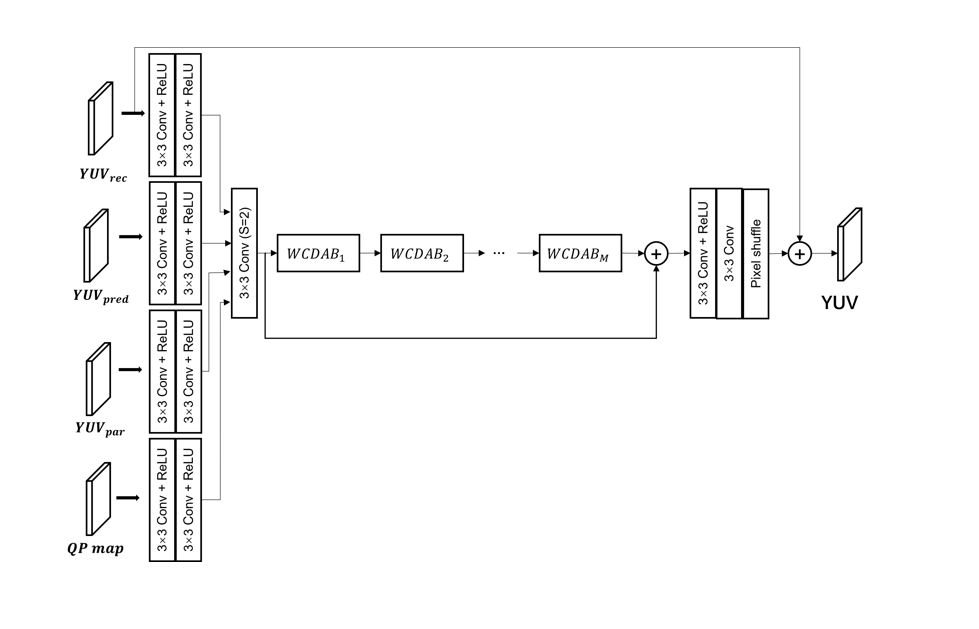
***NN-filter architecture not based on NCS-1.0***

**JVET-AB0054 EE1-1.3 “EE1-1.3: CNN Based In-Loop Filter with WCDANN”**

Filter architecture called WCDANN shown on Fig. 7. Instead of Residual Blocks which are typically used in enhancement filters WCDAB blocks are used. Number of WCDAB blocks is 4. Similar to **NCS-1.0 filter set #0,** all three color components are processed together. Prediction, partitioning are additional inputs to the network. Only “all Intra” configuration was tested. Computation complexity of this filter is 434 kMAC/pxl (lower that both filters in NCS).

In test EE1-1.3.1 “QP map is used as input, single model is used (total number of parameters in 1.3M), but in EE1-1.3.2 five models trained for different QP are used (total number of parameters is 6.5M).

Test results show that ×5 larger number of parameters and switchable models give 0.22% Y-BD rate gain. This complexity is unlikely to be justified by demonstrated moderate gain.



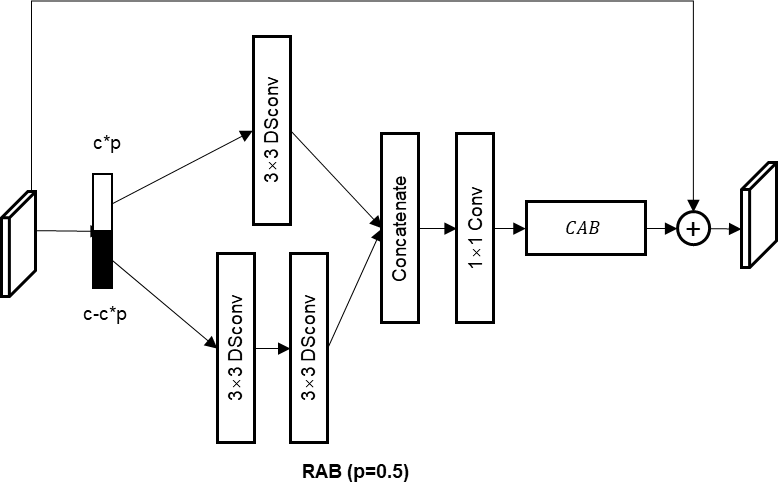
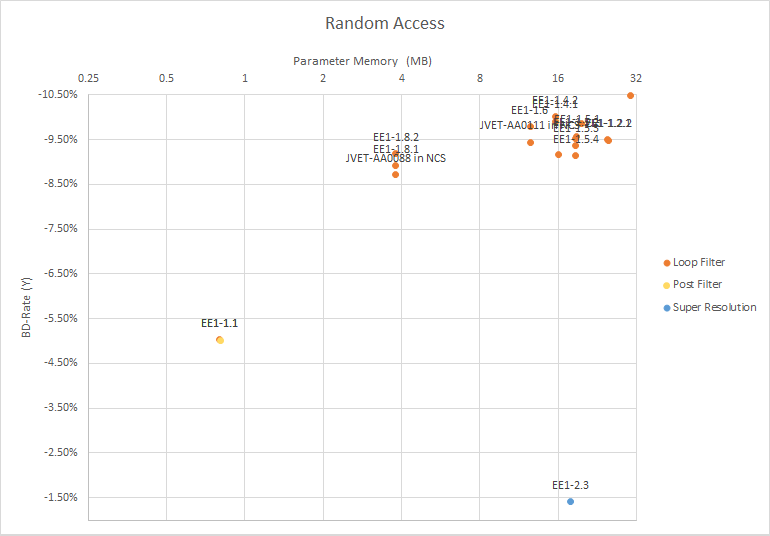


Fig. 7 Network architecture of the proposed Lighter WCDANN. The parameter M is set to 4 (EE1-1.3).

Next series of tests studies effect of weights between Luma and Chroma distortion in Loss function during training. Allocation of higher weights for Luma distortion “moves” gain from Chroma to Luma. Strangely {8,1,1} improves Luma gain with small reduction of Chroma performance, but {10,1,1} already has less Y BD-rate gain compared to {8,1,1}

|  |  |  |  |
| --- | --- | --- | --- |
| Weights  {Y,U,V} | BD-rate vs AhG11 anchor (all intra cfg) | | |
| Y | U | V |
| {6,1,1} | -4.36% | -12.59% | -11.95% |
| {8,1,1} | -4.78% | -10.95% | -10.24% |
| {10,1,1} | -4.38% | -11.11% | -9.20% |

From this study weights [8,1,1} in training can be recommended for better Luma/Chroma compression performance balance in all Intra configuration.



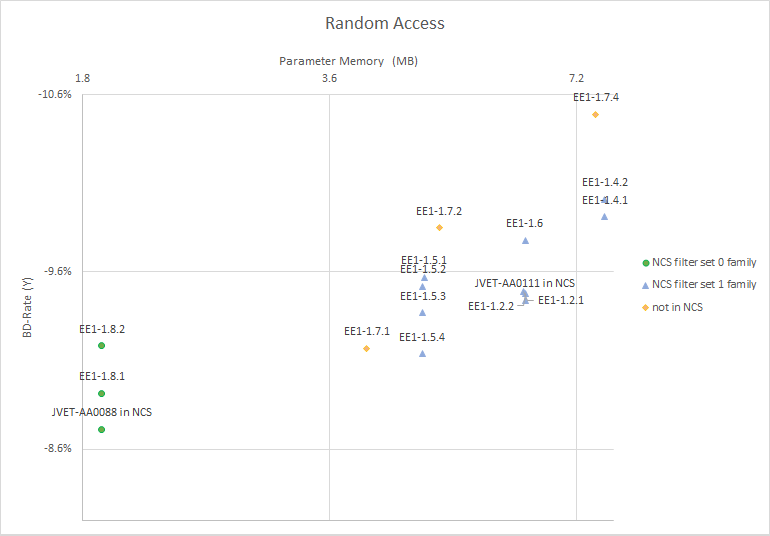


Fig. 2. Gain vs memory complexity. (a) Y-BD-rate gain in RA configuration vs total number of parameters, (b) in-loop filters in mode details.

**JVET-AB0164 EE1 1.7 “EE1-1.7: Capacity Ablation of CNN-based in-loop filtering**”

This test tries to compromise between **NCS-1.0 filter set #0** and **NCS-1.0 filter set #1,** re-using elements of both. Filter architecture is shown on FIG. 8. Additionally to reconstructed signal, prediction, partitioning, boundary strength (BS) are used. Luma and Chroma are processed separately. The goal of this study is to check performance difference between pytorch and SADL implementation. The difference on performance of float point implementation in pytorch and SADL is very low (within 0.05% Y-BD-rate in “all intra” configuration). Similary very little performance difference is observed between performance of model with float32 and quantized to int16 parameters.

Diagram

Description automatically generated

Fig. 8 NN-filter architecture of EE1-1.7

More interesting test was performed for different number of residual blocks in NN-based filter used for Luma and Chroma enhancement. The higher number of residual blocks the better performance, but, obviously the higher the complexity.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **N residual Blocks** | | **Total # params, M** | **kMAC/pxl** | **BD-rate vs AhG 11 anchor, RA** | | |
| **Luma** | **Chroma** | **Y** | **U** | **V** |
| 32 | 32 | 7.6 | 619.0 | -10.5% | -22.5% | -24.0% |
| 24 | 16 | 4.9 | 450.0 | -9.9% | -19.5% | -19.3% |
| 16 | 16 | 4.0 | 337.0 | -9.2% | -19.6% | -19.4% |
| NCS filter set #0 (N=32) | | 1.9 | 485.0 | -8.7% | -18.2% | -18.9% |
| NCS filter set #1 (N=8) | | 6.24 | 539.0 | -9.4% | -20.7% | -20.4% |

Looking at complexity/ performance plots (Fig. 9), EE1-1.7 test (16, 16) is in between the two NCS filter in BD-rate vs total number of parameters graph, and shows clearly better performance complexity trade-off on BD-rate vs kMAC/pxl graph.

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

Fig. 9. Comparison of EE1-1.7 tests with NCS filters.

There is no cross-check for this test, only all intra and partially RA test results are available for quantized model, but this test looks like very promising candidate for unified NN-based filter design. Can be recommended for training stage cross-check in the next meeting circle.

It is noted that targeting for a unified approach which takes up the elements from both filter sets #0 and #1 would be highly desirable. When the training crosscheck is performed in the upcoming EE cycle, it should also be considered if some of the elements that brought further improvements on top of #0 and #1 might also be beneficial in the unified approach. Investigate in EE

***NN-filter based post-filters***

**JVET-AB0048 EE1-1.1 “EE1-1.1: Content-adaptive post-filter with SADL inference and signalling of NN post-filter characteristics and activation SEI messages**”



Fig. 10. Post filter EE1-1.1 architecture.

It is reported that the BD-rate gain over NNVC 2.0 is 5.0% (Y), 20% (U), 17% (V) in RA configuration.

Extra time required for overfitting is comparable with one Intra Period Segment coding in RA configuration

**NN-based super-resolution**

All filters in this category work as post-filter. Some are combined with adaptive coding picture resolution. This is why an additional reference for tests in this category is [JVET-Z0065](https://jvet-experts.org/doc_end_user/current_document.php?id=11500) (by LGE) in which scaling factor for coded video was adaptively selected between 1.0 (full size coding), 2.0 (coding at quarter resolution) and 1.5 (coding of down sampled by ratio 1.5 video). Selection of scaling factor was performed at GOP basis, RPR resampling filters were used (no neural network based coding is involved.). Thanks to the wise strategy of scaling factor selection, this test shows gain in average over AhG11 anchor. This test is listed in EE1 summary table as additional reference.

For some tests in this category results for 4K resolutions only available. The comparison of test results for 4K only sequences (all intra cfg)

is shown in table below.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| SR techniques | Y-BD-rate | U-BD-rate | V-BD-rate | Total Params, 106 | kMAC/pxl | 10% rate matching |
| [RPR](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0065-v1.zip) ([JVET-Z0065](https://jvet-experts.org/doc_end_user/current_document.php?id=11500)) | **-2.18%** | 7.52% | 6.15% | 0 | 0 | YES |
| [EE1-2.1](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0065-v1.zip) | **-9.24%** | 8.82% | -16.39% | 12.5 | 854 | NO |
| [EE1-2.2](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0065-v1.zip) | **-8.50%** | 18.77% | -12.61% | 31.0 | 2044 | NO |
| EE1-2.3 | **-3.47%** | 0.61% | -0.15% | 4.5 | 469 | YES |

Note that “10% rate matching” means that the downsampled video used in the proposed technology shall not deviate by more than 10% from the full resolution anchor, as otherwise the BD rate numbers might be misleading. If that requirement is not met, the quality of the subsampled video should be increased, for which different options exist: Change of QP, use different subsampling.

Training crosscheck of EE1-2.3 in next EE.

From a more close review of PSNR vs. bitrate graphs, it appears that the most severe problem interpreting the results is that the quality at the lowest rate is significantly lower than for the full resolution anchor. The QP of the low resolution should be decreased such that the quality of upsampled SR and full resolution anchor matches.Further investigate 2.1 and 2.2 (replaced by improved version JVET-AB0093) in the EE, in particular

- Adapt for rate/quality matching with high resolution anchors

- Perform adaptation for enabling/disabling low resolution coding, depending on rate point, GOP basis, etc. (similar as 2.3 has been doing) – this will also allow testing with other classes (not only A)

- Reducing complexity / lower number of models.

It is also noted that all technologies in this experiment (including RPR) are operated as post processing.

In longer term, it would be interesting to investigate subsampling ratios other than 2 with neural networks. This would be very important in comparison with RPR (note that JVET-AB0080 and JVET-AB0102 propose usage of subsampling 4/5, 2/3, 1/2 to be selectable with RPR).

**JVET-AB0076 EE1 2.1 “EE1-2.1: RPR-Based Super-Resolution Guided by Partition Information”**

Unfortunately, for this test:

* no decoding run time provided,
* results shown only for 4K sequences,
* bit-rates are significantly different from anchor (BD-rate is not very reliable).

NN-based processing operates after RPR up-sampling (de facto works as enhancement filter, trained to reduce RPR artifacts).

For “all intra configuration” **9.2%** Y-BD-rate gain over NNVC2.0 anchor is reported. In comparison **2.2%** can be achieved by adaptive resolution selection and RPR re-sampling ([JVET-Z0065](https://jvet-experts.org/doc_end_user/current_document.php?id=11500)).

The proposed network is composed of four parts: feature extraction, reference information generation, mutual information processing and reconstruction (Fig. 11). Specifically, the feature extraction part consists of three convolutional layers, which are used to extract features of the inputs. The convolutional layer is followed by a ReLU activation function. The inputs are the predicted frame and chroma component after RPR up-sampling.

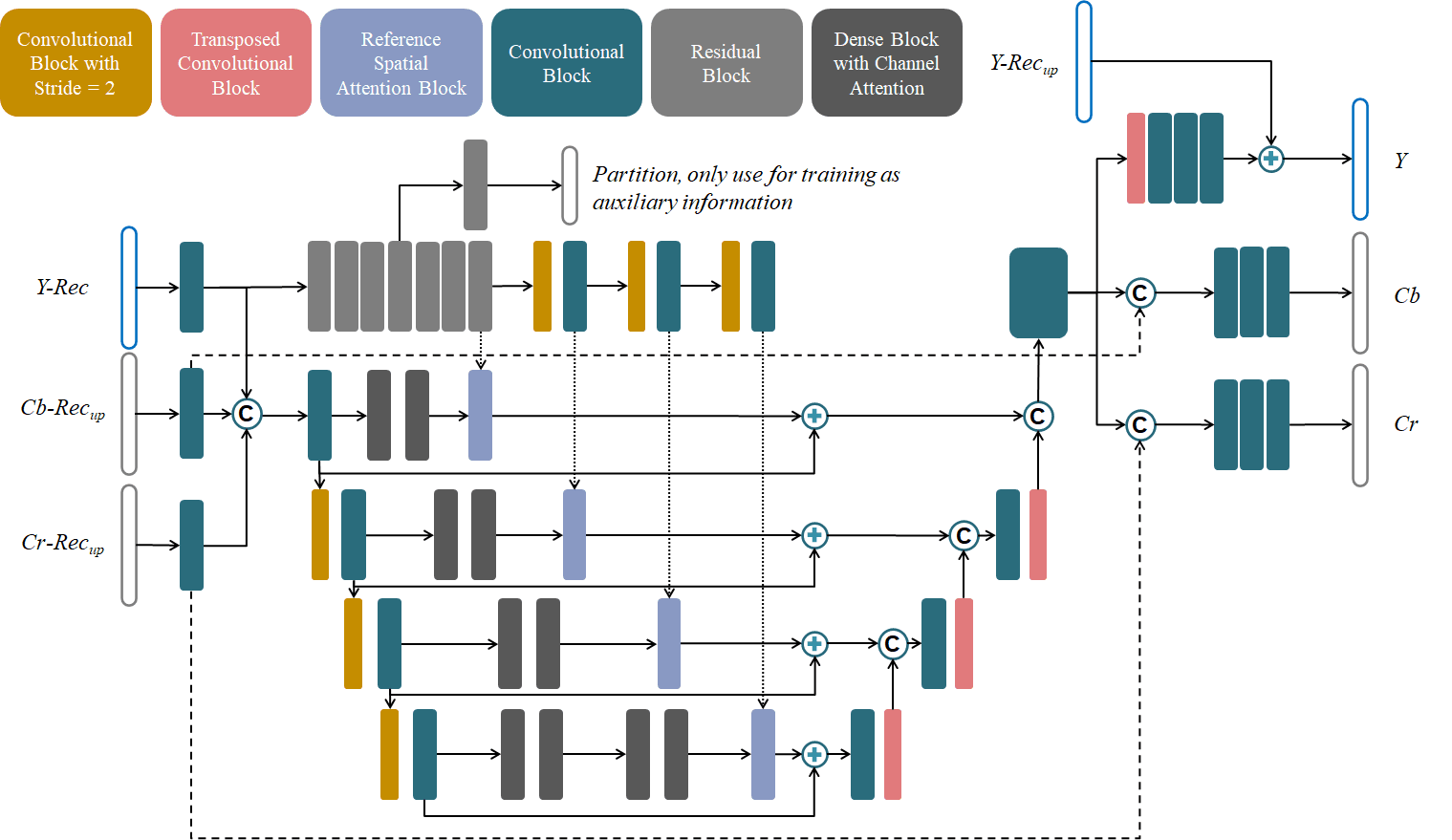


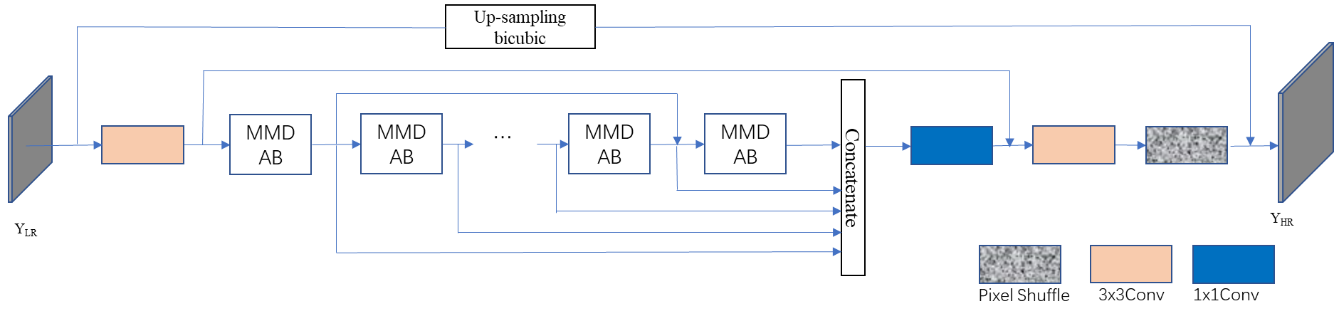
Fig. 11 Illustration of the proposed network architecture for Super-resolution guided by partition information (EE1-2.1).

The reference information generation (RIG) part consists of eight residual blocks. The first four blocks are performed for predicting the CTU partition information from reconstructed frame with an extra block, while the last three blocks are used for reference information generation. Sequentially, reference information features are input to several convolutional layer sets to generate different scales features as the input of the reference feature attention module. Each layer set consists of a convolutional layer with stride 2 and a convolutional layer followed by one ReLU.

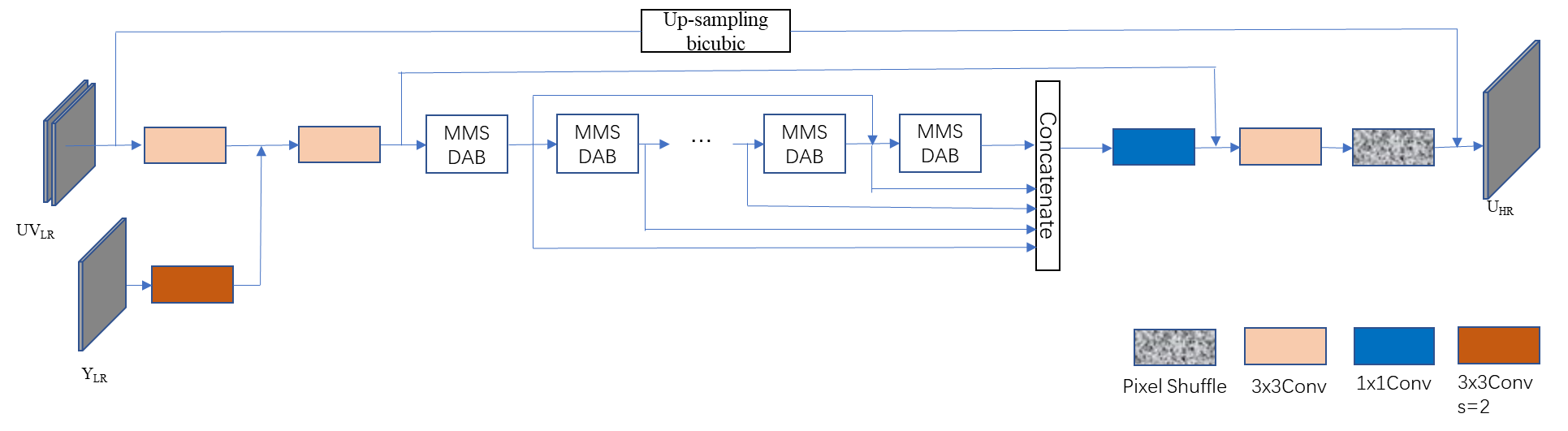
**JVET-AB0077 EE1 2.2 “EE1-2.2: CNN Filter for Super-Resolution with RPR functionality in VVC”**

In opposite to previous test (EE1-2.1) no RPR resampling is used here, reconstructed signal up-sampled using NN-based algorithm, called MMSDANet (Fig. 12). Y component is processed w/o extra information. U and V channels are up-sampled using all three low resolution components Y, U and V as inputs.

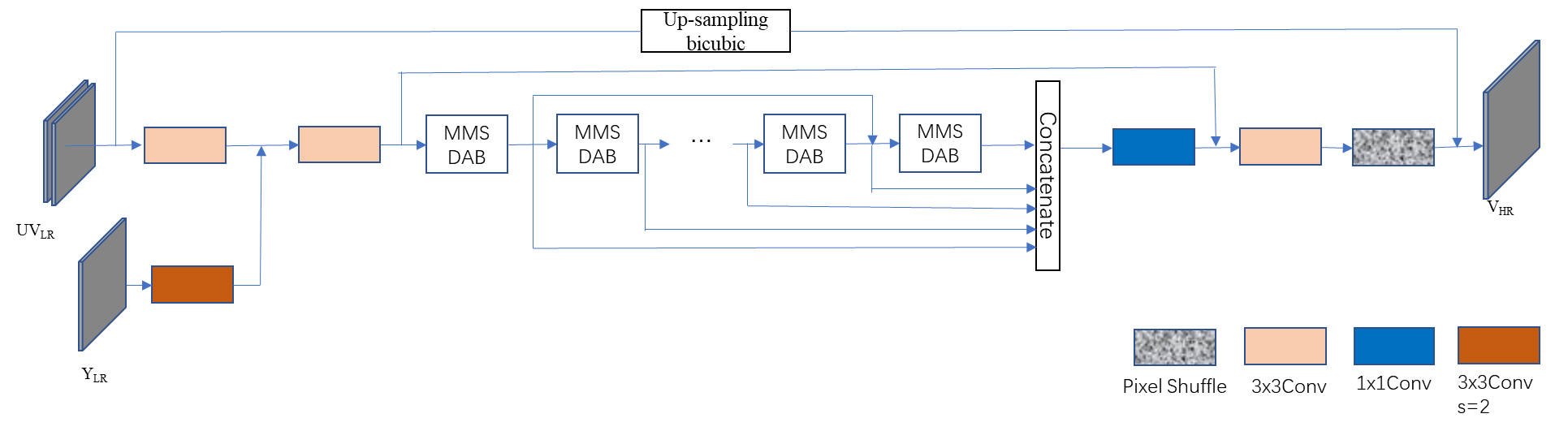
MMSDANet consists of 8 MMSDABs, three convolutional layers, one concatenate layer, one shuffle layer and five shortcut connections. Except for the 1x1 and last convolutional layers, all convolutional layers and MMSDABs have 64 input and output channels. 1x1 convolution is used to reduce the number of channels and reduce the complexity of the operation. The output channel of the last convolutional layer is 4 so that the pixel shuffle layer can do 2x up-sampling. Since the input and output resolutions are different, the input needs to be up-sampled when applying global connections.



Y channel



U channel



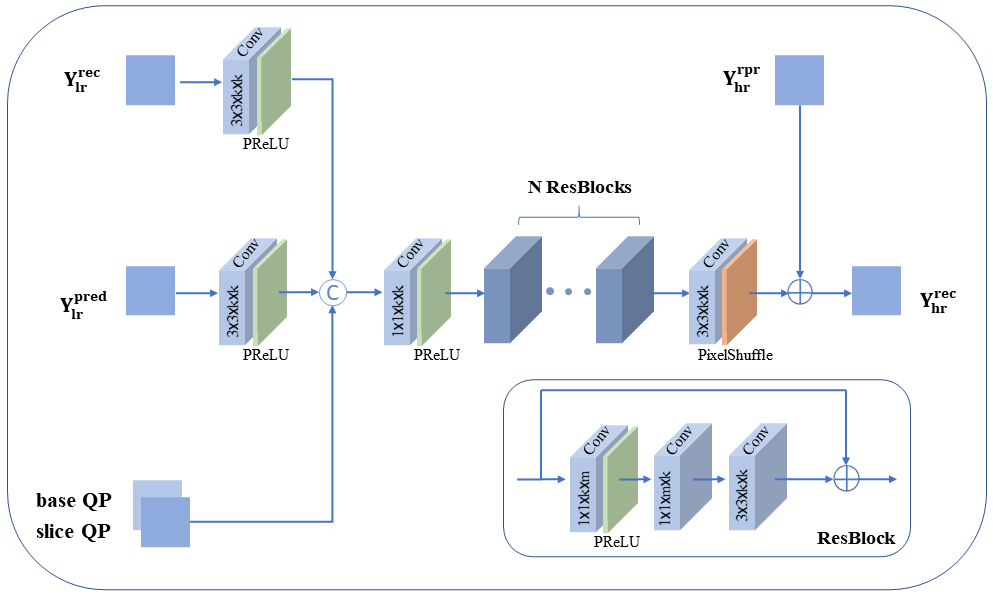
V channel

Fig. 12 Illustration of the proposed MMSDANet in EE1-2.2.

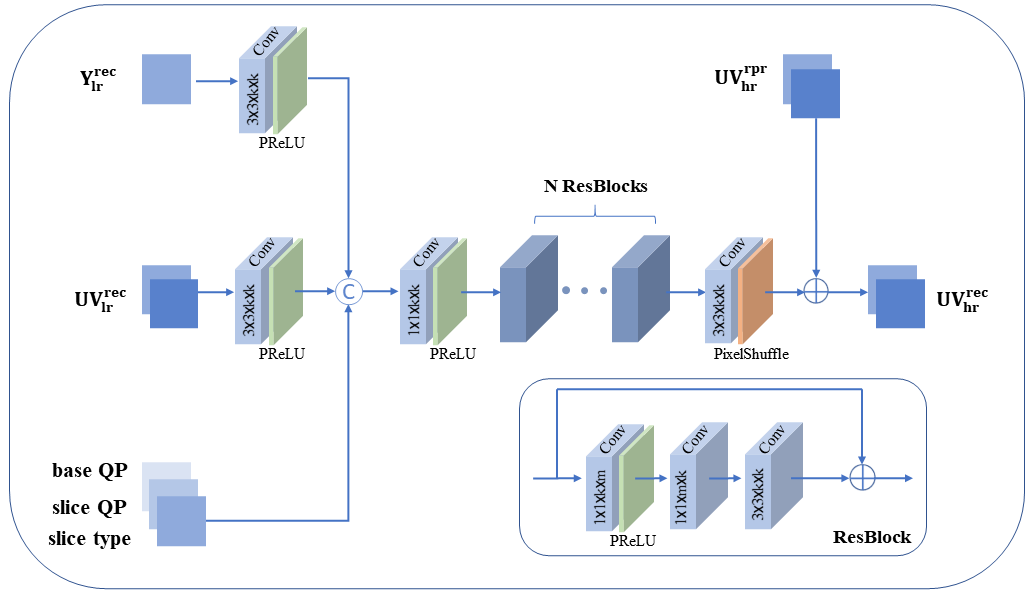
**JVET-AB0084 EE1 2.3 “A CNN-based Super Resolution Method with GOP Level Adaptive Resolution”**

In this test similarly to [JVET-Z0065](https://jvet-experts.org/doc_end_user/current_document.php?id=11500) scaling factor for coded video was adaptively selected between 1.0 (full size coding), 2.0 (coding at quarter resolution) and 1.5 (coding of down sampled by ratio 1.5 video). Selection of scaling factor was performed at GOP basis. Additionally to [JVET-Z0065](https://jvet-experts.org/doc_end_user/current_document.php?id=11500) NN-based enhancement is used to improve RPR up-sampling.

NN-based re-sampler fed (for luma) with low resolution reconstructed signal, prediction, base QP and slice QP information (for luma). UV component are up-sampled jointly and use Luma low resolution signal data as extra input to NN-based up-sampler.



***for Luma***



***for Chroma***

Fig. 13 The super resolution network architecture in EE1-2.3.

1. **End-to-End AI video coding**

Test EE1-3.1 was withdrawn by proponent.

1. **Cross-check status**

Performance deviation was observed in almost all cases (even when CPU inference is used). Several reasons for performance deviation were identified: computing environment difference, GCC compiler version difference, float point operations.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Test | proposal | tester | Cross-check | Cross-checker | Comment |
| EE1-1.2 | JVET-AB0053 | Ericsson | [JVET-AB0063](https://jvet-experts.org/doc_end_user/current_document.php?id=11978) | Nokia | **Training** and inference were crosschecked; training time reduction 14% is confirmed for proposed simplification. Performance difference up to 0.1%, computing environments are different |
| EE1-1.3 | JVET-AB0054 | Xidian/OPPO | JVET-AB0064 | Nokia | AI inference was cross-checked. Results match proponent ones. Very small differences (0.0x%) likely due to float inference. |
| EE1-1.5 | JVET-AB0052 | Ericsson | [JVET-AB0089](https://jvet-experts.org/doc_end_user/current_document.php?id=12016) | Tencent | full test results are not an exact match due to different GCC versions |
| EE1-1.6 | JVET-AB0068 | Bytedance | [JVET-AB0137](https://jvet-experts.org/doc_end_user/current_document.php?id=12064) | Ericsson | **Training** and inference were crosschecked, training by cross-checker was faster than reported by proponent.  SW was modified (with help from proponent) to be able to verify training  Performance difference is 0.0x% (both RA and LDB cfg). |
| EE1-1.8 | JVET-AB0083 | Tencent/OPPO | [JVET-AB0088](https://jvet-experts.org/doc_end_user/current_document.php?id=12015) | Ericsson | To match performance exactly same GCC complier shall be used (uses GCC 7.4.0 for the full simulation results and GCC 8.2.0 for class B under LDB configuration) |
| EE1-2.1 | JVET-AB0076 | Xidian/OPPO | [JVET-AB0105](https://jvet-experts.org/doc_end_user/current_document.php?id=12032) | Tencent | Results match proponent ones. Small differences in RA (likely due to float inference). |
| EE1-2.2 | JVET-AB0077 | Xidian/OPPO | [JVET-AB0106](https://jvet-experts.org/doc_end_user/current_document.php?id=12033) | Tencent | Results match proponent ones. Small differences in RA (likely due to float inference). |
| EE1-2.3 | JVET-AB0084 | Tencent/LGE | [JVET-AB0097](https://jvet-experts.org/doc_end_user/current_document.php?id=12024) | Ericsson | Results match to proponent for all intra, for RA 0.1% better than reported by proponent |
| EE1-3 | withdrawn | Xidian/OPPO |  | Dolby | Bug was discovered, to be fixed |

It was suggested that the next version of CTC and evaluation procedures (JVET-AB2016) should contain some rules about the deviations that would be allowed in training and inference crosschecks. This should however still allow deviations beyond some threshold if the reasons are sufficiently understood.

1. **Conclusions**

* Training
  + Training was cross-checked for tests EE1-1.2 and EE1-1.6.
  + Training strategy from EE1-1.2.2 ([JVET-AB0053](https://jvet-experts.org/doc_end_user/current_document.php?id=11968)) is more successful (faster, slight reduction of kMAC/pix) for **NCS-1.0 filter set #1** than one currently in NCS.
  + Weights {8,1,1} for {Y,U,V} distortion in loss-function was found in EE1-1.3.2 ([JVET-AB0090](https://jvet-experts.org/doc_end_user/current_document.php?id=12017)) to provide best balance for Luma / Chroma performance.
* Encoder only modification
  + Additional 0.4% / 0.3% gain (RA/LDB) with 8% / 29% encoding run time increase can be achieved if NN-based filter is considered in RDO (EE1-1.6) ([JVET-AB0068](https://jvet-experts.org/doc_end_user/current_document.php?id=11983)). Investigate training cross-check with similar methods for filter #0 (as suggested in JVET-AB0146) and EE1.1-7 in next EE. Decision(SW): Include the method from JVET-AB0068 in the next version of NCS, but don’t enable it in CTC (see further discussion under JVET-AB0146 on the latter aspect).
* NN-filter design
  + Removing partitioning information from **NCS-1.0 filter set #1** filter design (EE1-1.2.1 ([JVET-AB0053](https://jvet-experts.org/doc_end_user/current_document.php?id=11968)) helps reduction of training time 14%, computation complexity reduction and comes w/o performance degradation. Decision: Adopt the modified filter design (integerized version), and the software for modified training strategy to next version of NCS. It was further suggested to investigate a corresponding modification/simplification for chroma filters in next EE.
  + Gain of 0.5~0.6% in RA can be achieved if **NCS-1.0 filter set #1** extended to temporal filtration and the sample flipping is used. Complexity measured by EncT, DecT, and kMAC/pixel remain unchanged while total number of parameters increases by 25%, because one additional model is introduced (EE1-1.4.1, 1.4.2) ([JVET-AB0073](https://jvet-experts.org/doc_end_user/current_document.php?id=11988)). Perform training verification, and report gains for the two elements (extending filter input using samples from reference frames, sample flipping) in next EE.
  + In test EE1-1.5 variant EE1-1.5.2 (**IPB** model used for intra luma and inter luma slices in **NCS-1.0 filter set #1**) reduces of overall number of parameters 6.24 🡪4.68 M, shows same the worst case 682 kMAC/pix and similar performance as **NCS-1.0 filter set #1.** Is recommended for consideration in EE1 (training cross-check of 1.5.2, only luma, both FP and INT16 versions should be verified; usage of a similar approach with only one model for chrome to be investigated).
  + EE1-1.7 [JVET-AB0164](https://jvet-experts.org/doc_end_user/current_document.php?id=12091) shows promising performance/complexity trade-off, the model n=(24,16) is recommended for training verification in the next EE1 round. Both FP and INT16 versions should be verified.
  + EE1-1.8[JVET-AB0083](https://jvet-experts.org/doc_end_user/current_document.php?id=12010) has two aspects: modification of QP input to the network without changing the model gives approx. 0.2% for RA and 0.5% for LB (Decision: Adopt JVET-AB0083 test 1.8.1; 1.8.2 is an improvement of **NCS-1.0 filter set #0,**  shows 0.3%/0.2% better performance **NCS-1.0 filter set #0** in RA/LB having the same decoder complexity. Training crosscheck of 1.8.2 in the next EE.
* Post-filter SEI
  + Successful usage of the post-filter characteristics and activation SEI messages was demonstrated in test EE1-1.1, 5% (Y) gain in RA can be achieved. Training crosscheck is currently running, it was initially expected to be completed during the meeting. This was not the case, therefore the training crosscheck of the EE1-1.1 needs to be completed in the next round of EE1.

1. **Recommendation**

* Block base implementation of NN-based filters is more realistic, it is recommended to refine AhG 11 complexity assessment methodology and report kMAC/pxl per block (all EE1 proponents are doing so already).
* Discuss performance deviation between cross-checker and proponent, ideally specify maximum allowed performance deviation to be used for future NNVC work in JVET.

### EE1 contributions: Neural network-based video coding (11)

Beyond the EE report, selected contributions in this area were discussed at 0900–1200 on Saturday 22 October 2022 (chaired by JRO).

For actions to be taken, see section 5.2.1

[JVET-AB0048](https://jvet-experts.org/doc_end_user/current_document.php?id=11963) EE1-1.1: Content-adaptive post-filter with SADL inference and signalling of NN post-filter characteristics and activation SEI messages [M. Santamaria, R. Yang, F. Cricri, J. Lainema, H. Zhang, R. G. Youvalari, M. M. Hannuksela (Nokia)]

No need for presentation, sufficiently covered in EE summary report.

[JVET-AB0052](https://jvet-experts.org/doc_end_user/current_document.php?id=11967) EE1-1.5: One luma model with IPB and/or skip for filtering intra and inter luma slices [D. Liu, J. Ström, M. Damghanian, P. Wennersten, K. Andersson (Ericsson)]

This contribution reports the results of EE1-1.5, where one luma model with block type information and/or block skip information is used for NN-filtering of both intra luma and inter luma slices as described in JVET‑AA0090, instead of two separate luma models for intra and inter as in JVET-AA0111. The tested models use block type information IPB (intra/uni-predicted/bi-predicted) and/or block skip information (bypassed or not) as additional input(s) on top of the inter luma model from JVET‑AA0111. Comparing to the worst-case complexity (kMAC/pixel) in JVET‑AA0111, Test 1.5.1 the IPB+skip model is 2.0% higher, Test 1.5.2 the IPB model and Test 1.5.3 the skip model are of the same complexity, and Test 1.5.4 the original model is 2.0% lower than JVET‑AA0111. All the tests reduce the number of models from four to three, which saves the effort of training one extra model, and reduce the total number of parameters from 6.24M to 4.67~4.70M. It is reported that, for RA configuration, the IPB+skip model gives 0.17% higher luma gain than JVET‑AA0111, the IPB model gives 0.11% higher luma gain, the skip model gives 0.04% luma loss, and the original model (non-IPB/skip) gives 0.30% luma loss.

* Anchor JVET-AA0111 (equivalent to NCS-1.0 with NN filter set 1)
  + Worst case 682 kMAC/pixel (block-basis) and num. of para. 6.24M
* Test 1.5.1: Test the IPB+skip model used for intra luma and inter luma slices as proposed in JVET-AA0090.
  + Worst case 696 kMAC/pix and num. of para. 4.70M
  + Vs. NNVC-2.0, BDR-Y: -9.57% RA, -8.30% LDB, and -7.37% AI.
  + Vs. JVET-AA0111, BDR-Y: -0.17% RA, 0.13% LDB, and -0.13% AI.
* Test 1.5.2: Test the IPB model used for intra luma and inter luma slices.
  + Worst case 682 kMAC/pix and num. of para. 4.68M
  + Vs. NNVC-2.0, BDR-Y: -9.52% RA, -8.15% LDB, and -7.39% AI.
  + Vs. JVET-AA0111, BDR-Y: -0.11% RA, 0.29% LDB, and -0.16% AI.
* Test 1.5.3: Test the skip model used for intra luma and inter luma slices.
  + Worst case 682 kMAC/pix and num. of para. 4.68M
  + Vs. NNVC-2.0, BDR-Y: -9.37% RA, -8.23% LDB, and -7.36% AI.
  + Vs. JVET-AA0111, BDR-Y: 0.04% RA, 0.20% LDB, and -0.12% AI.
* Test 1.5.4: Test the original inter luma model from JVET-AA0111 (without IPB or skip) used for intra luma and inter luma slices.
  + Worst case 668 kMAC/pix and num. of para. 4.67M
  + Vs. NNVC-2.0, BDR-Y: -9.14% RA, -8.11% LDB, and -7.33% AI
  + Vs. JVET-AA0111, BDR-Y: 0.30% RA, 0.33% LDB, and -0.09% AI

Only one model for intra and inter luma, but the block type information (I,P,B) is input to the model. In some of the versions investigated, skip information is also input.

For chroma, still two models (from JVET-A0111) are used.

Integer conversion not yet done, but results could be provided during the meeting according to proponents.

The benefit of using skip is small. In terms of performance/complexity, 1.5.2 (only using IPB) is preferred by the proponents.

The new models were re-trained from scratch.

[JVET-AB0089](https://jvet-experts.org/doc_end_user/current_document.php?id=12016) Crosscheck of JVET-AB0052 (EE1-1.5: One luma model with IPB and/or skip for filtering intra and inter luma slices) [L. Wang (Tencent)] [late]

[JVET-AB0053](https://jvet-experts.org/doc_end_user/current_document.php?id=11968) EE1-1.2: NN intra model without attention and partitioning strength [J. Ström, D. Liu, K. Andersson, P. Wennersten, M. Damghanian, R. Yu (Ericsson)]

No need for presentation, sufficiently covered in EE summary report.

[JVET-AB0063](https://jvet-experts.org/doc_end_user/current_document.php?id=11978) Cross-check of JVET-AB0053 (EE1-1.2: NN intra model without attention and partitioning strength) [M. Santamaria, F. Cricri (Nokia)]

[JVET-AB0054](https://jvet-experts.org/doc_end_user/current_document.php?id=11969) EE1-1.3: CNN Based In-Loop Filter with WCDANN [H. Zhang, C. Jung (Xidian Univ.), D. Zou, M. Li (OPPO)]

No need for presentation, sufficiently covered in EE summary report.

[JVET-AB0064](https://jvet-experts.org/doc_end_user/current_document.php?id=11979) Cross-check of JVET-AB0054 (EE1-1.3: CNN Based In-Loop Filter with WCDANN) [M. Santamaria, F. Cricri (Nokia)] [late]

[JVET-AB0068](https://jvet-experts.org/doc_end_user/current_document.php?id=11983) EE1-1.6: RDO Considering Deep In-Loop Filtering [J. Li, Y.Li, K. Zhang, L. Zhang (Bytedance)]

No need for presentation, sufficiently covered in EE summary report.

[JVET-AB0137](https://jvet-experts.org/doc_end_user/current_document.php?id=12064) Crosscheck of JVET-AB0068 (EE1-1.6: RDO Considering Deep In-Loop Filtering) [J. Ström (Ericsson)]

[JVET-AB0073](https://jvet-experts.org/doc_end_user/current_document.php?id=11988) EE1-1.4: Deep In-Loop Filter with Additional Input Information [Y. Li, K. Zhang, L. Zhang (Bytedance)]

No need for presentation, sufficiently covered in EE summary report.

[JVET-AB0241](https://jvet-experts.org/doc_end_user/current_document.php?id=12170) Crosscheck of JVET-AB0073 (EE1-1.4: Deep In-Loop Filter with Additional Input Information) [S. Eadie (Qualcomm)] [late] [miss]

[JVET-AB0076](https://jvet-experts.org/doc_end_user/current_document.php?id=11991) EE1-2.1: RPR-Based Super-Resolution Guided by Partition Information [Q. Han, C. Jung (Xidian Univ.), Y. Liu, M. Li (OPPO)]

This contribution reports the EE1-2.1 test results, which are based on JVET-AA0076, presenting a CNN filter for RPR-based super-resolution guided by partition information (AA0076 filter). In the EE1-2.1 test, VTM-11.0\_NNVC-2.0 and JVET-Z0065 are adopted as anchors for comparison, and the performance of the AA0076 filter over each anchor under AI and RA configurations are as follows:

VTM-11.0\_NNVC-2.0: {-9.24%, 8.82%, -16.39%} and {-4.36%, -0.82%, -11.34%} BD-rate changes under AI and RA configurations for {Y, Cb, Cr} channels,

JVET-Z0065: {-5.25%, -5.69%, -23.92%} and {-0.83%, -10.99%, -19.34%} BD-rate changes under AI and RA configurations for {Y, Cb, Cr} channels.

[JVET-AB0105](https://jvet-experts.org/doc_end_user/current_document.php?id=12032) Crosscheck of JVET-AB0076 (EE1-2.1: RPR-Based Super-Resolution Guided by Partition Information) [R. Chang (Tencent)] [late]

[JVET-AB0077](https://jvet-experts.org/doc_end_user/current_document.php?id=11992) EE1-2.2: CNN Filter for Super-Resolution with RPR functionality in VVC [S. Huang, C. Jung (Xidian Univ.), Y. Liu, M. Li (OPPO)]

The EE1-2.2 test reports BD-rate gains of JVET-AA0065 (AA0065 filter) over VTM-11.0\_NNVC-2.0 anchor with RPR functionality in VVC. Compared with VTM-11.0\_NNVC-2.0, AA0065 filter achieves {-8.51% (Y), 18.78% (U), -12.62% (V)} and {-3.89% (Y), 5.45% (U), -9.09% (V)} BD-rate gains (average on A1 and A2 Classes) in AI and RA configurations, respectively. Moreover, this contribution also reports the BD-rate gains of AA0065 filter over JVET-Z0065 on top of VTM-11.0-NNVC-2.0. JVET-Z0065 provided a method that the encoder can adaptively select a scale factor from ×1.0, ×1.5, and ×2.0 at the GOP level. Compared with JVET-Z0065, AA0065 filter achieves {-4.36% (Y), -1.73% (U), -21.26% (V)} and {-0.34% (Y), -9.94% (U), -18.46% (V)} BD-rate gains (average on A1 and A2 Classes) in AI and RA configurations, respectively.

[JVET-AB0106](https://jvet-experts.org/doc_end_user/current_document.php?id=12033) Crosscheck of JVET-AB0077 (EE1-2.2: CNN Filter for Super-Resolution with RPR functionality in VVC) [R. Chang (Tencent)] [late]

[JVET-AB0083](https://jvet-experts.org/doc_end_user/current_document.php?id=12010) EE1-1.8: More refinements on NN based in-loop filter with a single model [L. Wang, X. Xu, S. Liu (Tencent), Z. Xie, Y. Yu, H. Yu, D. Wang (OPPO)]

No need for presentation, sufficiently covered in EE summary report.

[JVET-AB0088](https://jvet-experts.org/doc_end_user/current_document.php?id=12015) Crosscheck of JVET-AB0083 (EE1-1.8: More refinements on NN based in-loop filter with a single model) [D. Liu (Ericsson)]

[JVET-AB0084](https://jvet-experts.org/doc_end_user/current_document.php?id=12011) EE1-2.3: A CNN-based Super Resolution Method with GOP Level Adaptive Resolution [R. Chang, L. Wang, X. Xu, S. Liu (Tencent), J. Nam, S. Yoo, J. Lim, S. Kim (LGE)]

No need for presentation, sufficiently covered in EE summary report.

[JVET-AB0097](https://jvet-experts.org/doc_end_user/current_document.php?id=12024) Crosscheck of JVET-AB0084 (EE1-2.3：A CNN-based Super Resolution Method with GOP Level Adaptive Resolution) [D. Liu (Ericsson)] [late]

[JVET-AB0164](https://jvet-experts.org/doc_end_user/current_document.php?id=12091) EE1-1.7: Capacity Ablation of CNN-based in-loop filtering [S. Eadie, H. Wang, M. Coban, M. Karczewicz (Qualcomm)]

In this contribution, the model complexity-performance trade-off of JVET-AA0131 is analysed independently for its luma and chroma models based on the channels’ complexities and subsequent required processing capacities. Specifically, the number of residual blocks is ablated across {16, 24, 32} with n=32 representing the previous model. For the libtorch floating-point, SADL floating-point and SADL int16 inference runtimes, the BD-rate savings for Y, Cb, Cr components in RA and AI configurations, respectively, are {Y%, U%, V%}, {Y%, U%, V%}, {Y%, U%, V%} and {Y%, U%, V%}, {Y%, U%, V%}, {Y%, U%, V%} for 32 residual blocks, {Y%, U%, V%}, {Y%, U%, V%}, {Y%, U%, V%} and {Y%, U%, V%}, {Y%, U%, V%}, {Y%, U%, V%} for 24 residual blocks, {Y%, U%, V%}, {Y%, U%, V%}, {Y%, U%, V%} and {Y%, U%, V%}, {Y%, U%, V%}, {Y%, U%, V%} for 24 and 16 residual blocks for the luma and chroma models, respectively, and {Y%, U%, V%}, {Y%, U%, V%}, {Y%, U%, V%} and {Y%, U%, V%}, {Y%, U%, V%}, {Y%, U%, V%} for 16 residual blocks.

In total, four models are used (different for luma/chroma, inter/intra)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Network Information in Inference Stage** | | | | | | | | | |
| Mandatory | Test | | n=16 | n=(24,16) | | n=24 | | n=32 | |
| HW environment: | | | | | | | | |
| GPU Type | | N/A | | | | | | |
| Framework: | | Libtorch, SADL | | | | | | |
| Number of GPUs per Task | | 0 | | | | | | |
| Total Parameter Number | | 1.0 M/model | (1.0,1.45)M | | 1.45M/model | | 1.9M/model | |
| Parameter Precision (Bits) | | 32 or 16 | | | | | | |
| Memory Parameter (MB) | |  |  | |  | |  | |
| Multiply Accumulate (kMAC/pixel) picture (block) based | | 337.05 (426.58) | 450.38 (570.02) | | 478.38 (605.45) | | 619.70 (784.31) | |
| Total Conv. Layers | | 57 /model | (81,57) | | 81 /model | | 105 / model | |
| Total FC Layers | | 0 | | | | | | |
| Total Memory (MB) | Float | 4.0 /model | | (5.8,4.0) | | 5.8 /model | | 7.5 /model |
| Int16 | 2.0 /model | | (2.9,2.0) | | 2.9 /model | | 7.5 /model |
| Batch size: | | 1 | | | | | | |
| Patch size | | 128128, 256256 | | | | | | |
| Changes to network configuration or weights required to generate rate points | |  | | | | | | |
| Peak Memory Usage | |  |  | |  | |  | |
| Other information: | |  |  | |  | |  | |

|  |  |  |
| --- | --- | --- |
| **Network Information in Training Stage** | | |
| Mandatory | GPU Type | GPU: NVIDIA Tesla V100-SXM2-32GB |
| Framework: | PyTorch v1.8 |
| Number of GPUs per Task | 1 |
|  |  |
| Epoch: | 120 |
| Batch size: | 64 |
| Training time: | ~120h |
| Training data information: | DIV2K, BVI-DVC |
| Training configurations for generating compressed training data (if different to VTM CTC): | QP {22, 27, 32, 37, 42} |
|  | Loss function: | L1, L2 |
| Optional | Number of iterations |  |
| Patch size | 144x144 |
| Learning rate: | 1e-4 |
| Optimizer: | ADAM |
| Preprocessing: |  |
| Other information: |  |

The BD-rate savings using the libtorch, floating-point inference runtime for the Y, Cb and Cr components, respectively, in the RA and AI configurations, for 16, (24 for luma,16 for chroma), 24 and 32 residual blocks are shown in Tables 1-4, respectively.

*Table 1: n=16*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | **All Intra Main10** | | |
|  | **BD-rate Over VTM-11.0\_nnvc-1.0** | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-PSNR | U-PSNR | V-PSNR |
| Class A1 | -8.44% | -15.88% | -18.16% | -6.61% | -17.26% | -18.27% |
| Class A2 | -9.55% | -19.23% | -14.29% | -6.61% | -17.94% | -13.85% |
| Class B | -8.71% | -20.47% | -21.46% | -6.84% | -17.99% | -20.07% |
| Class C | -10.00% | -21.49% | -21.49% | -7.61% | -16.69% | -19.31% |
| Class E |  |  |  | -9.89% | -18.52% | -20.50% |
| **Overall** | -9.17% | -19.58% | -19.37% | -7.44% | -17.66% | -18.63% |
| Class D | -11.82% | -23.07% | -24.30% | -7.53% | -15.89% | -19.18% |
| Class F | #VALUE! | #VALUE! | #VALUE! | -5.69% | -16.39% | -14.83% |

*Table 2: n=(24, 16)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | **All Intra Main10** | | |
|  | **BD-rate Over VTM-11.0\_nnvc-1.0** | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-PSNR | U-PSNR | V-PSNR |
| Class A1 | -9.17% | -15.65% | -17.77% | -7.01% | -17.06% | -18.21% |
| Class A2 | -10.27% | -19.26% | -14.30% | -7.05% | -17.99% | -13.84% |
| Class B | -9.39% | -20.47% | -21.52% | -7.28% | -16.33% | -18.95% |
| Class C | -10.64% | -21.40% | -21.54% | -8.09% | -16.50% | -19.12% |
| Class E |  |  |  | -10.54% | -18.51% | -20.55% |
| **Overall** | -9.85% | -19.51% | -19.33% | -7.92% | -17.13% | -18.28% |
| Class D | -12.24% | -22.94% | -24.34% | -7.92% | -15.67% | -19.02% |
| Class F | -5.86% | -15.22% | -15.12% | -6.19% | -16.45% | -14.90% |

*Table 3: n=24*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | **All Intra Main10** | | |
|  | **BD-rate Over VTM-11.0\_nnvc-1.0** | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-PSNR | U-PSNR | V-PSNR |
| Class A1 | #VALUE! | #VALUE! | #VALUE! | -7.01% | -17.00% | -18.37% |
| Class A2 | #VALUE! | #VALUE! | #VALUE! | -7.05% | -18.24% | -14.14% |
| Class B | -9.50% | -20.89% | -21.65% | -7.28% | -16.37% | -19.11% |
| Class C | -10.71% | -21.89% | -21.84% | -8.09% | -16.50% | -19.39% |
| Class E |  |  |  | -10.54% | -18.57% | -20.47% |
| **Overall** | #VALUE! | #VALUE! | #VALUE! | -7.92% | -17.18% | -18.45% |
| Class D | -12.25% | -23.54% | -24.90% | -7.92% | -15.83% | -19.32% |
| Class F | -5.90% | -15.42% | -15.08% | -6.17% | -16.78% | -15.15% |

*Table 4: n=32*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | **All Intra Main10** | | |
|  | **BD-rate Over VTM-11.0\_nnvc-1.0** | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-PSNR | U-PSNR | V-PSNR |
| Class A1 | -9.87% | -17.14% | -19.77% | -7.24% | -17.14% | -17.90% |
| Class A2 | -10.99% | -21.85% | -26.20% | -7.15% | -18.84% | -16.15% |
| Class B | -10.04% | -24.92% | -24.46% | -7.55% | -18.44% | -18.60% |
| Class C | -11.14% | -23.89% | -24.78% | -8.40% | -18.22% | -20.55% |
| Class E |  |  |  | -10.86% | -19.99% | -19.44% |
| **Overall** | -10.49% | -22.48% | -23.96% | -8.17% | -18.50% | -18.65% |
| Class D | -12.69% | -25.39% | -27.46% | -8.18% | -17.95% | -20.32% |
| Class F | -5.50% | -15.18% | -14.58% | -6.50% | -11.89% | -11.51% |

The BD-rate savings using the SADL, int16 inference runtime for the Y, Cb and Cr components, respectively, in the RA and AI configurations, for 16, (24 for luma,16 for chroma), 24 and 32 residual blocks are shown in Tables 9-12, respectively.

*Table 9: n=16*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | **All Intra Main10** | | |
|  | **BD-rate Over VTM-11.0\_nnvc-1.0** | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-PSNR | U-PSNR | V-PSNR |
| Class A1 | #VALUE! | #VALUE! | #VALUE! | -6.55% | -16.68% | -18.28% |
| Class A2 | #VALUE! | #VALUE! | #VALUE! | -6.61% | -17.77% | -13.81% |
| Class B | -9.14% | -20.98% | -22.25% | -6.75% | -17.63% | -19.91% |
| Class C | -10.28% | -22.20% | -22.21% | -7.53% | -16.63% | -19.22% |
| Class E |  |  |  | -9.80% | -18.17% | -20.34% |
| **Overall** | #VALUE! | #VALUE! | #VALUE! | -7.37% | -17.36% | -18.54% |
| Class D | -11.99% | -23.50% | -25.07% | -7.45% | -15.80% | -19.21% |
| Class F | #VALUE! | #VALUE! | #VALUE! | -5.14% | -15.92% | -14.48% |

*Table 10: n=(24, 16)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | **All Intra Main10** | | |
|  | **BD-rate Over VTM-11.0\_nnvc-1.0** | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-PSNR | U-PSNR | V-PSNR |
| Class A1 | #VALUE! | #VALUE! | #VALUE! | -6.93% | -16.55% | -18.24% |
| Class A2 | #VALUE! | #VALUE! | #VALUE! | -6.99% | -17.70% | -13.61% |
| Class B | #VALUE! | #VALUE! | #VALUE! | -7.20% | -17.19% | -19.77% |
| Class C | #VALUE! | #VALUE! | #VALUE! | -8.09% | -16.51% | -19.13% |
| Class E |  |  |  | -10.49% | -18.19% | -20.34% |
| **Overall** | #VALUE! | #VALUE! | #VALUE! | -7.87% | -17.19% | -18.44% |
| Class D | -12.47% | -23.55% | -25.19% | -7.92% | -15.54% | -19.01% |
| Class F | #VALUE! | #VALUE! | #VALUE! | -6.23% | -16.41% | -14.98% |

*Table 11: n=24*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | **All Intra Main10** | | |
|  | **BD-rate Over VTM-11.0\_nnvc-1.0** | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-PSNR | U-PSNR | V-PSNR |
| Class A1 | #VALUE! | #VALUE! | #VALUE! | -6.92% | -16.36% | -18.22% |
| Class A2 | #VALUE! | #VALUE! | #VALUE! | -7.00% | -17.82% | -13.96% |
| Class B | #VALUE! | #VALUE! | #VALUE! | -7.20% | -17.35% | -20.08% |
| Class C | -10.92% | -22.69% | -22.85% | -8.08% | -16.55% | -19.50% |
| Class E |  |  |  | -10.49% | -18.23% | -19.79% |
| **Overall** | #VALUE! | #VALUE! | #VALUE! | -7.86% | -17.23% | -18.57% |
| Class D | -12.44% | -24.11% | -25.75% | -7.92% | -15.80% | -19.38% |
| Class F | -6.03% | -15.56% | -15.11% | -6.22% | -16.76% | -15.34% |

*Table 12: n=32*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | **All Intra Main10** | | |
|  | **BD-rate Over VTM-11.0\_nnvc-1.0** | | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | Y-PSNR | U-PSNR | V-PSNR |
| Class A1 | #VALUE! | #VALUE! | #VALUE! | -7.27% | -15.88% | -18.50% |
| Class A2 | #VALUE! | #VALUE! | #VALUE! | -7.23% | -17.29% | -13.98% |
| Class B | #VALUE! | #VALUE! | #VALUE! | -7.48% | -17.24% | -20.07% |
| Class C | #VALUE! | #VALUE! | #VALUE! | -8.39% | -15.96% | -18.91% |
| Class E |  |  |  | -10.81% | -18.13% | -20.60% |
| **Overall** | #VALUE! | #VALUE! | #VALUE! | -8.16% | -16.89% | -18.62% |
| Class D | -12.38% | -24.76% | -27.04% | -8.17% | -15.57% | -18.99% |
| Class F | #VALUE! | #VALUE! | #VALUE! | -6.59% | -16.36% | -15.28% |

Luma and chroma models are trained independently, loss function is combination of L1/L2. No balance of luma and chroma quality was considered in training.

Block-level on/off is used. Residual scaling is used with adjustment at picture level.

[JVET-AB0239](https://jvet-experts.org/doc_end_user/current_document.php?id=12168) Crosscheck of JVET-AB0164 (EE1-1.7: Capacity Ablation of CNN-based in-loop filtering) [Y. Li (Bytedance)] [late] [miss]

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

Contributions in this area were discussed at 1200–1310 on Saturday 22 October 2022 (chaired by JRO), and at 1400–1515 on Sunday 23 October 2022 (chaired by JRO).

[JVET-AB0090](https://jvet-experts.org/doc_end_user/current_document.php?id=12017) EE1-1.3 related: Lightweight and Efficient CNN In-loop Filter [H. Zhang, C. Jung (Xidian Univ.), Y. Liu, M. Li (OPPO)]

This contribution presents a convolutional neural network (CNN)-based in-loop filter. In this contribution, the CNN-based in-loop filter proposed in JVET-AA0074 (AA0074 filter) is made lightweight and efficient. Moreover, a multi-stage training strategy under AI configuration is proposed to achieve an optimal trade-off between performance and complexity. The training strategy of AA0074 filter, i.e. use ground truth as the label, is used to verify the performance of the proposed CNN filter in RA configuration. Compared with EE1-anchor-2.0 (VTM-11.0\_NNVC-2.0), the proposed CNN filter achieves average {7.08%, 12.46%, 12.75%} BD-rate reductions for {Y, Cb, Cr} under AI configuration and average {5.43%, 15.34%, 14.79%} BD-rate reductions for {Y, Cb, Cr} under RA configuration.

The network architecture of the proposed CNN filter is shown in Fig. 1. Compared with AA0074 filter, the following improvements are mainly made. First, we reduce the convolution layer of the feature extraction part and the reconstruction part, and reduce the number of channels in the backbone. Then, some minor changes are made to WCDAB and RAB. Specifically, we reduce the number of skip connections in WCDAB and the number of convolution layers in RAB. The modified WCDAB and RAB are shown in Fig. 2(a) and Fig. 2(b), respectively.

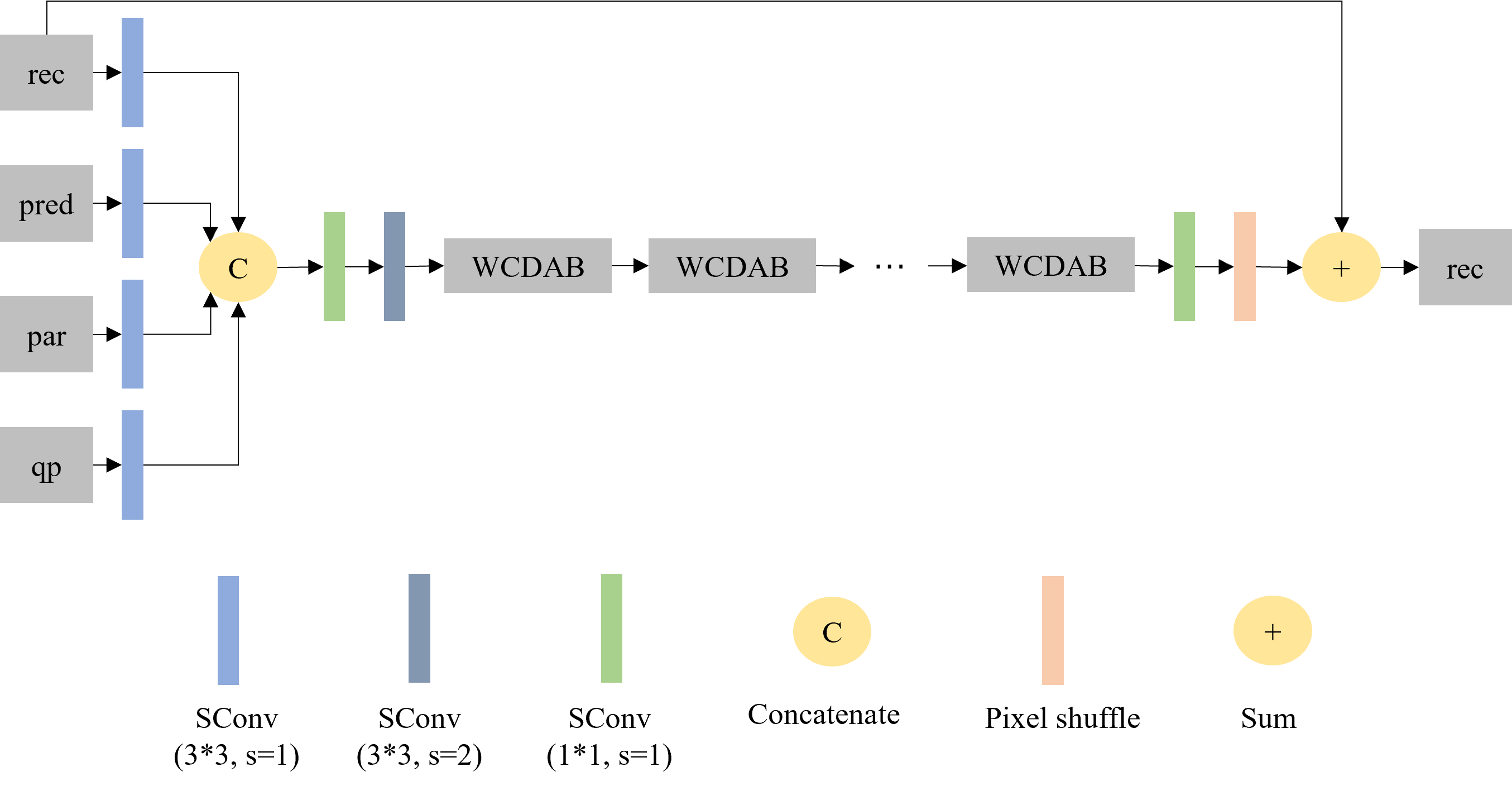


Fig. 1 Illustration of the proposed NN-based in-loop filter. (SConv is the standard convolution)

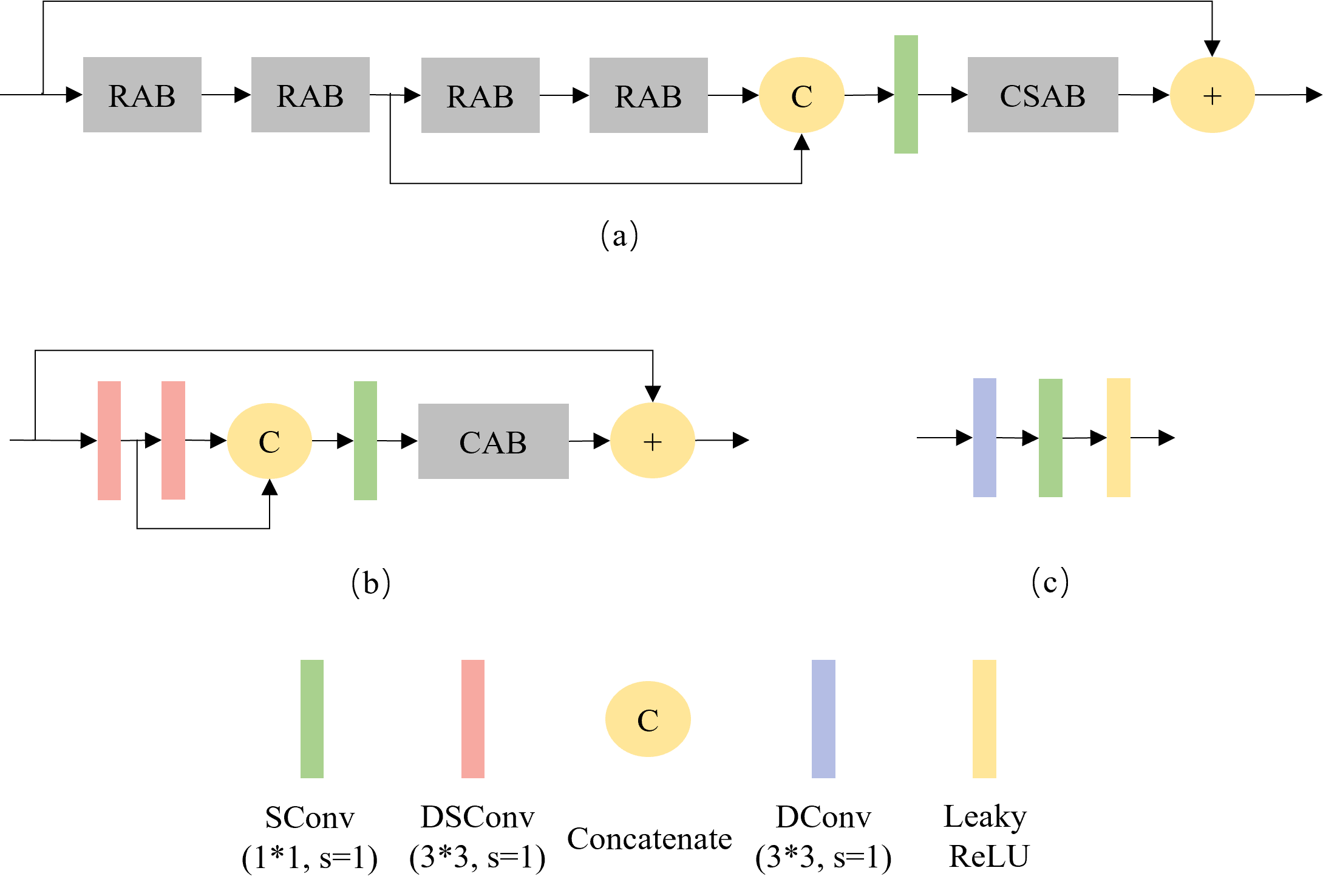


Fig.2 (a) The architecture of WCDAB. (b) The architecture of RAB. (c) The architecture of Depthwise separable convolution. (SConv is standard convolution, DSConv is depthwise separable convolution, and DConv is depthwise convolution)

In training. first, qp\_dis is set to 5 and trained the network with L1 and L2 loss functions successively. Then, qp\_dis is increased to 10, and then trained the proposed network again with L1 and L2 loss functions. After that, continue to increase qp\_dis and use the same method to continue training.

|  |  |  |
| --- | --- | --- |
| **Network Information in Training Stage** | | |
| Mandatory | GPU Type | GPU: GeForce RTX 3090 |
| Framework: | PyTorch v1.9.0 |
| Number of GPUs per Task | 1 |
|  |  |
| Epoch: | 800 |
| Batch size: | 32 |
| Training time: | ~200h/model |
| Training data information: | BVI-DVC, DIV2K |
| Training configurations for generating compressed training data (if different to VTM CTC): | VTM-11.0, QP {12,17,22, 27, 32, 37, 42} |
|  | Loss function: | Weighted L1 and L2 |
| Optional |  |  |
| Number of iterations | 3100 |
| Patch size | 144x144 |
| Learning rate: | 1e-4 |
| Optimizer: | ADAM |
| Preprocessing: | random cropping |
| Other information: |  |
|  |  |

|  |  |  |
| --- | --- | --- |
| **Network Information in Inference Stage** | | |
| Mandatory | HW environment: | |
| GPU Type | CPU only |
| Framework: | Libtorch v1.9.0 |
| Number of GPUs per Task | 0 |
|  |  |
| Number of Parameters (Each Model) | 0.78M |
| Total Number of Parameters (All Models) | 0.78M |
| Parameter Precision (Bits) | 32 |
| Memory Parameter (MB) | 3.12M |
| Multiply Accumulate (MAC)/pixel | 200K |
| Optional |  |  |
| Total Conv. Layers | 64 Depthwise separable convolution  56 Common convolution |
| Total FC Layers | 72 |
| Total Memory (MB) |  |
| Batch size: | 1 |
| Patch size | 144x144 |
| Changes to network configuration or weights required to generate rate points |  |
| Peak Memory Usage (Total) |  |
| Peak Memory Usage (per Model) |  |
| Border handling |  |
| Other information: |  |
|  |  |

What is the benefit (in terms of gain) of the multistage training? Approximately 0.5%

Why is a switch between L1 and L2 loss performed in different stages of training? Was found experimentally to give advantage.

No specific action.

[JVET-AB0093](https://jvet-experts.org/doc_end_user/current_document.php?id=12020) EE1-2.2 related: Lightweight CNN Filter for Super-Resolution with RPR functionality in VVC [S. Huang, C. Jung (Xidian Univ.), Y. Liu, M. Li (OPPO)]

This contribution presents a lightweight super-resolution filter that combines convolutional neural networks (CNNs) with existing RPR functionality in VVC. In this contribution, we modify the basic unit of MMSDAB proposed by JVET-AA0065 and add spatial attention into MMSDAB to make MMSDANet lightweight, called LMSDANet. Compared with VTM-11.0\_NNVC-2.0, the proposed CNN filter achieves average {-9.16%, 17.03%, -7.61%} and {-4.14%, 6.34%, -2.25%} BD-rate gains (average on A1 and A2) in AI and RA configurations, respectively.

In LMSDAB, the branch used to extract multi-scale information is simplified from MMSDAB. As shown in Fig. 3, LMSDAB is as the basic unit of LMSDANet. LMSDAB aims to extract multi-scale and depth features from a large receptive field and stacking convolutional layers and emphasize important spatial and channel information by MSSAB and channel attention block (CAB) from the extracted features. LMSDAB consists of three parts: the first is feature extraction part, the second is feature fusion part, and the third is attention enhancement part. The feature extraction part contains one 1x1 convolution layer and three 3x3 convolution layers. The feature fusion part concatenates features in the channel dimension and uses a 1x1 convolution layer for fusion and dimension reduction. The attention enhancement part uses MSSAB and CAB to enhance the fused features in both spatial and channel dimensions. The structures of LMSDANet and MMSDANet (AA0065 filter) are consistent. This contribution only modifies MMSDAB and presents LMSDAB.

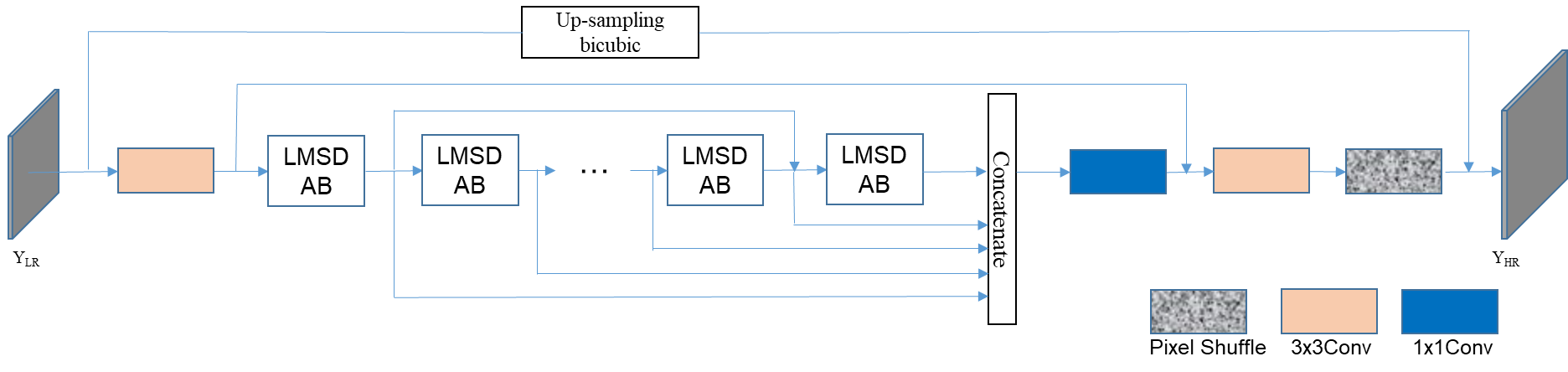


Fig. 1 Network architecture of the proposed LMSDANet for Y channel.

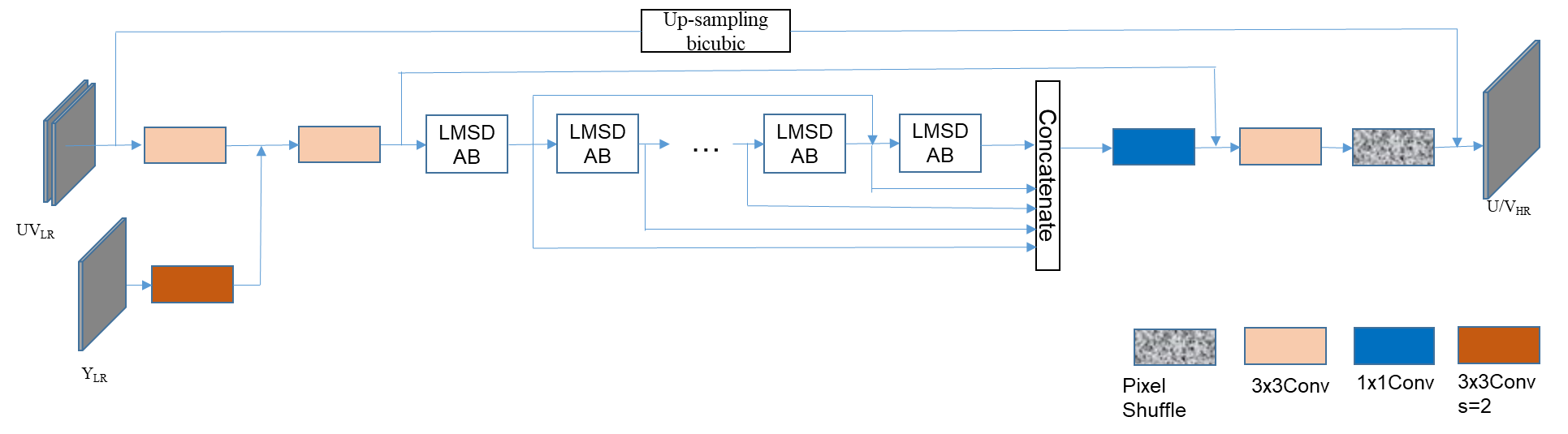


Fig. 2 Network architecture of the proposed LMSDANet for UV channel.

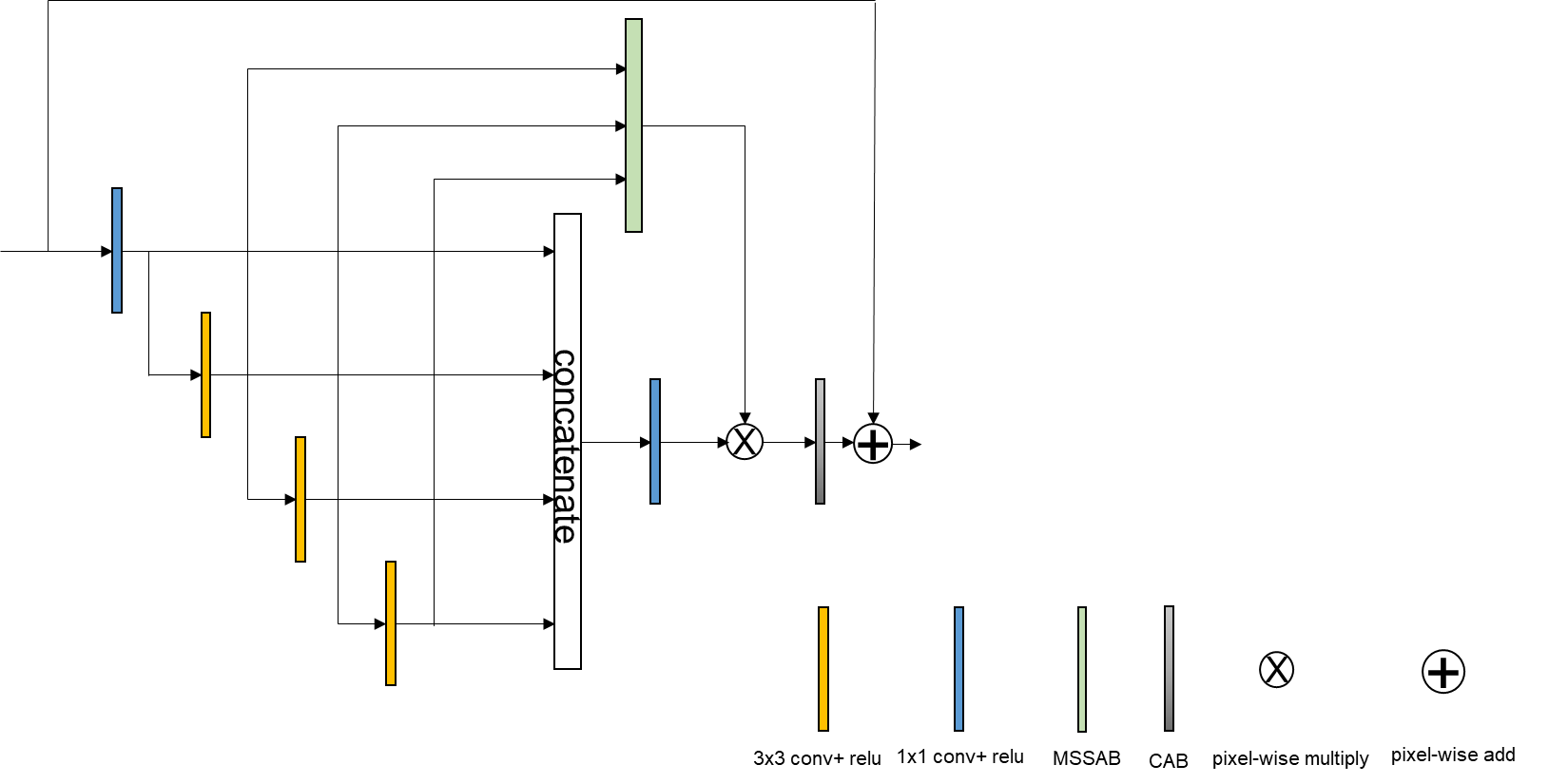


Fig. 3 Network architecture of the modified MMSDAB (LMSDAB).

As shown in Figure 4, the multi-scale spatial attention is mainly composed of three spatial attention blocks, a concatenation layer and a convolution layer. The input of three SABs is respectively from the output of three branches of the feature extraction part in LMSDAB, and each proposed feature is subject to a spatial attention operation. Then, three spatial attention maps are concatenated together and go through a 3x3 convolution layer for fusion to obtain the final multi-scale spatial attention map.

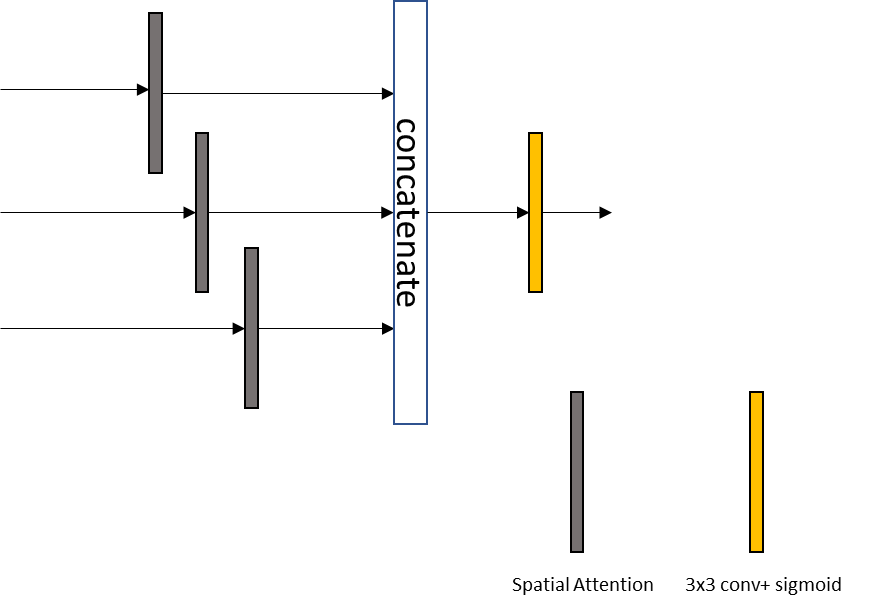


Fig. 4 Network architecture of the MSSAB.

|  |  |  |
| --- | --- | --- |
| **Network Information in Inference Stage** | | |
| Mandatory | HW environment: | Intel Core i9 12900k @3.9GHz |
| Framework: | LibTorch v1.8 |
| Number of GPUs per Task | 0 |
| Number of Parameters (Each Model) | luma up-sampling model: 1.13M/model  chroma up-sampling model: 1.18M/model |
| Total Parameter Number | 17.5M |
| Parameter Precision (Bits) | 32 (F) |
| Memory Parameter (MB) | 15 models in total: 67 MB |
| Multiply Accumulate (MAC) | 964kMAC/pixel |
| Optional | Total Conv. Layers | 91 for up-sampling the luma, 92 for up-sampling the chroma |
| Total FC Layers | 0 |
| Batch size: | 1 |
| Patch size | Whole frame |

|  |  |  |
| --- | --- | --- |
| **Network Information in Training Stage** | | |
| Mandatory | GPU Type | GPU: NVIDIA 3090 24GB |
| Framework: | PyTorch v1.9 |
| Number of GPUs per Task | 1 |
| Epoch: | luma:50, chroma:50 |
| Batch size: | 64 |
| Training time: | 26h/model |
| Training data information: | DIV2K, BVI\_DVC |
| Training configurations for generating compressed training data (if different to VTM CTC): | VTM-11.0\_NNVC-2.0 + new MCTF, QP {22, 27, 32, 37, 42} |
|  | Loss function: | L2 |
| Optional | Number of iterations |  |
| Patch size | 128128 |
| Learning rate: | 1e-4 |
| Optimizer: | ADAM |
| Preprocessing: |  |
| Other information: |  |

Investigate in EE (replacing the previous EE1-2.2 which was a much more complex version)

[JVET-AB0102](https://jvet-experts.org/doc_end_user/current_document.php?id=12029) AHG11/EE1-related: Updates on RPR encoder and filters [J. Nam, S. Yoo, J. Lim, S. Kim (LGE)]

This contribution proposes an RPR encoder with four available scale factors {x2.0, x1.5, x1.25, x1.0}. The scale factor is selected at GOP level based on the initial QP and PSNR value of the first picture without multi-pass coding, and it is applied to all pictures within the GOP. This contribution also suggests to increase RPR filters for NNVC software to align with ECM where the luma and chroma interpolation filters are using 12-tap (non-affine luma), 10-tap (affine luma), and 6-tap (chroma), respectively. Lastly, it is proposed to increase the post-filter taps used for re-sampling the current picture to the original picture size. The experimental results are shown as follows;

* Test 1 (Adding x1.25 scale factor)
  + AI: -0.55% (Y), 5.02% (U), 3.83% (V), 99% (EncT), 90% (DecT)
  + RA: -1.07% (Y), 3.76% (U), 3.23% (V), 92% (EncT), 79% (DecT)
* Test 2 (Test1 + increasing RPR filters)
  + AI: -0.55% (Y), 5.00% (U), 3.86% (V), 100% (EncT), 91% (DecT)
  + RA: -1.11% (Y), 3.71% (U), 3.18% (V), 92% (EncT), 79% (DecT)
* Test 3 (Test2 + increasing post-filter)
  + AI: -0.57% (Y), 4.60% (U), 3.55% (V), 99% (EncT), 91% (DecT)
  + RA: -1.12% (Y), 3.36% (U), 2.81% (V), 91% (EncT), 79% (DecT)
* Test 4 (Test1 + increasing RPR filters + increasing post-filter)
  + AI: -0.57% (Y), 4.60% (U), 3.55% (V), 99% (EncT), 91% (DecT)
  + RA: -1.12% (Y), 3.34% (U), 2.80% (V), 92% (EncT), 79% (DecT)

Investigate in EE1-2 as another more optimized approach of lower-resolution coding with “conventional” upsampling, still retain the previous results from JVET-Z0065 as comparison point.

It would be desirable to align the encoder decision in coordination with the proponents of JVET-AB0080, which is a very similar approach implemented on top of VTM18 (whereas JVET-AB0102 was implemented in the NNVC2.0 used in EE1).

[JVET-AB0242](https://jvet-experts.org/doc_end_user/current_document.php?id=12171) Cross-check of JVET-AB0102 (AHG11/EE1-related: Updates on RPR encoder and filters) [K. Andersson (Ericsson)] [late]

[JVET-AB0141](https://jvet-experts.org/doc_end_user/current_document.php?id=12068) EE1-related: QP-based loss function design for NN-based in-loop filter [C. Zhou, Z. Lv, J. Zhang (vivo), W. Chen, J. Guo, B. Ai (BJTU)]

In this contribution, a new loss function is defined where quantization parameter is used as a factor for the loss calculation in the training stage. In this study, the intra models from NCS-1.0 with NN filter set 1 are retrained using the latest published scripts. BD-rate changes of {Y, Cb, Cr} are summarized as follows:

* Compared with NNVC2.0

AI: { -7.22 %, -20.64 %, -21.34 % }

* Compared with NCS-1.0 with the retrained intra model

AI: { -0.12%, -0.69%, -1.05 % }

In the latter case, the model as well as the anchor were retrained from scratch. Would be more interesting to compare against the actual NCS-1.0 filter model. It was verbally reported that comparing against NCS-1.0 would give a small loss.

It was suggested to update the contribution, correcting the weight factor which was found to be incorrect, and adding the duration of training.

Study in EE

[JVET-AB0146](https://jvet-experts.org/doc_end_user/current_document.php?id=12073) EE1-1.8-related: encoder-only optimization for NN based in-loop filter with a single model [L. Wang, X. Xu, S. Liu (Tencent)]

In this contribution, an encoder-only tool for EE1-1.8, which is derived based on EE1-1.6, is studied on top of NCS-1.0 (filter set #0). The specific encoder implementation is same to one of EE1-1.6, but the filters in the RDO process are retained based on the network structure of in-loop filter in EE1-1.8. Then, the well-trained models are used to further enhance the solution in EE1-1.8. Based on the NNVC-2.0, the test results are shown in order of RA and AI configurations as follows.

RA: -9.43% -19.68% -19.12% EncT: 158% DecT: 65467%

AI : -7.06% -15.37% -16.46% EncT: 172% DecT: 34972%

Compared with EE1-1.8 under RA and AI configurations, additional 0.25% (9.18% to 9.43%) and 0.43% (6.63% to 7.06) BD-rate luma gains are observed. In terms of decoder side, there is not any increase in both MAC result and memory size. As for the runtime at encoder side, there are about 8% (150% to 158%) and 26% (146% to 172%) increase additionally under RA and AI configurations, respectively.

The decoder model of filter #0 is the same as in EE1-1.8.2.

Investigate in EE. It is agreed that no extensive training cross-check is necessary for including encoder-side optimizations in the software, but training scripts shall be made available in order to avoid that somebody overtrains an encoder for CTC.

It was further agreed that extensive usage of trained encoder optimizations is undesirable in the EE experimentation. Methods of encoder optimization should be included in the software as optional part, but not enabled in CTC. Otherwise it is hard to judge whether a benefit reported for a new element is due to encoder optimization or due to its own benefit. Furthermore, if usage of an additional encoder optimization training would become common, it is an additional burden in the development of technology.

[JVET-AB0263](https://jvet-experts.org/doc_end_user/current_document.php?id=12192) Crosscheck of JVET-AB0146 (EE1-1.8-related: encoder-only optimization for NN based in-loop filter with a single model) [Z. Xie (OPPO)] [late]

[JVET-AB0147](https://jvet-experts.org/doc_end_user/current_document.php?id=12074) EE1-1.8-related: using additional models for higher temporal layers [L. Wang, X. Xu, S. Liu (Tencent)]

In this contribution, a NN based in-loop filter with reference images as one of inputs, which is derived based on EE1-1.4, is studied on top of NCS-1.0 (filter set #0) for higher temporal layers. The luma inter mode with reference images and the chroma inter model from EE1-1.4 are directly used to replace the in-loop filter of JVET-AB0146, which is implemented on top of EE1-1.8, for the higher temporal layers. Based on NNVC-2.0, the test results are shown in order of RA and AI configurations as follows.

RA: -9.91% -19.72% -19.15% EncT: 152% DecT: 51248%

AI : -7.06% -15.37% -16.46% EncT: 173% DecT: 34633%

Compared with the base method in JVET-AB0146 under RA configuration, an additional 0.48% (9.43% to 9.91%) BD-rate luma gain is observed.

Compared with EE1-1.8 under RA configuration, an additional 0.73% (9.18% to 9.91%) BD-rate luma gain is observed.

|  |  |  |
| --- | --- | --- |
| **Network Information in Training Stage** | | |
| Mandatory | GPU Type | Tesla V100 32GB |
| Framework: | Pytorch v1.9.0 |
| Number of GPUs per Task | 1 |
|  |  |
| Epoch: | ~100 |
| Batch size: | 64 |
| Loss function: | L1 |
| Training time (for 1 model): | ~60h |
| Training data information: | DIV2K, TVD, BVI-DVC |
| Training configurations for generating compressed training data (if different to VTM CTC): |  |
| Optional |  |  |
| Number of iterations |  |
| Patch size |  |
| Learning rate: | 1e-4 |
| Learning rate update strategy |  |
| Optimizer: | ADAM |
| Preprocessing: |  |
| Mini-batch selection process: |  |
| Training data update strategy: |  |
| Other information: |  |
|  |  |

|  |  |  |
| --- | --- | --- |
| **Network Information in Inference Stage** | | |
| Mandatory | HW environment: | |
| GPU Type | CPU only |
| Framework: | SADL |
| Number of GPUs per Task | 0 |
|  |  |
| Number of Parameters (Each Model) | 1.9M x 1, 1.56M x 2 |
| Total Number of Parameters (All Models) | 5.02M |
| Parameter Precision (Bits) | 16 |
| Memory Parameter (MB) | 10.04MB in all |
| Multiply Accumulate (kMAC/pixel) | 537K (w/o block extension)  680K (w/ block extension) |
| Calculation Method | block basis |
| Optional |  |  |
| Total Conv. Layers |  |
| Total FC Layers |  |
| Total Memory (MB) |  |
| Batch size: |  |
| Patch size |  |
| Changes to network configuration or weights required to generate rate points |  |
| Peak Memory Usage (Total) |  |
| Peak Memory Usage (per Model) |  |
| Border handling |  |
| Other information: |  |
|  |  |

Two of the base models are from JVET-AB0073.

Proponents suggest further study outside of EE.

[JVET-AB0179](https://jvet-experts.org/doc_end_user/current_document.php?id=12106) EE1-related: Deep In-Loop Filter with Wide Activation and Large Receptive Field [Y. Li, K. Zhang, L. Zhang (Bytedance)]

This contribution proposes a deep in-loop filter which is constructed based on basic residual blocks with wide activation and large receptive field. The proposed filter is implemented on top of NNVC common software NCS-1.0. Two tests are conducted with regular and compact models, where both models have lower kMAC/pixel than filter set #1 in NCS-1.0. BD-rate changes of {Y, Cb, Cr} compared with NCS-1.0 (filter set #1) and NNVC-2.0 are reportedly summarized as below:

Compared with NCS-1.0 (filter set #1):

Test #1 (Regular model): RA: {-2.26%, -5.24%, -6.56%}, LB: {-2.90%, -6.57%, -9.03%}, AI: {-1.55%, -1.94%, -2.12%}

Test #2 (Compact model): RA: {-1.98%, -5.24%, -6.67%}, LB: {-2.48%, -6.64%, -8.86%}, AI: {-1.37%, -2.09%, -2.34%}

Compared with NNVC-2.0:

Test #1 (Regular model): RA: {-11.48%, -24.60%, -25.41%}, LB: {-11.11%, -21.61%, -22.79%}, AI: {-8.68%, -21.49%, -22.09%}

Test #2 (Compact model): RA: {-11.23%, -24.61%, -25.50%}, LB: {-10.72%, -21.67%, -22.66%}, AI: {-8.51%, -21.61%, -22.26%}

The architecture of the proposed CNN filter is shown below, which comprises three types of basic blocks known as HeadBlock, BackboneBlock, and TailBlock. The design of these blocks follows the principle of wide activation, large receptive field, and multi-scale feature extraction.

HeadBlock is responsible for extracting features from input. denotes the number of input channels, and is equal to 5 for intra model (rec, pred, split, bs, qp) and 3 for inter model (rec, pred, qp). stands for the basic number of feature maps and is set as 64. {, } represent numbers of output channels in large activation branch and large receptive field branch, and are set as {160, 32}. means the stride of convolution and is set as 2 to achieve feature down-sampling. Backbone of the proposed network containing a series of BackboneBlocks achieves feature embedding. , the number of BackboneBlocks, is set as 22 and 19 for the regular model and compact model. In the end, there is a TailBlock mapping the embedded features from backbone to the final output.



*Architecture of the proposed deep in-loop filter.*

|  |  |  |
| --- | --- | --- |
| **Network Information in Inference Stage** | | |
| Mandatory | HW environment: | |
| GPU Type | N/A |
| Framework: | SADL |
| Number of GPUs per Task | 0 |
|  |  |
| Total Parameter Number | regular:  1.85M/luma model, 0.93M/chroma model, 5.56M in total  compact:  1.62M/luma model, 0.93M/chroma model, 5.1M in total |
| Parameter Precision (Bits) | 16 (I) |
| Memory Parameter (MB) | regular:  3.7MB/luma model, 1.9MB/chroma model, 11.2M in total  compact:  3.3MB/luma model, 1.9MB/chroma model, 10.4M in total |
| Multiply Accumulate (kMAC/pixel) | regular:  537K/pixel (frame-level input)  680K/pixel (block-level input)  compact:  479K/pixel (frame-level input)  606K/pixel (block-level input) |
| Optional |  |  |
| Total Conv. Layers | regular: 97  compact: 85 |
| Total FC Layers | 0 |
| Total Memory (MB) |  |
| Batch size: | 1 |
| Patch size | 144144, 272272 |
| Changes to network configuration or weights required to generate rate points |  |
| Peak Memory Usage |  |
| Other information: |  |

|  |  |  |
| --- | --- | --- |
| **Network Information in Training Stage** | | |
| Mandatory | GPU Type | GPU: A100-SXM-80GB |
| Framework: | PyTorch v1.6 |
| Number of GPUs per Task | 4 |
|  |  |
| Epoch: | 60 |
| Batch size: | 64 |
| Training time: | 70h/model |
| Training data information: | DIV2K, BVI-DVC |
| Training configurations for generating compressed training data (if different to VTM CTC): | VTM-11.0 + new MCTF, QP {17, 22, 27, 32, 37, 42} |
|  | Loss function: | L1, L2 |
| Optional |  |  |
| Number of iterations |  |
| Patch size | 128128 |
| Learning rate: | 1e-4 |
| Optimizer: | ADAM |
| Preprocessing: |  |
| Other information: |  |

4 Models (Intra/Inter, Luma/Chroma)

It was pointed out that the parallel branches of the residual blocks have an asymmetry, which might have some impact for the GPU memory allocation. Aspects like may currently not be considered in the computations for memory requirements.

Proponents suggest further study outside of EE.

### Improvements of NNVC technology and/or base software beyond EE1 (13)

Contributions in this area were discussed at 1515–1900 on Sunday 23 October 2022 (chaired by JRO until 1800, documents after JVET-AB0136 chaired by A. Segall).

[JVET-AB0098](https://jvet-experts.org/doc_end_user/current_document.php?id=12025) EE1-2.3 related: GOP Level Adaptive Resampling with CNN-based Super Resolution [R. Chang, L. Wang, X. Xu, S. Liu (Tencent)]

The CNN-based super resolution methods for video coding show considerable performance, but there is still room to further improve the coding gains especially for chroma components. In order to obtain higher compression performance, this contribution presents a GOP level adaptive resampling method with CNN-based super resolution. At the GOP level, the proposed method can adaptively select a scale factor from ×1.0 (original size) and ×2.0 (half size) to determine the encoding resolution, and the designed CNN-based super-resolution will be used for the half size. Compared with VTM-11.0-NNVC-v2, the experimental results show {-4.42%, -4.83%, -4.11%} and {-7.53%, -8.14%, -7.24%} BD-rate savings for {Y, Cb, Cr} under RA and AI configurations, respectively.

Based on the method JVET-AA0084, model not changed. Compared to EE1-2.3, the bit rate saving is more than doubled. There is probably gain due to the better encoder decision. However, some of the gain may be due to the fact that the rate/quality is no longer as closely matching as in the previous proposal (see notes under JVET-AB0023).

Investigate in EE, replacing the previous EE1-2.3.

[JVET-AB0101](https://jvet-experts.org/doc_end_user/current_document.php?id=12028) AHG11: Lightweight CNN filter for RPR-based SR with Wavelet Decomposition [H. Lan, C. Jung (Xidian Univ.), Y. Liu, M. Li (OPPO)]

This contribution proposes a convolutional neural network (CNN) filter for reference picture resampling (RPR) based super-resolution (SR) with wavelet decomposition. The proposed CNN filter takes the LR reconstructed frame (), LR prediction frame () and RPR upsampled frame () as the input for RPR-based SR. We adopt wavelet decomposition to make the same size as and as well as obtain the relationship between high frequency and low frequency components. Thus, the proposed CNN filter not only learns a mapping function between LR and HR images, but also effectively removes blocking artifacts in the reconstructed frame. Experimental results show that the proposed CNN filter in Y channel achieves -8.98% and -4.05% BD-rate reductions in AI and RA configurations over VTM-11.0\_NNVC-2.0 anchor, respectively.

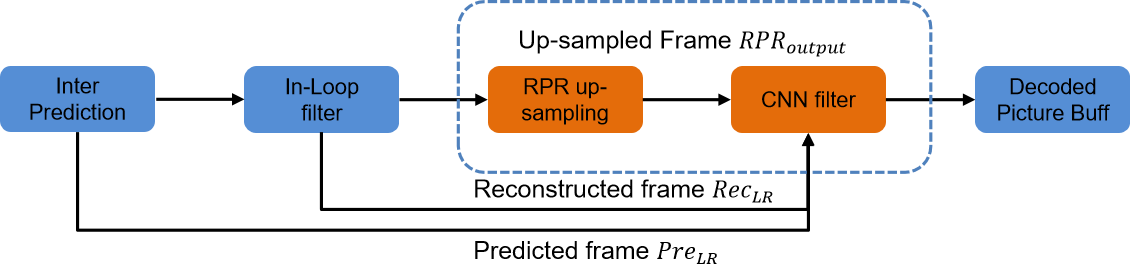


Illustration of the proposed CNN filter for RPR-based SR. The proposed CNN filter takes , and as the input.

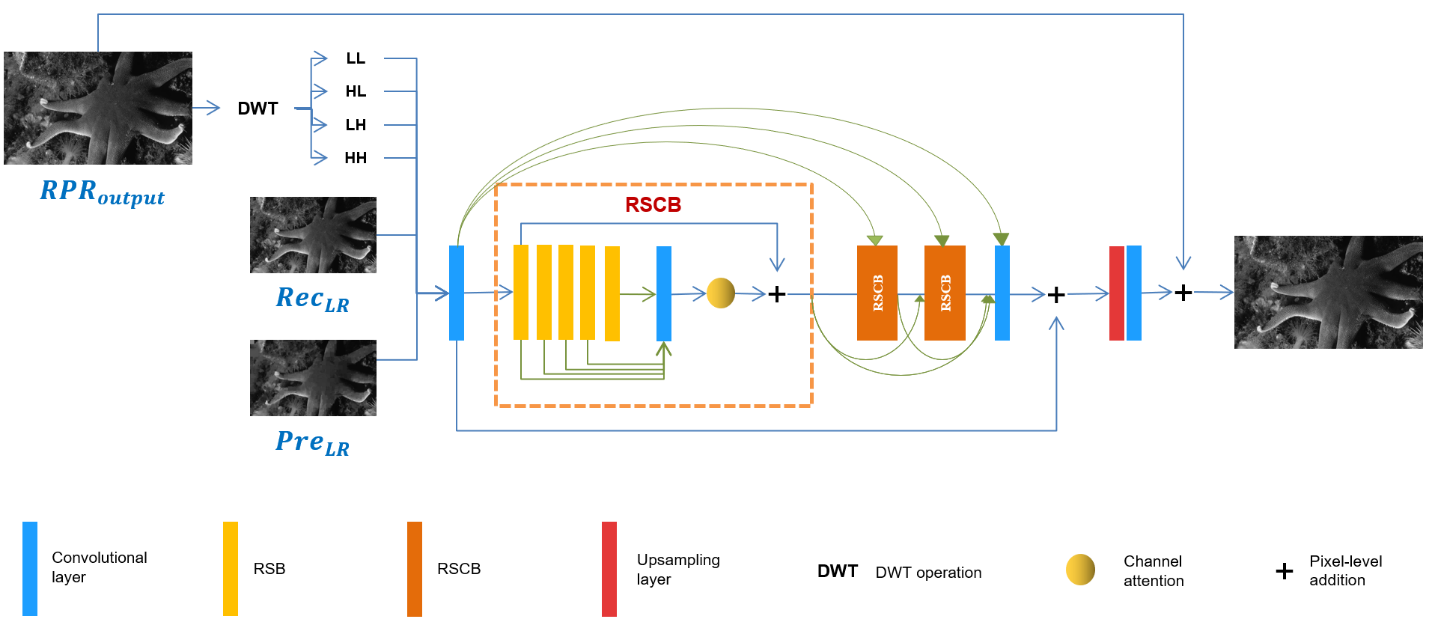
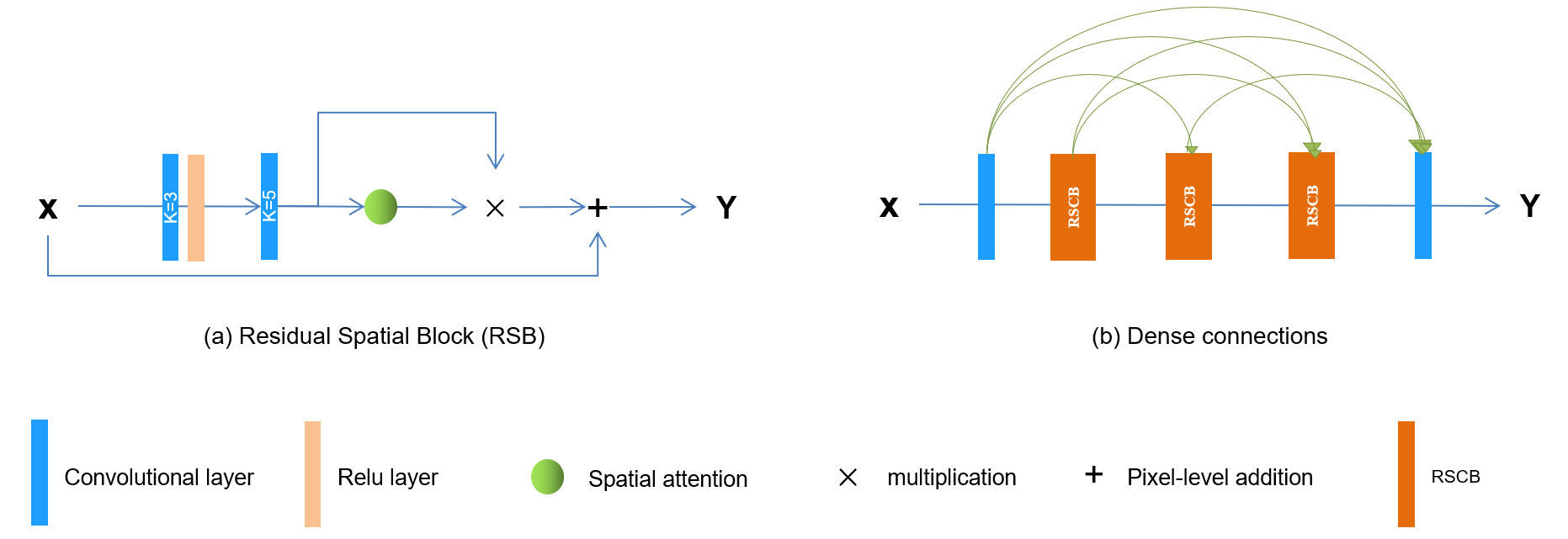
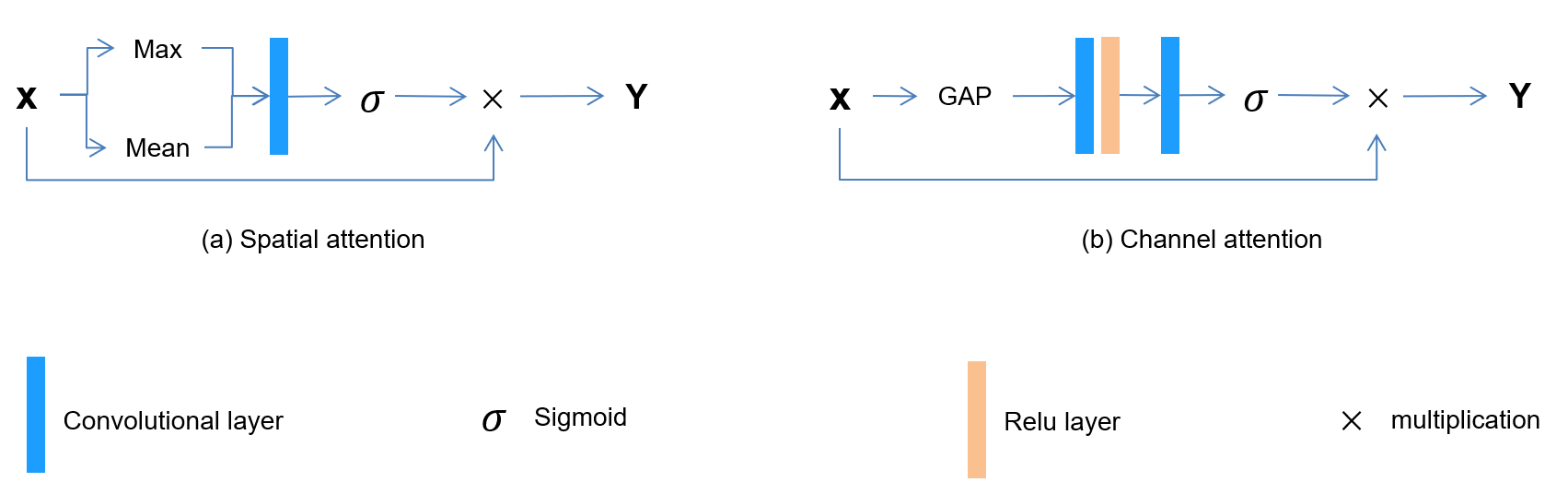


Illustration of the proposed network architecture for RPR based super-resolution for video compression with wavelet decomposition. DWT: Discrete wavelet transform.



(a) Residual spatial block (RSB). (b) Dense connections.



(a) Spatial attention block. (b) Channel attention block

|  |  |  |
| --- | --- | --- |
| **Network Information in Training Stage** | | |
| Mandatory | GPU Type | GPU: NVIDIA GTX 3090 |
| Framework: | PyTorch v1.1.0 |
| Number of GPUs per Task | 1 |
|  |  |
| Epoch: | 300 |
| Batch size: | 64 |
| Training time: | ~36h/model |
| Training data information: | BVI-DVC, DIV2K |
| Training configurations for generating compressed training data (if different to VTM CTC): | VTM-11.0, NNVC-2.0 QP {22, 27, 32, 37, 42} |
|  | Loss function: | Weighted L1 and L2 |
| Optional |  |  |
| Number of iterations | 1000/epoch |
| Patch size | 128x128 |
| Learning rate: | 1e-4 |
| Optimizer: | ADAM |
| Preprocessing: | random flipped |
| Other information: |  |
|  |  |

|  |  |  |
| --- | --- | --- |
| **Network Information in Inference Stage** | | |
| Mandatory | HW environment: | |
| GPU Type | CPU only |
| Framework: | Libtorch v1.9.0 |
| Number of GPUs per Task | 0 |
|  |  |
| Number of Parameters (Each Model) | 2.38 M |
| Total Number of Parameters (All Models) | 14 M total |
| Parameter Precision (Bits) | 32 |
| Memory Parameter (MB) | 9.6 M total |
| Multiply Accumulate (MAC)/pixel | 2193.9 K/pixel |
| Optional |  |  |
| Total Conv. Layers | 57 Common convolution |
| Total FC Layers | 0 |
| Total Memory (MB) |  |
| Batch size: | 1 |
| Patch size | Whole frame |
| Changes to network configuration or weights required to generate rate points |  |
| Peak Memory Usage (Total) |  |
| Peak Memory Usage (per Model) |  |
| Border handling |  |
| Other information: |  |
|  |  |

According to the first two figures above, the upsampling would be in the loop. Is this really the case? At least in the first overall figure, a downsampling from the reference picture to the prediction seems missing.

Information about encoder/decoder run time. Also, from PSNR plots shown, rate/quality matching with the full-resolution anchor is not observed. Quality is below the FR anchor for some higher rates.

The approach of having the NN compute a residual on top of the RPR upsampling could be interesting. However, the method is not mature enough compared to other methods that are in the EE – further study recommended.

[JVET-AB0107](https://jvet-experts.org/doc_end_user/current_document.php?id=12034) Non-EE1: CNN-based super resolution with luma-only rescaling [C. Lin, Y. Li, J. Li, K. Zhang, L. Zhang (Bytedance)]

This contribution combines CNN-based super resolution and CNN-based loop filter (CNNLF) in the neural common software (NCS), to support luma-only rescaling. When CNNLF (filter set #1) in NCS-1.0 is turned on, the simulation results reportedly show {-8.63%, -7.82%, -8.34%} BD-rate savings for {Y, Cb, Cr} under AI configurations, respectively. When CNNLF is turned off, the simulation results reportedly show {-4.91%, -7.06%, -7.84%} BD-rate savings for {Y, Cb, Cr} under AI configurations, respectively, compared with NNVC-2.0.



The proposed model for luma up-sampling.

|  |  |  |
| --- | --- | --- |
| **Network Information in Inference Stage** | | |
| Mandatory | HW environment: | Intel® Xeon ® Platinum 8260 CPU @ 2.40GHz |
| Framework: | LibTorch v1.10.1 |
| Number of GPUs per Task | 0 |
| Number of Parameters (Each Model) | 1.37 M |
| Total Parameter Number | 1.37 M |
| Parameter Precision (Bits) | 32 (F) |
| Memory Parameter (MB) | 5.22 |
| Multiply Accumulate (MAC) | 342 kMAC/pixel |
| Optional | Total Conv. Layers | 35 for up-sampling the luma, 37 for up-sampling the chroma |
| Total FC Layers | 0 |
| Batch size: | 1 |
| Patch size | Whole frame |

|  |  |  |
| --- | --- | --- |
| **Network Information in Training Stage** | | |
| Mandatory | GPU Type | GPU: Tesla-V100-SXM2-32GB |
| Framework: | PyTorch v1.10 |
| Number of GPUs per Task | 1 |
| Epoch: | 120 |
| Batch size: | 16 |
| Training time: | 64 h/model |
| Training data information: | DIV2K |
| Training configurations for generating compressed training data (if different to VTM CTC): | VTM-11.0 + new MCTF, QP {22, 27, 32, 37, 42} |
|  | Loss function: | L1 |
| Optional | Number of iterations |  |
| Patch size | 128\*128 |
| Learning rate: | 1e-4 |
| Optimizer: | ADAM |
| Preprocessing: |  |
| Other information: |  |

Low-resolution coding is performed in 4:4:4 domain. This might be difficult to realize practically (at lest with current principles of profiles and related DPB management). Chroma QP offset is adjusted with the goal of shifting bit rate from luma to chroma. Rate/quality matching is not observed in the current results.

In general, it is difficult to compare approaches which do the low resolution coding with different luma/chroma sampling schemes. All methods considered so far (including RPR) are using 4:2:0 coding in the downsampled video (RPR experiences losses in chroma by doing that, while some of the NN based methods also gain on chroma).

Difficult to decide if the proposed method is having benefit in terms of NN based SR, as it cannot be compared to other methods from the presented results.

[JVET-AB0108](https://jvet-experts.org/doc_end_user/current_document.php?id=12035) AHG11: ALF-SPLIT for NCS2.0 [W. Zou, Y. Zhou, C. M. Gu (Xidian University), C. Huang, Y. X. Bai, Y. J. Zhang (ZTE Corporation)]

An ALF split scheme was proposed in JVET-Y0046 which proposed to signal ALF enabled/disable flag separately for luma and chroma components. It is reported that, with NCS-1.0 as anchor and test platform, the overall performance are as follows:

NCS-1.0 (filter set #1)

* AI: 0.02%/-0.71%/-0.87%

NCS-1.0 (filter set #0)

* AI: 0.01%/-0.25%/-0.14%

Gain only for some sequences. Modification of conventional tools should not be a primary goal in the exploration on neural networks.

No action.

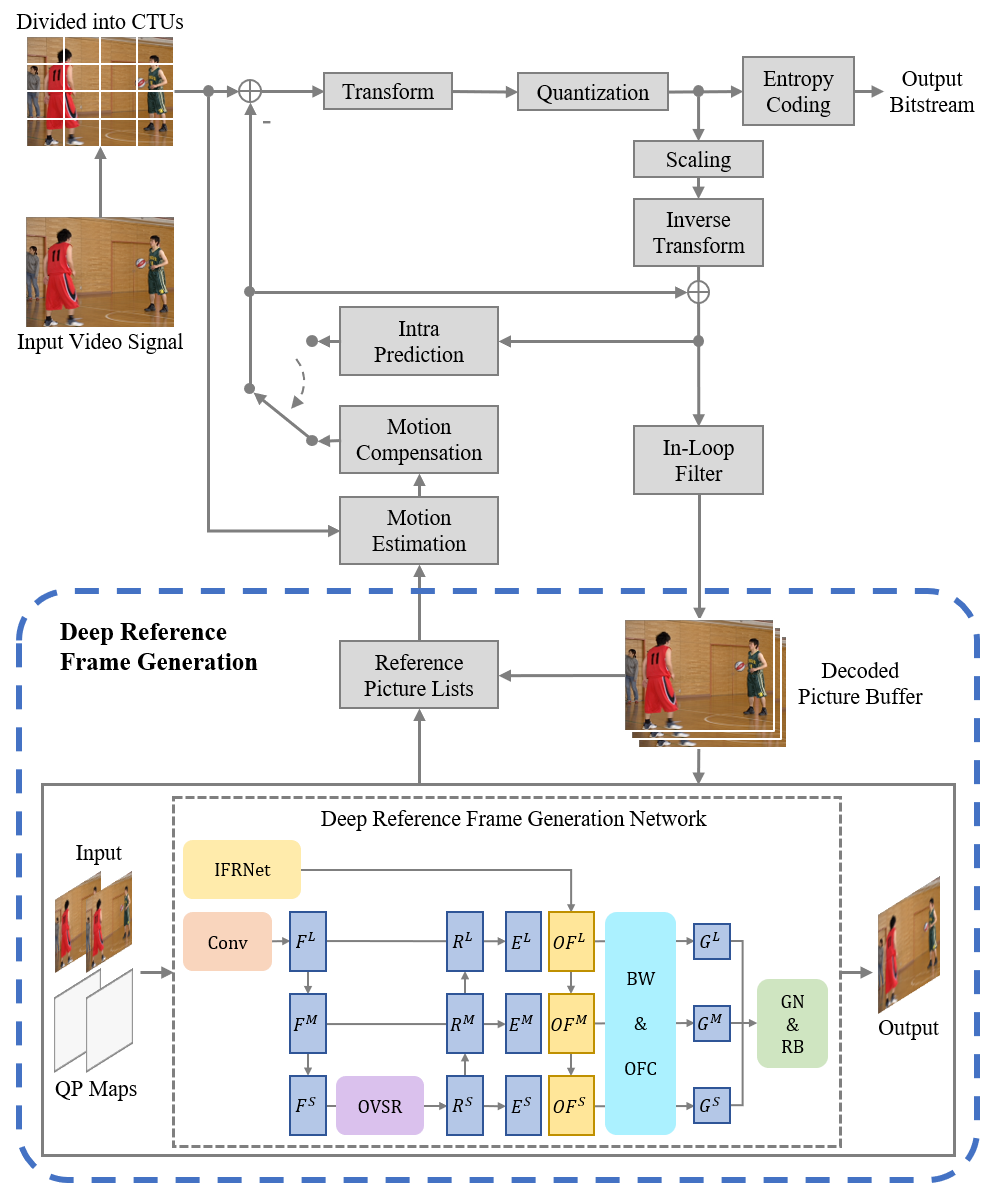
[JVET-AB0109](https://jvet-experts.org/doc_end_user/current_document.php?id=12036) AHG11: Content-adaptive NN post-filter with new QP normalisation [M. Santamaria, F. Cricri, M. M. Hannuksela (Nokia)]

This contribution proposes a change in the normalisation of the slice QP used in the neural-network post-filter in JVET-Z0082. Compared to VTM 11.0 NNVC 1.0, in the RA configuration, it is reported that the coding gains for class B are -5.42% (Y), -21.75% (U), -17.93% (V); and for class C are -4.00% (Y), -19.5% (U), -16.44 (V) compared to VTM 11.0 NNVC 1.0. Compared to the JVET-Z0082, in the RA configuration, it is reported that the BD-rates for class B are -0.09% (Y), -0.66% (U), -0.58% (V); and for class C are -0.01% (Y), 0.05% (U) and 0.28% (V).

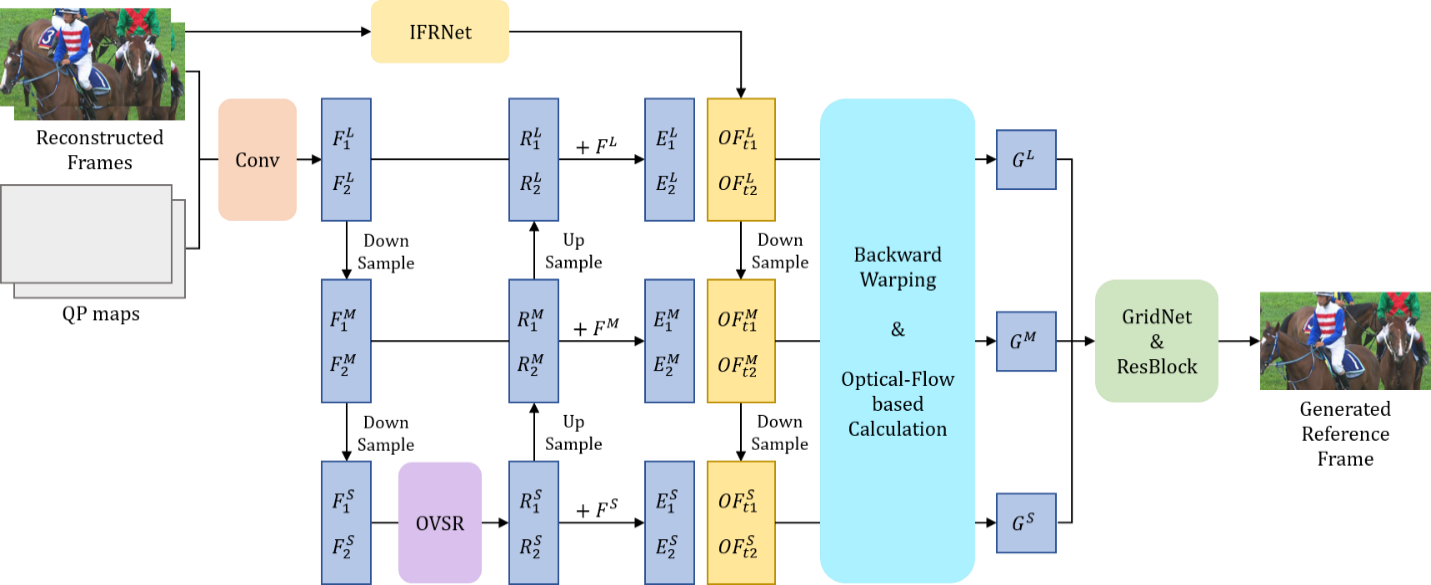
Contribution for information – no specific action requested.

[JVET-AB0114](https://jvet-experts.org/doc_end_user/current_document.php?id=12041) AHG11: Deep Reference Frame Generation for Inter Prediction Enhancement [J. Jia, Y. Zhang, H. Zhu, Z. Chen (Wuhan Univ.), Z. Liu, L. Wang, X. Xu, S. Liu (Tencent)]

This contribution presents a deep reference frame generation (DRF) method for VVC inter prediction enhancement. During the encoding and decoding process, the network receives two reconstructed frames from the decoded picture buffer as inputs and generates a new frame. The generated frame will be put into the reference picture lists as an additional reference for the current encoding picture, which enhances the inter prediction. Different from the previous work JVET-AA0082, this contribution proposes a new network to synthesize the reference frame both for RA and LDB. It is reported that this method can bring {LDB: -2.57%/-12.87%/-11.43%, RA: -3.24%/-10.66%/-10.18%} bitrate savings for Y/U/V components respectively, based on VTM-11.0\_nnvc-2.0.



The frame work of the proposed DFR method



The architecture of the DFR network

Different models used for RA and LB.

|  |  |  |
| --- | --- | --- |
| **Network Information in Training Stage** | | |
| Mandatory | GPU Type | GTX 1080ti 12GB |
| Framework: | PyTorch v1.9.0 |
| Number of GPUs per Task | 8 |
|  |  |
| Epoch: | 100 |
| Batch size: | 256x256 |
| Loss function: | L1 |
| Training time (for 1 model): | 7 Days |
| Training data information: | Vimeo (compressed) |
| Training configurations for generating compressed training data (if different to VTM CTC): | QP = 22, 27, 32, 37, 42 |
| Optional |  |  |
| Number of iterations |  |
| Patch size | 256x256 |
| Learning rate: | 1e-4 |
| Learning rate update strategy |  |
| Optimizer: | ADAM |
| Preprocessing: |  |
| Mini-batch selection process: |  |
| Training data update strategy: |  |
| Other information: |  |
|  |  |

|  |  |  |
| --- | --- | --- |
| **Network Information in Inference Stage** | | |
| Mandatory | HW environment: | |
| GPU Type | CPU only |
| Framework: | PyTorch v1.9.0 |
| Number of GPUs per Task | 0 |
|  |  |
| Number of Parameters (Each Model) | 11749K (RA) 12607K (LDB) |
| Total Number of Parameters (All Models) | 24356K |
| Parameter Precision (Bits) | 32 |
| Memory Parameter (MB) | 44.9 (RA) 48.2 (LDB) |
| Multiply Accumulate (kMAC/pixel) | 659 (RA) 765(LDB) |
| Calculation Method | On a block basis |
| Optional |  |  |
| Total Conv. Layers |  |
| Total FC Layers |  |
| Total Memory (MB) |  |
| Batch size: |  |
| Patch size | 256 x 256 |
| Changes to network configuration or weights required to generate rate points |  |
| Peak Memory Usage (Total) |  |
| Peak Memory Usage (per Model) |  |
| Border handling |  |
| Other information: |  |
|  |  |

Proponents would like to further improve performanc and reduce the complexity before joining EE.

[JVET-AB0121](https://jvet-experts.org/doc_end_user/current_document.php?id=12048) AHG11: Assistant Reference Picture Method for NNVC [C. M. Gu, W. Zou, Y. Zhou, J. W. Fan (Xidian Univ.), Y. X. Bai, C. Huang, Y. J. Zhang (ZTE)]

This contribution presents an assistant reference picture for NNVC to improve the coding efficiency. Besides the reconstructed picture in NNVC, the picture after DBF will put into the decoded picture buffer, using as an assistant reference picture for the future pictures. It is reported that, with NCS-1.0 (NnlfOption=2) as anchor, the performance under RA configuration are as follows:

RA：ClassB -0.01%/-0.03%/0.08%

ClassC -0.02%/0.01%/0.04%

Not directly related to NN based tools.

Is the additional reference ever used? Was not investigated.

No benefit – no action.

[JVET-AB0125](https://jvet-experts.org/doc_end_user/current_document.php?id=12052) AHG11 - CompressAI models integration using SADL [F. Galpin, F. Levebvre, F. Racapé (InterDigital)]

In JVET-AA00063, a hybrid video codec was presented were Intra frames were coded using an end-to-end auto-encoder. To facilitate the evaluation of such approaches and comparing with traditional compression methods, this contribution presents a SADL implementation of the inference of an auto-encoder trained with CompressAI [1].

Results on very low-complexity auto-encoders suitable for hardware implementation are presented. Improvement on entropy coding of the model also shows improvement up to -18% bdrate gains on the kodak dataset, compared to the original entropy coding.

[1] J. Bégaint; F. Racapé; S. Feltman; A. Pushparaja - Compressai: a pytorch library and evaluation platform for end-to-end compression research, 2020/11/5, arXiv preprint arXiv:2011.03029

RDOQ optimization consists of three passes, no back propagation.

Both float and integer models of autoencoders could be used.

It was suggested to investigate the integerization in EE (intra coding)

Decision(SW): Adopt JVET-AB0125 in new version of SADL

[JVET-AB0126](https://jvet-experts.org/doc_end_user/current_document.php?id=12053) AhG11 - SADL update [F. Galpin, T. Dumas, P. Bordes, E. François (InterDigital)]

This contribution presents updates in the Small AdHoc Deep Learning (SADL) library already described in JVET-W0181, JVET-Y0110, JVET-Z0161 and JVET-AA0086.

No need for presentation – just for information

It was suggested to introduce a versioning of SADL.

[JVET-AB0136](https://jvet-experts.org/doc_end_user/current_document.php?id=12063) AHG11: Complexity Reduction on Neural-Network Loop Filter [J. N. Shingala, S. Kadaramandalgi, A. Shyam (Ittiam), T. Shao, A. Arora, P. Yin, Sean McCarthy (Dolby)]

JVET-AA0080 proposes to use CP decomposition plus fusing adjacent 1x1 convolution to reduce the complexity of NNLF. In addition, it also proposes to split architecture for luma and chroma components between input network and output network. Simulation shows that using JVET-X0140 low complexity model (EE1-1.4.1) as baseline, whose worse case block level complexity is 33.6 KMAC/Pixel, the proposal can achieve the AI luma gain about 5% with worst case complexity of 17.7 KMAC/pixel for (24L,8C) split. This contribution applies the same techniques proposed in JVET-AA0080 and investigates the RA case. The same baseline model shows the BD-rate saving of 4.98%,5.08%,3.36% (Y, Cb, Cr, respectively) for RA compared to EE1 NNVC-2.0 anchor. For the proposed complexity reduction techniques, when compared against EE1 NNVC-2.0 anchor, the BD-Rate saving for RA (Y, Cb Cr, respectively) is:

* CP decomposition with fusing: 4.45%,5.68%,5.19% with worst case block level complexity of 16.2 KMAC/Pixel
* Split Luma Chroma architecture (24L, 8C) + CP decomposition with fusion: 4.66%,8.78%,7.81% with worst case block level complexity of 17.7 KMAC/Pixel
* Split Luma Chroma architecture (20L, 8C) + CP decomposition with fusion: 4.11%,8.15%,5.62% with worst case block level complexity of 14.2 KMAC/Pixel

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

1x1 conv 1x1xRxK

1x3 Sep conv 1x3xRxR

CP decomposition

Layers approximated with CP decomposition

Separable Convolution layers

1x1 conv 1x1xMxR

1x1 conv 1x1xRxM

* CP Decomposition + fusion of 1x1 conv layers of JVET-X0140 Baseline Model

Ky

NxN

KC

Channel Concat

LY

NxN

LC

3x3 conv 3x3x10xM

Leaky ReLu

3x3 conv 3x3xKYxLY

1x1 conv 1x1xMxM

Leaky ReLu

3x3 conv 3x3xKxK

1x1 conv 1x1xMxK

Leaky ReLu

3x3 conv 3x3xKYxKY

1x1 conv 1x1xMYxKY

BS info

QStep

Yx4+U+V 72x72x10

NxN

1x1 conv 1x1xKYxMY

Leaky ReLu

3x3 conv 3x3xKYxKY

1x1 conv 1x1xMYxKY

Y’x4+U’+V‘ 64x64x6

N’xN’

***nY*** Hidden Layers (Luma)

***nC*** Hidden Layers (Chroma)

1x1 conv 1x1xMCxKC

1x1 conv 1x1xKCxMC

3x3 conv 3x3xKCxKC

1x1 conv 1x1xMCxKC

3x3 conv 3x3xKCxKC

1x1 conv 1x1xKCxMC

Leaky ReLu

3x3 conv 3x3xKCxLC

Leaky ReLu

1x1 conv 1x1xKYxMY

Layers approximated with CP decomposition

Separable Convolution layers

3x3 conv 3x3xKYxKY

Rank RY CP decomposition

1x1 conv 1x1xKYxRY

3x1 Sep conv 3x1xRYxRY

1x1 conv 1x1xRYxKY

1x3 Sep conv 1x3xRYxRY

Illustration of CP decomposition in Luma layer hidden unit

* **Split luma and chroma model + CP Decomposition + fusion of 1x1 conv layer**
* **Block-level complexity measure in KMacs/Pixel for split luma and chroma model + CP + fusion**

|  |  |
| --- | --- |
| Split luma+chroma model | complexity in KMAC/Pixel |
| CP + fusion | 16.2 |
| (24L, 8C) | 17.7 |
| (20L, 8C) | 14.2 |
| (16L, 8C) | 11.3 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Network Information in Training Stage** | | | | | | | | |
| Mandatory | | GPU Type | Nvidia GeForce RTX3090 | | | | | |
| Framework: | Tensorflow 2.8.0 | | | | | |
| Number of GPUs per Task | 1 | | | | | |
|  |  | | | | | |
| Epoch: | (30-50 per each RA stage)  Baseline CP Fused - Stage1:50 Stage2:50  Split LC CP Fused (24,8) - Stage1:50 Stage2:30  Split LC CP Fused (20,8) - Stage1:30 (Stage2 pending) | | | | | |
| Batch size: | 16 | | | | | |
| Training time: | 200hours per each RA stage | | | | | |
| Training data information: | BVI-DVC | | | | | |
| Training configurations for generating compressed training data (is different to VTM CTC): | QP=20,25,30,35,40,45 | | | | | |
| Optional | |  |  | | | | | |
| Number of iterations |  | | | | | |
| Patch size | 72x72 (4 luma and 2 chroma planes) | | | | | |
| Learning rate: | Epoch : LR, 0-20 1E-4, 20 to 30 1E-5, >30 1E-6 | | | | | |
| Optimizer: | ADAM | | | | | |
| Loss function: | 12-1-1 Weighted L2 | | | | | |
| Preprocessing: | Convert 144x144 YUV420 signal to 6 72x72 blocks. Normalize to 0~1 | | | | | |
| Other information: |  | | | | | |
|  |  | | | | | |
| **Network Information in Inference Stage** | | | | | | | |
|  | | | | **Split LC (16L, 8C)** | **Split LC (20L, 8C)** | **Split LC (24L, 8C)** | |
| Mandatory | HW environment: | | | | | | |
| GPU Type | | | CPU only | | | |
| Framework: | | | Tensorflow 2.8.0 | | | |
| Number of GPUs per Task | | | 0 | | | |
|  | | |  | | | |
| Number of Parameters (Each Model) | | | 36974 | 46066 | | 57014 |
| Total Parameter Number | | | 73948 | 92132 | | 114028 |
| Parameter Precision (Bits) | | | Floating (32bits) | | | |
| Memory Parameter (MB) | | | 0.282 | 0.368 | | 0.435 |
| MAC/Pixel (Kilo) | | | 11.3 | 14.2 | | 17.7 |
| Optional |  | | |  | | | |
| Total Conv. Layers | | | 88 | | | |
| Total FC Layers | | | 0 | | | |
| Total Memory (MB) | | |  | | | |
| Batch size: | | |  | | | |
| Patch size | | | 72x72 (4 luma and 2 chroma planes ) | | | |
| Changes to network configuration or weights required to generate rate points | | |  | | | |
| Peak Memory Usage (Total) | | |  |  | | |
| Peak Memory Usage (per Model) | | |  |  | | |
| Other information: | | |  | | | |

No implementation in SADL so far.

Luma/chroma split in hidden layers for reduced complexity by independent filtering.

CPU runtime approx. 50x compared to NNVC2.0

Number of layers increased, one expert mentions that this might increase latency in hardware.

0.5% coding efficiency lost relative to original method.

Investigate in EE. Would be desirable to also test with integerization.

[JVET-AB0149](https://jvet-experts.org/doc_end_user/current_document.php?id=12076) Non-EE1: neural network-based intra prediction with learned mapping to VVC intra prediction modes [T. Dumas, F. Galpin, P. Bordes (InterDigital)]

This contribution proposes a neural network-based intra prediction mode with learned mapping to the VVC intra prediction modes. Thanks to this learned mapping, for the current luma CB predicted with the neural network-based intra prediction mode, the neural network-based intra prediction mode can be represented by a VVC intra prediction mode, and this VVC mode index can be involved in the generation the MPM list of a luma CB being neighbor of the current luma CB.

A low-complexity neural network-based intra prediction mode is proposed with a complexity of 7.7kMAC/pixel (worst-case). The inference uses 16 bits integer model.

On top of VTM-11-NNVC, this neural network-based intra prediction, using the proposed learned mapping yields -3.61%, -3.16%, -3.27% and -1.75%, -0.75%, -1.08% BD-rate gains for Y, Cb and Cr in AI and RA respectively.

Results were also provided using the NCS-1.0 software as a base (with the NN-filter-set 0). The method yields -3.35%, -3.60%, -3.65% and -1.58%, -1.02%, -1.20% of average BD-rate gains for Y, Cb, and CR in AI and RA respectively.

More information about the network includes:

|  |  |  |
| --- | --- | --- |
| **Network information in training stage** | | |
| Mandatory | GPU type | GPU: Tesla-P100-16GB |
| Framework | Tensorflow 2.6 |
| Number of GPUs per task | 1 |
| Epoch | 15 |
| Batch size | 128 |
| Training time | 8 hours |
| Training data information | The set of images used to create the training sets combines frames extracted from the BVI-DVC dataset, the TVD dataset, the UVG dataset, and the RGB images in the ILSVCR2012 dataset set and DIV2-K, converted into YCbCr. |
| Training configurations for generating compressed training data | VTM-8.0; QP {22, 27, 32, 37} |
| Patch size | The neural network predicting blocks is trained on pairs of a block and its context. |
| Learning rate | 1e-4 |
| Optimizer | ADAM |
| Loss function | Mean-squared error. A l1-norm weight decay term is added to the loss function for training the neural network. |
| Preprocessing | See [1]. |

One participant asked if the proponent had evaluated the performance of the method with filter set 1. It was reported that this hadn’t been evaluated.

It was commented that a similar method had been included in a previous EE. However, the previous EE had not included training cross-checks.

Decision: Include in EE1.

[JVET-AB0158](https://jvet-experts.org/doc_end_user/current_document.php?id=12085) [AHG11] On chroma order adjustment in NNLF [Z. Dai, Y. Yu, H. Yu, D. Wang (OPPO)]

This contribution proposes a chroma order adjustment method based on JVET-AA0088. In JVET-AA0088, the chroma information including U and V components are part of the inputs/outputs to the neural network. It is proposed to allow the input/output order of U and V components to be switched at the frame level. The switch is signaled in the bit-stream. Compared with JVET-AA0088, the proposed method demonstrates {-0.01%, -1.29%, -0.59%} and {-0.10%, -2.51%, -1.20%} BD-rate savings with the RA and LDB configurations, respectively, when compared to the NCS-1.0 software with NnalfOption=1 (filter set 0).

It was reported that the encoder uses approximately 25% of the CTUs when determining the value of the flag. This was done to reduce encoder complexity.

It was reported that the chroma order adjustment decision was made during inference at the encoder.

It was proposed to include the approach in the next version of NCS for the “filter set 0” configuration.

See additional discussion with JVET-AB0060.

[JVET-AB0204](https://jvet-experts.org/doc_end_user/current_document.php?id=12131) Crosscheck of JVET-AB0158 ([AHG11] On chroma order adjustment in NNLF) [L. Wang (Tencent)] [late]

It was reported that the method was only used for B-slces.

Minor differences were reported for random access. The low delay results were reported to match perfectly.

[JVET-AB0159](https://jvet-experts.org/doc_end_user/current_document.php?id=12086) [AHG11] On adjustment of residual for NNLF [Z. Dai, Y. Yu, H. Yu, D. Wang (OPPO)]

This contribution proposes an adjustment to the residual that is output by the NNLFin JVET-AA0088to further improve the coding performance. In JVET-AA0088, a residual network based NNLF was proposed to predict the residual between the reconstructed image and the original image to get better coding performance. It is proposed to apply an offset to the residuals at the frame level, which is signaled by a flag. Compared with JVET-AA0088, the proposed method provides {-0.09%, -1.06%, -0.32%} and {-0.59%, -1.63%, -0.49%} Y, Cb and Cr BD-rate savings with the RA and LDB configurations, respectively, when compared to the NCS-1.0 software with NnalfOption=1 (filter set 0).

The residual adjustment is performed by decreasing the absolute value of the network output by a constant factor. Thus, the constant factor is subtracted from positive values, and the constant factor is added to negative values. In the experiments, a constant factor of 1 or 2 was considered.

It was observed that the method increases the decoding time. The proponent suggested that this could be because the NNLF was being used more frequencly.

It was reported that different offset values are used for luma and chroma.

A participant asked if the proponents had compared the proposal with residual scaling. It was reported that the proposal was enabled in addition to residual scaling.

[JVET-AB0205](https://jvet-experts.org/doc_end_user/current_document.php?id=12132) Crosscheck of JVET-AB0159 ([AHG11] On adjustment of residual for NNLF) [L. Wang (Tencent)] [late]

It was reported that the method was only used for B-slces.

Minor differences were reported for random access. The low delay results were reported to match perfectly.

[JVET-AB0160](https://jvet-experts.org/doc_end_user/current_document.php?id=12087) [AHG11] Combination of chroma order adjustment and residual adjustment for NNLF [Z. Dai, Y. Yu, H. Yu, D. Wang (OPPO)]

This contribution proposes to combine the method of chroma order adjustment (from JVET-AB0158) and residual adjustment (from JVET-AB0159). Compared with JVET-AA0088, the proposed method demonstrates {-0.08%, -1.97%, -0.90%} and {-0.70%, -3.74%, -1.95%} BD-rate savings for Y, Cb and Cr BD-rate savings with the RA and LDB configurations, respectively, when compared to the NCS-1.0 software with NnalfOption=1 (filter set 0).

One participant commented that the approach could also benefit the filter set 1 configuration. The proponent responded that they have only investigated filter set 0, but that using in the filter set 1 configuration should also be possible and can be tested in the future.

One participant suggested studying the method in the EE.

One participant commented that it would be desireable to reduce the increase in encoding time for the JVET-AB0158 proposal. (The JVET-AB0158 is reported to have an encoding time of 107% of the anchor.)

It was suggested that the JVET-AB0159 proposal may be more appealing since its does not show an increase in encoding time on average.

One proponent suggested included JVET-AB0159 in the next version of the NCS software.

Decision: Include JVET-AB0158 and JVET-AB0159 in the EE1. Test the combination if feasible.

[JVET-AB0206](https://jvet-experts.org/doc_end_user/current_document.php?id=12133) Crosscheck of JVET-AB0160 ([AHG11] Combination of chroma order adjustment and residual adjustment for NNLF) [L. Wang (Tencent)] [late]

It was reported that the method was only used for B-slces.

Minor differences were reported for random access. The low delay results were reported to match perfectly.

The crosschecker expressed some preference for JVET-AB0159.

## AHG12: Enhanced compression beyond VVC capability (72)

### Summary and BoG reports

Contributions in this area were discussed at 1415–1915 on Friday 21 October 2022 (chaired by JRO).

[JVET-AB0024](https://jvet-experts.org/doc_end_user/current_document.php?id=12155) EE2: Summary Report on Enhanced Compression beyond VVC capability [V. Seregin, J. Chen, G. Li, K. Naser, J. Ström, M. Winken, X. Xiu, K. Zhang (EE coordinators)]

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

The software basis for this EE is ECM-6.0, released at <https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tags/ECM-6.0>. ECM-6.0 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-aa-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-aa-ee2/simulation-results> if cross-check reports are not submitted as they are optional for EE tests.

**List of tests**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Tests** | **Tester** | **Cross-checker** |
| **1 Intra prediction** | | | |
| 1.1 | Reduced complexity spatial GPM | InterDigital  K. Naser  [JVET-AB0129](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0129-v1.zip) | Qualcomm  B. Ray  [JVET-AB0212](https://jvet-experts.org/doc_end_user/current_document.php?id=12139) |
| 1.2 | Reducing the number of GPM partition types and intra mode derivation | Ittiam  J. Arumugam  Dolby  T. Lu  [JVET-AB0163](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0163-v1.zip) | OPPO  F. Wang  JVET-AB0218 |
| 1.3 | Adaptive blending for spatial GPM | OPPO  F. Wang  [JVET-AB0154](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0154-v1.zip) | Dolby  T. Lu  JVET-AB0221 |
| 1.4 | Smaller and bigger block sizes for spatial GPM | OPPO  F. Wang  [JVET-AB0154](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0154-v1.zip) | Dolby  T. Lu  JVET-AB0221 |
| 1.5 | Test 1.3 + Test 1.4 | OPPO  F. Wang  [JVET-AB0154](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0154-v1.zip) | Dolby  T. Lu  JVET-AB0221 |
| 1.6 | Combination of spatial GPM tests  1.6a: Test 1.5 + Test 1.2  1.6b: Test 1.6a + Test 1.1 | InterDigital  K. Naser  Ittiam  J. Arumugam  Dolby  T. Lu  OPPO  F. Wang  [JVET-AB0155](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0155-v1.zip) | Alibaba  R.-L. Liao  [JVET-AB0208](https://jvet-experts.org/doc_end_user/current_document.php?id=12135)  Qualcomm  B. Ray  [JVET-AB0212](https://jvet-experts.org/doc_end_user/current_document.php?id=12139)  KDDI  Y. Kidani  [JVET-AB0240](https://jvet-experts.org/doc_end_user/current_document.php?id=12169) |
| 1.7a | CCLM with non-linear term | Alibaba  X. Li  [JVET-AB0091](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0091-v1.zip) | OPPO  L. Xu  [JVET-AB0196](https://jvet-experts.org/doc_end_user/current_document.php?id=12123) |
| 1.7b | CCLM with non-linear term as an additional mode | Alibaba  X. Li  [JVET-AB0091](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0091-v1.zip) | OPPO  L. Xu  [JVET-AB0196](https://jvet-experts.org/doc_end_user/current_document.php?id=12123) |
| 1.7c | CCLM with non-linear term with CCCM parameter derivation | Alibaba  X. Li  [JVET-AB0091](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0091-v1.zip) | OPPO  L. Xu  [JVET-AB0196](https://jvet-experts.org/doc_end_user/current_document.php?id=12123) |
| 1.8a | Gradient linear model with luma value | Alibaba  X. Li  [JVET-AB0092](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0092-v1.zip) | Kwai  C. -W. Kuo  JVET-[AB0227](https://jvet-experts.org/doc_end_user/current_document.php?id=12156)  OPPO  F. Wang  JVET-AB0247 |
| 1.8b | Gradient linear model with luma value as additional modes | Alibaba  X. Li  [JVET-AB0092](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0092-v1.zip) | Kwai  C. -W. Kuo  JVET-[AB0227](https://jvet-experts.org/doc_end_user/current_document.php?id=12156)  OPPO  F. Wang  JVET-AB0247 |
| 1.9 | Self-aware filter estimation for CCLM | Bytedance  K. Zhang  [JVET-AB0169](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0169-v1.zip) | Alibaba  X. Li |
| 1.10 | Template-based multiple reference line intra prediction | OPPO  L. Xu  [JVET-AB0156](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0156-v1.zip) | Alibaba  X. Li  [JVET-AB0199](https://jvet-experts.org/doc_end_user/current_document.php?id=12126) |
| 1.11a | Intra prediction fusion | Qualcomm  H. Wang  [JVET-AB0148](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0148-v1.zip) | Xiaomi  F. Le Léannec  JVET-AB0216 |
| 1.11b | Intra reference fusion | Qualcomm  H. Wang  [JVET-AB0148](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0148-v1.zip) | Xiaomi  F. Le Léannec  JVET-AB0216 |
| 1.12a | Test 1.10 and Test 1.11a | OPPO  L. Xu  Qualcomm  H. Wang  [JVET-AB0157](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0157-v1.zip) | Alibaba  X. Li  [JVET-AB0200](https://jvet-experts.org/doc_end_user/current_document.php?id=12127) |
| 1.12b | Test 1.10 and Test 1.11b | OPPO  L. Xu  Qualcomm  H. Wang  [JVET-AB0157](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0157-v1.zip) | Alibaba  X. Li  [JVET-AB0200](https://jvet-experts.org/doc_end_user/current_document.php?id=12127) |
| 1.13a | Template selection scheme for CCCM modes | Qualcomm  Y.-J. Chang  [JVET-AB0143](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0143-v1.zip) | Nokia  J. Lainema |
| 1.13b | Test 1.13a and CCCM fusion instead of MMLM fusion | Qualcomm  Y.-J. Chang  [JVET-AB0143](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0143-v1.zip) | Nokia  J. Lainema |
| 1.14 | IntraTMP adaptation for camera-captured content | InterDigital  K. Naser  [JVET-AB0130](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0130-v1.zip) | Ofino  D. Ruiz Coll  [JVET-AB0211](https://jvet-experts.org/doc_end_user/current_document.php?id=12138) |
| 1.15 | Horizontal and vertical planar modes | Alibaba  X. Li  [JVET-AB0127](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0127-v1.zip) | WILUS  K. Kim  [JVET-AB0197](https://jvet-experts.org/doc_end_user/current_document.php?id=12124) |
| 1.16 | Picture-level geometry transform | Bytedance  W. Jia  [JVET-AB0165](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0165-v1.zip) | Xiaomi  P. Andrivon |
| **2 Inter prediction** | | | |
| 2.1a | AmvpMerge for the low-delay picture | LGE  H. Jang  [JVET-AB0078](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0078-v1.zip) | ETRI  W. Lim |
| 2.1b | Test 2.1a without template matching-based MV refinement | LGE  H. Jang  [JVET-AB0078](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0078-v1.zip) | ETRI  W. Lim |
| 2.2 | Template matching based BCW index derivation for merge mode | Alibaba  R.-L. Liao  [JVET-AB0079](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0079-v1.zip) | OPPO  K. Sato  JVET-AB0224  Qualcomm  Z. Zhang |
| 2.3 | POC based BCW weights derivation | Qualcomm  Z. Zhang  [JVET-AB0124](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0124-v1.zip) | Alibaba  R. -L. Liao  [JVET-AB0209](https://jvet-experts.org/doc_end_user/current_document.php?id=12136) |
| 2.4a | Test 2.2 + Test 2.3 with removing negative BCW weights | Alibaba  R.-L. Liao  Qualcomm  Z. Zhang  [JVET-AB0140](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0140-v1.zip) | Kwai  W. Chen |
| 2.4b | Test 2.2 + Test 2.3 with negative BCW weights | Alibaba  R.-L. Liao  Qualcomm  Z. Zhang  [JVET-AB0140](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0140-v1.zip) | Kwai  W. Chen |
| 2.5a | Enhanced temporal motion information derivation | Bytedance  L. Zhao  [JVET-AB0118](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0118-v1.zip) | Tencent  L.-F. Chen  [JVET-AB0150](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0150-v1.zip) |
| 2.5b | Enhanced temporal motion information derivation without template matching for SbTMVP | Bytedance  L. Zhao | withdrawn |
| 2.6 | DMVR for affine merge coded blocks | Alibaba  J. Chen  [JVET-AB0112](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0112-v1.zip) | Bytedance  B. Vishwanath  [JVET-AB0226](https://jvet-experts.org/doc_end_user/current_document.php?id=12154) |
| 2.7 | Extended weights for MHP | OPPO  K. Sato  [JVET-AB0153](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0153-v1.zip) | ETRI  D. Kim  [JVET-AB0235](https://jvet-experts.org/doc_end_user/current_document.php?id=12164) |
| **3 Screen content coding** | | | |
| 3.1a | IntraTMP for chroma component | InterDigital  K. Naser  [JVET-AB0131](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0131-v1.zip) | Xiaomi  F. Le Léannec  [JVET-AB0213](https://jvet-experts.org/doc_end_user/current_document.php?id=12140) |
| 3.1b | IntraTMP for chroma components using luma block vector | InterDigital  K. Naser | withdrawn |
| 3.2 | Using block vector derived from IntraTMP for IBC | ETRI  W. Lim  [JVET-AB0061](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0061-v1.zip) | Xiaomi  F. Le Léannec  [JVET-AB0214](https://jvet-experts.org/doc_end_user/current_document.php?id=12141) |
| 3.3 | Using luma and chroma block vectors derived from IntraTMP for IBC | InterDigital  K. Naser  ETRI  W. Lim  [JVET-AB0132](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0132-v1.zip) | Xiaomi  F. Le Léannec  [JVET-AB0215](https://jvet-experts.org/doc_end_user/current_document.php?id=12142) |
| **4 Transform** | | | |
| 4.1 | Modification of LFNST for MIP coded blocks | Xidian MMC Lab.  J.-Y. Huo  W.-H. Qiao  [JVET-AB0067](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0067-v2.zip) | Alibaba  X. Li  [JVET-AB0198](https://jvet-experts.org/doc_end_user/current_document.php?id=12125) |
| **5 In-loop filtering** | | | |
| 5.1 | Extended fixed-filter-output based taps for ALF | Bytedance  W. Yin  [JVET-AB0184](https://jvet-experts.org/doc_end_user/documents/28_Mainz/wg11/JVET-AB0184-v1.zip) | OPPO  L. Xu  JVET-AB0195 |

**Description of tests**

***Intra prediction***

**Test 1.1~1.6: spatial GPM tests**

These tests are based on the spatial GPM mode described in JVET-AA EE2-1.4 (JVET-AA0118), with the original algorithm as follows: a candidate list is built which includes partition split and two intra prediction modes shown in Figure 1. 26 partition modes and up to 11 MPMs of intra prediction modes are used to form the combinations, the length of the candidate list is set equal to 16. The selected candidate index is signalled.

**Figure 1. Spatial GPM candidates**

The list is reordered using template shown in Figure 2. GPM blending process is not used in the template, and SAD between the prediction and reconstruction of the template is used for ordering.



**Figure 2. GPM template**

The SGPM mode is applied to blocks whose width and height meet the same restrictions as in inter GPM.

On top of this base algorithm, Tests 1.1 through to 1.6 test various aspects as summarised in the below table:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Partition modes** | **Intra prediction modes** | **Template size** | **Extended block size** | **Adaptive blending** |
| **Test 1.1** | 26 predefined modes | IPM list with TIMD | 4 | Yes | No |
| **Test 1.2** | Adaptive derivation algorithm | IPM list without TIMD | 1 | Yes | Yes |
| **Test 1.3** | 26 predefined modes | MPM list | 4 | No | Yes |
| **Test 1.4** | 26 predefined modes | MPM list | 4 | Yes | No |
| **Test 1.5** | 26 predefined modes | MPM list | 4 | Yes | Yes |
| **Test 1.6a** | 26 predefined modes | IPM list without TIMD | 1 | Yes | Yes |
| **Test 1.6b** | 26 predefined modes | IPM list with TIMD | 1 | Yes | Yes |

In the Tests 1.1 to 1.6, the following modifications are considered:

* Spatial GPM partition modes:

26 predefined modes

Adaptive derivation algorithm based on the horizontal and vertical gradients ratio

* Intra prediction mode selection:

IPM list with and without TIMD:

For each partition mode, an IPM list is derived for each part using intra-inter GPM list derivation. The IPM list size is 3. In the list, TIMD derived mode is replaced by 2 derived modes with horizontal and vertical orientations (using top or left templates) or TIMD derived mode is excluded.

MPM list:

A uniform MPM list, up to 11 elements, is used for all partition modes.

* Template size (left and above): 1 or 4
* Extended block size:

Spatial GPM is extended to be further applied to 4x8, 8x4, 4x16 and 16x4 blocks, which can be described as 4<=width<=64, 4<=height<=64, width<height\*8, height<width\*8, width\*height>=32.

* Adaptive blending:

Adaptive blending is tested for spatial GPM, where blending depth τ shown in Figure 3 is derived as follows:

* If min(width, height)==4, 1/2 *τ* is selected
* else if min(width, height)==8, *τ* is selected
* else if min(width, height)==16, 2 *τ* is selected
* else if min(width, height)==32, 4 *τ* is selected
* else, 8 *τ* is selected

Chart

Description automatically generated

**Figure 3. GPM blending**

All the Tests 1.1~1.6 do not add additional RDO in the encoder side.

**Test 1.7: CCLM with non-linear term**

In ECM-6.0, CCLM model can be represented as follows:

where represents the predicted value of a chroma sample, represents the reconstructed value of the collocated luma samples which are down-sampled for the case of non-4:4:4 color format and the linear model parameters and are derived based on one row and column reconstructed samples that are adjacent to the current block by the linear minimum mean square error (LMMSE) method.

In the test, a non-linear term of the reconstructed luma sample value into the CCLM method, so that the chroma sample can be predicted by a non-linear model as:

where the model parameters , and are derived from the reconstructed values of the top 2 rows and left 2 columns adjacent samples based on the LDL decomposition method used in CCCM.

In Test 1.7a, the non-linear model is applied to all the 6 CCLM modes when slope adjustment factor is equal to 0 without any additional signaling.

In Test 1.7b, the non-linear model is signaled as an additional mode. When one of the 6 CCLM mode is enabled to the current block, an extra flag is signaled to indicate whether to use the non-linear model.

In Test 1.7c, based on Test 1.7b, the CCCM parameter derivation method is used when the non-linear model is selected, that is 6 rows and columns adjacent samples are used to derive the model parameters.

**Test 1.8: Gradient linear model with luma value**

In ECM-6.0, GLM utilizes the gradient of luma samples to predict a chroma sample as:

where represents the predicted value of a chroma sample, represents the gradient of the corresponding reconstructed luma samples, and the linear model parameters and are derived by adjacent reconstructed samples based on the linear minimum mean square error (LMMSE) method as CCLM.

In the tests, a new GLM mode is evaluated that a chroma sample is predicted based on both the gradient of luma samples and the reconstructed value of the down-sampled luma sample with different parameters:

where the model parameters , and are derived from 6 rows and columns adjacent samples based on the LDL decomposition method as the CCCM mode in ECM-6.0.

For signalling, one flag is signaled to indicate whether GLM is enabled to both Cb and Cr components, and the syntax element that indicates the gradient pattern is coded by truncated unary code.

In Test 1.8a, the original GLM mode is replaced by the tested GLM mode.

In Test 1.8b, the original GLM mode is reserved and the new GLM mode is signalled as an additional mode by signaling one extra flag in the bitstream.

**Test 1.9: Self-aware filter estimation for CCLM**

In the test, N luma downsampling candidates are predefined for CCLM mode. For each downsampling candidate, SAD cost is calculated between the reconstructed neighbors (1 above row and 1 left column) and their corresponding predictor derived by using CCLM with the downsampling candidate. The filter candidate with the least SAD cost is selected as the down-sampling filter to perform the CCLM prediction for the current block as shown in Figure 4.



**Figure 4. Self-aware filter estimation for CCLM**

In the test, number of the downsampling filters is set to 32 and is specified in the below table.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ***F0*** | ***F1*** | ***F2*** | ***F3*** | ***F4*** | ***F5*** | ***F6*** | ***F7*** |
|  |  |  |  |  |  |  |  |
| ***F8*** | ***F9*** | ***F10*** | ***F11*** | ***F12*** | ***F13*** | ***F14*** | ***F15*** |
|  |  |  |  |  |  |  |  |
| ***F16*** | ***F17*** | ***F18*** | ***F19*** | ***F20*** | ***F21*** | ***F22*** | ***F23*** |
|  |  |  |  |  |  |  |  |
| ***F24*** | ***F25*** | ***F26*** | ***F27*** | ***F28*** | ***F29*** | ***F30*** | ***F31*** |
|  |  |  |  |  |  |  |  |

**Test 1.10: Template-based multiple reference line intra prediction**

In template-based multiple reference line intra prediction, instead of signalling the reference line and the intra mode directly, an index to the candidate list is coded to indicate which combination of the reference line and prediction mode is used for coding the current block, a truncated Golomb-Rice coding with a divisor 4 is employed to code selected combinations from the combination list.

The list of 20 candidates is constructed by combining an MPM with the reference line {1, 3, 5, 7, 12}.

The MPM list construction is modified comparing to the regular intra MPM as follows:

* PLANAR mode is excluded from the intra-prediction-mode candidate list
* DC mode is added after the 5 neighboring modes and DIMD modes
* The delta angles from to added to the already included to the list angular modes.

There are 5x10 = 50, which are sorted in the ascending order by SAD cost in the template area shown in Figure 5. Since the extended reference line starts from reference line 1, the area covered by reference line 0 is used for the template cost calculation. The 20 combinations with the least SAD cost form the candidate list.

Text

Description automatically generated with medium confidence

**Figure 5. Template area**

**Test 1.11: Intra prediction fusion**

In this test, intra prediction is formed by fusion intra prediction derived from different reference lines as follows:

* For angular intra prediction modes including the single mode case of TIMD and DIMD, the proposed method derives intra prediction by weighting intra predictions obtained from multiple reference lines represented as , where is the intra prediction from the default reference line and is the prediction from the line above the default reference line. The weights are set as and .
* For TIMD mode with blending, is used for the 1st mode () and is used for the 2nd mode ().
* For DIMD mode with blending, the number of predictors selected for a weighted average is increased from 3 to 6.

Intra prediction fusion is applied to luma blocks when angular intra mode has non-integer slope (required reference samples interpolation) and the block size is greater than 16, it is used with MRL and not applied for ISP coded blocks. PDPC is applied for the intra prediction mode using the closest to the current block reference line.

In Test 1.11b, intra prediction fusion is performed on reference lines instead of prediction blocks. Two reference lines (referred to as and ) are used for intra prediction fusion. of the corresponding intra prediction angle is considered in the fusion process. Each value in the fused reference line () is derived from:

**Test 1.12: Combination of Test 1.10 and Test 1.11**

The combinations of intra MRL related tests are evaluated.

Test 1.12a: Test 1.10 + Test 1.11a

Test 1.12b: Test 1.10 + Test 1.12b

**Test 1.13: Template selection scheme for CCCM modes**

In ECM, there are 3 types of neighboring samples used for CCLM mode derivation: top only, left only, and top and left. The predictors of non-CCLM intra chroma modes, i.e., DM, DIMD, and 4 default modes, can be fused with the predictors of MMLM.

In the tests, two changes are evaluated:

* Extend the CCLM template selection to CCCM to use left, top or top and left templates for model derivation.
* Existed in ECM MMLM fusion is replaced with multi-model CCCM fusion. More specifically, the predictors of a non-CCLM chroma mode can be fused with the predictors of multi-model CCCM mode. The derivation of fusion weights is kept unchanged.

In Test 1.13a, the first change is evaluated. In Test 1.13b, both changes are evaluated together.

**Test 1.14: IntraTMP adaptation for camera-captured content**

In the test, IntraTMP is enabled for camera-captured content with the speedup method applied, where the search area is sub-sampled by a factor of 2, which reduces the template matching search by a factor of 4. After finding the best match, a second refinement process is performed in which another template matching search is performed around the best match with a reduced search range defined as min(width, height)/2 of the current block.

**Test 1.15: Horizontal and vertical planar modes**

In ECM, planar mode predicts the current sample from the reconstructed values of 4 reference samples as shown in Figure 6. Specifically, linear interpolation in the horizontal direction and vertical direction are performed respectively, and the two results are averaged to obtain the predicted sample, as shown in the following equations:



**Figure 6. The reference samples used in planar mode**

In Test 1.15, two additional planar mode: horizontal planar mode and vertical planar mode are tested.

For horizontal planar mode, only the horizontal linear interpolation is performed based on the left reference sample and the top-right reference sample to predict the current sample as:

For vertical planar mode, only the vertical linear interpolation is performed based on the above reference sample and the bottom-left reference sample to predict the current sample as:

Two additional planar modes are only applied to the luma component and are not used for ISP coded blocks. When the current block enables one of the two proposed planar modes, the block's propagation mode is set to the original planar mode.

**Test 1.16: Picture-level geometry transform**

In the test, three geometry transforms are supported at picture level: horizontal flip, vertical flip, and 180°-rotation. It is signaled in the picture header to indicate whether the geometry transform is applied, and which transform is used.

The best geometry transform is selected by rate distortion optimization (RDO) at encoder. To accelerate the encoder, a simplified encoding process is used in RDO and only the selected geometry transform is encoded with the original encoding process. To recover the geometry correlations between luma and chroma components after geometry transforms, the chroma samples may be adjusted correspondingly to a new phase. The encoder will disable geometry transform with resolution smaller than 1080p.

When a geometry transform is selected, the picture will be restored to the original orientation after reconstruction.



In the category of GPM proposals (1.1..1.6), best tradeoff is achieved by 1.6a/1.6b. Both are almost identical except that 1.6b uses the TIMD mechanism for the IPM list derivation (same is used in intra/inter GPM of ECM6). Though 1.6a could be asserted to be slightly less complex, and considering that at this stage of exploration JVET is not performing a detailed optimization in terms of hardware implementability, it was agreed to adopt 1.6b. Decision: Adopt JVET-AB0155 test 1.6b.

1.7x (non-linear CCLM) is not asserted to be a good tradeoff in terms of encoder runtime vs. compression benefit.

1.8x is asserted to be a straightforward extension of the gradient linear model in ECM6 (which is knowmn to be mainly beneficial for screen content, but also enabled for camera captured content, where it gives some benefit in chroma. Both 1.8a and 1.8b have a decent tradeoff, where 1.8b has slightly increased encoder runtime, as it has to decide between two different GLM modes. It has however benefit in compression for screen content. Decision: Adopt JVET-AB0092 test 1.8b.

For 1.9, the gain in compression is lower than in the last meeting (likely due to the inclusion of CCCM and GLM in ECM6). Though not fully reflected in run time, the derivation of subsampling filters adds additional processing, and in general, the tradeoff is not good (even chroma loss for general CTC, only some gain fro screen content). No action to be taken.

1.10-1.12: Combinaton 1.12a provides a reasonable tradeoff (4% enc. RT increase, 0.26% luma rate reduction). Intra reference fusion (1.11b, 1.12b) has losses, gains of 1.10 (MRL with TM, using a similar template method as TIMD) and 1.11a (prediction fusion) are additive. Decision: Adopt JVET-AB0157 test 1.12a.

1.13a is asserted to have an attractive tradeoff compression vs. run time. The additional benefit of 1.13b is marginal. Decision: Adopt JVET-AB0143 test 1.13a

1.14 is a simplification of intra TMP search, which also provides gain for camera-captured content with relatively small increase of encoder run time (Intra TMP is only used for screen content in CTC so far). For screen content, run time of encoder and decoder is reduced with marginal change in performance.

Decision: Adopt JVET-AB0130 test 1.14. Decision(CTC): Enable Intra TMP in classes A-E.

1.15: Significant increase in encoder run time, no good tradeoff.

1.16: The gain is not uniform over the different sequences and classes. Practically, implementation would require an additional picture buffer, and it might introduce some additional delay in encoding and decoding (kind of two-pass coding for intra pictures, but first pass with fast algorithm). Due to the rotation, a shift of chroma phase occurs, which may have some positive impact on cross-component methods for some sequences. There some sequences where losses occur (which may be due to the fast algorithm). It is also commented that this approach does not really fall into the category of block level compression tools that are of prior importance of this exploration. No action.

***Inter prediction***

**Test 2.1: AmvpMerge for the low-delay picture**

In ECM, AmvpMerge mode is designed to apply bilateral matching based candidate reordering. Therefore, AmvpMerge has not been applied for the low-delay picture due to the reference pictures being all forward pictures.

In the test, AmvpMerge is enabled for low-delay pictures by bypassing bilateral matching and template matching based candidate reordering, and MV refinement.

In Test 2.1a, AmvpMerge is enabled with bypassing bilateral matching merge candidate reordering process.

In Test 2.1b, on top of Test 2.1a template matching based MV refinement is bypassed as well.

**Test 2.2: Template matching based BCW index derivation for merge mode**

In the test, BCW index for merge coded CUs is derived according to template matching (TM) cost and bi-predicted weights are extended for merge mode.

Firstly, the bi-predicted weights for merge mode is extended from {-2, 3, 4, 5, 10} to {-4, -3, -2, -1, 1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12} if negative bi-predicted weights are supported and to {1, 2, 3, 4, 5, 6, 7} when negative bi-predicted weights are not supported. The negative bi-predicted weights for non-merge mode {-2, 10} are replaced with positive weights {1, 7}.

Secondly, for a CU coded in one of regular merge, template matching, adaptive decoder-side motion vector refinement and MMVD modes, the choice of bi-predicted weight is derived based on TM cost instead of inferring from neighboring blocks. Given a selected merge candidate, the TM cost values are calculated with different bi-predicted weights, and then, the bi-predicted weight with minimum TM cost value is used to predict the merge CU.

Two tests with (Test 2.2b) and without (Test 2.2a) negative weights are performed.

**Test 2.3: POC based BCW weights derivation**

In the test, BCW weights derived based on POC difference are added as follows:

* If both reference pictures are from the past or from the future relatively to the current picture, and the current picture is a non-low-delay picture, the weight pair (-2, 10) is added.
* Otherwise, the weight pair (2, 6) is added.

The larger value from the weight pair is assigned to the closest POC reference picture or list 0 reference picture when the POC distance is the same. The proposed method is not used with MHP for AMVP mode. One bypass bin is added to indicate whether to use default BCW weight or POC based BCW weight.

**Test 2.4: Combination of BCW tests**

In this test, extended BCW weights from the tests 2.2 and 2.3 are reordered based on TM cost with (Test 2.4b) and without (Test 2.4a) negative weights.

**Test 2.5: Enhanced temporal motion information derivation**

In ECM, TMVP for AMVP mode is derived by fetching motion information from the center or the bottom-right of the collocated block. And a similar strategy is also applied to SbTMVP mode, where the motion information from the left neighboring position is used as a motion shift, which is then employed to obtain TMVPs at sub-CU level.

In the test, two aspects are evaluated:

* Two collocated pictures are used for TMVP derivation, the choices of the collocated pictures are signalled and two reference pictures with the least POC distance relative to the current picture are selected in the encoder implementation.
* The motion shift to locate TMVP is adaptively determined from multiple locations according to the template costs. Two motion shift candidate lists are constructed respectively for the two collocated pictures, only the shift candidates, which point to the collocated picture for which list is being constructed, are considered.

At most 4 SbTMVP candidates are included in the sub-block-based merge list and ARMC is modified such that the SbTMVP candidate with the least template matching cost derived from the first collocated picture is placed in the first entry without reordering, while other SbTMVP candidates are sorted together with AFFINE candidates.

**Test 2.6: DMVR for affine merge coded blocks**

In the test, DMVR is applied to affine merge coded blocks. Affine motion consists of translation and non-translation parts. The first stage of multi-stage DMVR is applied to the translation part of the affine motion such that a translation MV offset is added to all the CPMVs of the candidate in the affine merge list if the candidate meets the DMVR condition. And the MV offset is derived by minimizing the cost of bilateral matching which is the same as conventional DMVR.

The first stage refinement process consists of 3x3 square search pattern used to loop through the search range which is set as [-3, 3] to find the best integer MV offset. And then half-pel search is conducted around the best integer position and an error surface estimation is performed at last to find an optimal MV offset with 1/16 precision.

**Test 2.7: Extended weights for MHP**

In ECM, multi-hypothesis prediction (MHP) has weights {2, -1} / 8.

Extended weight table represented by syntax element add\_hyp\_weight\_idx for MHP is tested in the following configurations:

Test 2.7a: wt = {2, 1, -1, -2} / 8

Test 2.7b: wt = {3, 2, 1, -1, -2, -3} / 8

Test 2.7c: wt = {2, -1, 3, 1, -2, -3} / 8, but for block size smaller than 256 using only {2, -1} / 8

RA LB



Test 2.1a/b: Several experts expressed opinion that enabling AmvpMerge for LB gives some attractive gain, though encoding time is increased. 2.1a is also using TM based MV refinement (which was originally designed for bilateral case), which gives some additional gain. Decision(CTC): JVET-AB0078, enable AmvpMerge (including TM based MV refinement) for LB, test 2.1a

Tests 2.2-2.4: In the combination tests, encoder run time is increased by 7% in RA, 4% in LB. This does not appear to be a reasonable tradeoff compared to <0.2% bit rate reduction. It is reported that currently another encoder is investigated for the combination test, however only partial results are available. Comparing 2.2 and 2.3 against each other, 2.2 has clearly the better tradeoff (around 0.1% both in RA and LB, with no encoder run time increase). Among the two 2.2 variants, 2.2a has only 7 weights, 2.2b has 15. Template matching is used in both, so they both require additional processing, but 2.2a less than 2.2b. Better performance of 2.2b is mainly in LB chroma, which does not appear to be relevant considering the number of weights is more than doubled. It is also noted that the number of weights is increased compared to ECM which uses 5 weights (one of them negative). Decision: Adopt JVET-AB0079, test 2.2a

Test 2.5 (only a was kept) uses two collocated pictures and more offsets for temporal motion info derivation. The cross-check JVET-AB0150 reports an issue related to SBTmvp TM process, and suggests a solution to that. It was further commented by another expert that the proposal seems to remove the CTU row constraint of TMVP that is currently used in ECM. It was asked how much of the gain would be retained if such a constraint would be imposed. After review of JVET-AB0150, the method was not confirmed by the cross-checkers, such that it was concluded not to be included in ECM. Further study.

Test 2.6: First pass of DMVR (CU based) is used in affine coded blocks, as from ECM. The proposal is well understood, and gives a reasonable tradeoff (2% enc/dec run time increase, 0.15% rate reduction). Decision: Adopt JVET-AB0112 Test 2.6.

Test 2.7x: The increased number of weights imposes increases in encoder runtime, which are not justified by the relatively small gains (<0.05% in CTC, slightly more screen content). No action.

***Screen content coding***

**Test 3.1: IntraTMP for chroma component**

In ECM, IntraTMP is applied to luma only. In the test, the mode is extended to chroma components, a single flag is used to indicate whether it is applied for both Cb and Cr components.

**Test 3.2: Using block vector derived from IntraTMP for IBC**

In the test, block vector (BV) derived from the intra template matching prediction (IntraTMP) is used for intra block copy (IBC). The stored IntraTMP BV of the neighboring blocks along with IBC BV are used as spatial BV candidates in IBC candidate list construction.

**Test 3.3: Combination of Test 3.1 and Test 3.2**

This test is a combination of IBC merge mode with block vector differences (Test 3.1) and IBC with reconstruction reordering (Test 3.2), where IBC MBVD coded block does not inherit flip type from a neighbor block.



The IntraTMP used in this experiment is still the version from ECM6 (not the version adopted from EE1.14). The latter one is less complex, but had also very small losses in SC classes such as 0.04-0.06%. Both methods appear to be straightforward extensions, where 3.1 is probably increasing decoder complexity for chroma processing. For RA, test 3.1 does not provide any significant benefit (small loss in class F, small gain in class TGM). Test 3.2 is allowing a BV from a neighboring IntraTMP block to be used in the current IBC block’s BV candidate list, which appears very simple. Decision: Adopt JVET-AB0061 test 3.2.

***Transform***

**Test 4.1: Modification of LFNST for MIP coded blocks**

In the test, LFNST is enabled for MIP coded blocks of width and height greater than or equal to 4 and DIMD process is applied to MIP prediction before the upsampling of the current block to derive intra prediction mode as shown in Figure 7 to identify LFNST kernels.

****

**Figure 7. MIP prediction samples to build histogram of gradients**



Encoder runtime increase in AI is likely due to the additional checks for using small blocks. DIMD processing is the same as already used, but the input is different (MIP prediction before upsampling)

Encoder run time increase in RA is comparably low with almost 0.1% bit rate reduction. For AI, the tradeoff is worse than for some other intra tool adoptions made at this meeting, but still acceptable.

Decision: Adopt JVET-AB0067 Test 4.1

***In-loop filtering***

**Test 5.1: Extended fixed-filter-output based taps for ALF**

In ECM-6.0, an ALF signalled filter consists of 3 filter inputs: spatial samples (i.e., tap #0 ~ #19), samples before deblocking (i.e., tap #20, #21, #24) and output samples of 2 fixed filters (i.e., tap #22, #23) resulting in 25 taps in total as shown in Figure 8.

图片包含 文本

描述已自动生成

**Figure 8. ALF filter in ECM**

In the test, the output samples of the first fixed filter are extended to 5x5 diamond shape as shown Figure 9.

图表

描述已自动生成

**Figure 9. Extended output samples of the first fixed filter**

Results:



This is assessed to be a straightforward change, and reasonable tradeoff enc/dec run time versus compression benefit. Decision: Adopt JVET-AB0184 Test 5.1

### EE2 contributions: Enhanced compression beyond VVC capability (27)

Contributions in this area were discussed at XXXX–XXXX on XXday XX October 2022 (chaired by JRO).

[JVET-AB0061](https://jvet-experts.org/doc_end_user/current_document.php?id=11976) EE2-3.2: Using block vector derived from IntraTMP for IBC [W. Lim, D. Kim, J. Kim, S.-C. Lim (ETRI)]

[JVET-AB0214](https://jvet-experts.org/doc_end_user/current_document.php?id=12141) Cross-check of JVET-AB0061: "EE2-3.2: Using block vector derived from IntraTMP for IBC" [F. Le Léannec (Xiaomi)] [late]

[JVET-AB0067](https://jvet-experts.org/doc_end_user/current_document.php?id=11982) EE2-4.1: Modification of LFNST for MIP coded block [J.-Y. Huo, W.-H. Qiao, X. Hao, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.), J. Ren, M. Li, L.-H. Xu (OPPO)]

[JVET-AB0198](https://jvet-experts.org/doc_end_user/current_document.php?id=12125) Crosscheck of JVET-AB0067 (EE2-4.1: Modification of LFNST for MIP coded block) [X. Li (Alibaba)] [late]

[JVET-AB0078](https://jvet-experts.org/doc_end_user/current_document.php?id=11993) EE2-2.1: AmvpMerge for low delay [H. Jang, N. Park, J. Nam, J. Lim, S. Kim (LGE)]

[JVET-AB0256](https://jvet-experts.org/doc_end_user/current_document.php?id=12185) Crosscheck of JVET-AB0078 (EE2-2.1: AmvpMerge for low delay) [W. Lim, S.-C. Lim (ETRI)] [late] [miss]

[JVET-AB0079](https://jvet-experts.org/doc_end_user/current_document.php?id=12006) EE2-2.2: Template matching based BCW index derivation for merge mode [R.-L. Liao, J. Chen, Y. Ye, X. Li (Alibaba)]

[JVET-AB0224](https://jvet-experts.org/doc_end_user/current_document.php?id=12152) Crosscheck Report of EE2 Test 2.2 (Template matching based BCW index derivation for merge mode [S. Kazushi (OPPO) [late]

[JVET-AB0248](https://jvet-experts.org/doc_end_user/current_document.php?id=12177) Crosscheck of JVET-AB0079 (EE2-2.2: Template matching based BCW index derivation for merge mode) [Z. Zhang (Qualcomm)] [late] [miss]

[JVET-AB0091](https://jvet-experts.org/doc_end_user/current_document.php?id=12018) EE2-1.7: CCLM with non-linear term [X. Li, Y. Ye, R.-L. Liao, J. Chen (Alibaba)]

[JVET-AB0196](https://jvet-experts.org/doc_end_user/current_document.php?id=12123) Crosscheck of JVET-AB0091 (EE2-1.7: CCLM with non-linear term) [L. Xu (OPPO)] [late]

[JVET-AB0092](https://jvet-experts.org/doc_end_user/current_document.php?id=12019) EE2-1.8: Gradient linear model with luma value [X. Li, Y. Ye, R.-L. Liao, J. Chen (Alibaba)]

[JVET-AB0227](https://jvet-experts.org/doc_end_user/current_document.php?id=12156) Crosscheck of JVET-AB0092 (EE2-1.8: Gradient linear model with luma value) [C.-W. Kuo, H.-J. Jhu (Kwai)] [late]

[JVET-AB0247](https://jvet-experts.org/doc_end_user/current_document.php?id=12176) Crosscheck of JVET-AB0092 (EE2-1.8: Gradient linear model with luma value) [F. Wang (OPPO)] [late]

[JVET-AB0112](https://jvet-experts.org/doc_end_user/current_document.php?id=12039) EE2-2.6: DMVR for affine merge coded blocks [J. Chen, R.-L. Liao, X. Li, Y. Ye (Alibaba)]

[JVET-AB0226](https://jvet-experts.org/doc_end_user/current_document.php?id=12154) Crosscheck of JVET-AB0112 (EE2-2.6: DMVR for affine merge coded blocks) [B. Vishwanath (Bytedance)] [late]

[JVET-AB0118](https://jvet-experts.org/doc_end_user/current_document.php?id=12045) EE2-2.5a: Enhanced temporal motion information derivation [L. Zhao, K. Zhang, L. Zhang (Bytedance)]

[JVET-AB0150](https://jvet-experts.org/doc_end_user/current_document.php?id=12077) Crosscheck of JVET-AB0118 (EE2-2.5a: Enhanced temporal motion information derivation) [L.-F. Chen, G. Li, X. Xu, X. Zhao, S. Liu (Tencent)]

This contains an element of a technical proposal

As a general remark, this should have better been submitted as a separate document.

Was presented Tue 25 Oct. 1545.

During the study of the software related to JVET-AB0118, an issue of the template matching process for SbTMVP mode was identified in EE2-2.5a, and a modified version is proposed in this contribution. The modified version on top of EE2-2.5a was tested and the following results were observed over Test2.5a anchor.

RA: {0.00%, 0.06%, 0.01%; 100%, 100%};

LDB: {-0.01%, -0.25%, -0.02%; 99%, 99%}.

See further notes under EE2-2.5a in the EE summar report.

[JVET-AB0232](https://jvet-experts.org/doc_end_user/current_document.php?id=12161) Crosscheck of JVET-AB0150 (Crosscheck of JVET-AB0118 (EE2-2.5a: Enhanced temporal motion information derivation) [L. Zhao (Bytedance)] [late]

[JVET-AB0124](https://jvet-experts.org/doc_end_user/current_document.php?id=12051) EE2-2.3: POC based BCW weights derivation [Z. Zhang, H. Huang, C.-C. Chen, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-AB0209](https://jvet-experts.org/doc_end_user/current_document.php?id=12136) Crosscheck of JVET-AB0124 (EE2-2.3: POC based BCW weights derivation) [R.-L. Liao (Alibaba)] [late]

[JVET-AB0127](https://jvet-experts.org/doc_end_user/current_document.php?id=12054) EE2-1.15: Horizontal and vertical planar modes [X. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba)]

[JVET-AB0197](https://jvet-experts.org/doc_end_user/current_document.php?id=12124) Crosscheck of JVET-AB0127 (EE2-1.15: Horizontal and vertical planar modes) [K. Kim (WILUS)] [late]

[JVET-AB0129](https://jvet-experts.org/doc_end_user/current_document.php?id=12056) EE2-1.1: Reduced Complexity Spatial GPM [K. Naser, Y. Chen, A. Robert, K. Reuzé (InterDigital)]

[JVET-AB0212](https://jvet-experts.org/doc_end_user/current_document.php?id=12139) Crosscheck of JVET-AB0129 (EE2-1.1) and JVET-AB0155 (EE2-1.6) [B. Ray (Qualcomm)] [late]

[JVET-AB0130](https://jvet-experts.org/doc_end_user/current_document.php?id=12057) EE2-1.14: IntraTMP adaptation for camera-captured content [K. Naser, T. Poirier, F. Galpin, A. Robert (InterDigital)]

[JVET-AB0211](https://jvet-experts.org/doc_end_user/current_document.php?id=12138) Crosscheck of JVET-AB0130 (EE2-1.14: IntraTMP adaptation for camera-captured content) [D. Ruiz Coll, V. Warudkar (Ofinno)] [late]

[JVET-AB0131](https://jvet-experts.org/doc_end_user/current_document.php?id=12058) EE2-3.1a: IntraTMP for chroma component [K. Naser, A. Robert, T. Dumas, T. Poirier, F. Galpin (InterDigital)]

[JVET-AB0213](https://jvet-experts.org/doc_end_user/current_document.php?id=12140) Cross-check of JVET-AB0131: "EE2-3.1a: IntraTMP for chroma component" [F. Le Léannec (Xiaomi)] [late]

[JVET-AB0132](https://jvet-experts.org/doc_end_user/current_document.php?id=12059) EE2-3.3: Combination of EE2-3.1a and EE2-3.2 [K. Naser, A. Robert, T. Dumas, T. Poirier, F. Galpin (InterDigital), W. Lim, D. Kim, S-C Lim (ETRI)]

[JVET-AB0215](https://jvet-experts.org/doc_end_user/current_document.php?id=12142) Cross-check of JVET-AB00132: "EE2-3.3: Combination of EE2-3.1a and EE2-3.2" [F. Le Léannec (Xiaomi)] [late]

[JVET-AB0140](https://jvet-experts.org/doc_end_user/current_document.php?id=12067) EE2-2.4: Combined test of Test 2.2 and Test 2.3 on BCW weights derivation [Z. Zhang, H. Huang, C.-C. Chen, V. Seregin, M. Karczewicz (Qualcomm), R.-L. Liao, J. Chen, Y. Ye, X. Li (Alibaba)]

[JVET-AB0236](https://jvet-experts.org/doc_end_user/current_document.php?id=12165) Crosscheck of JVET-AB0140 (EE2-2.4: Combined test of Test 2.2 and Test 2.3 on BCW weights derivation) [W. Chen (kwai)] [late]

[JVET-AB0143](https://jvet-experts.org/doc_end_user/current_document.php?id=12070) EE2-1.13: On CCCM improvement [Y.-J. Chang, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-AB0254](https://jvet-experts.org/doc_end_user/current_document.php?id=12183) Crosscheck of JVET-AB0143 (EE2-1.13: On CCCM improvement) [J. Lainema (Nokia)] [late]

[JVET-AB0148](https://jvet-experts.org/doc_end_user/current_document.php?id=12075) EE2-1.11: Intra prediction fusion [H. Wang, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-AB0216](https://jvet-experts.org/doc_end_user/current_document.php?id=12143) Cross-check of JVET-AB0148: "EE2-1.11: Intra prediction fusion" [F. Le Léannec (Xiaomi)] [late]

[JVET-AB0153](https://jvet-experts.org/doc_end_user/current_document.php?id=12080) EE2-Test2.7: Extended weights for MHP [K. Sato, Y. Yu, H. Yu, D. Wang (OPPO)]

[JVET-AB0235](https://jvet-experts.org/doc_end_user/current_document.php?id=12164) Crosscheck of JVET-AB0153 (EE2-Test2.7: Extended weights for MHP) [D. Kim, S.-C. Lim (ETRI)] [late]

[JVET-AB0154](https://jvet-experts.org/doc_end_user/current_document.php?id=12081) EE2-1.3, 1.4, 1.5: Spatial GPM tests [F. Wang, Y. Yu, H. Yu, D. Wang (OPPO)]

[JVET-AB0221](https://jvet-experts.org/doc_end_user/current_document.php?id=12148) Crosscheck of JVET-AB0154 (EE2-1.3, 1.4, 1.5: Spatial GPM tests) [T. Lu (Dolby)] [late]

[JVET-AB0155](https://jvet-experts.org/doc_end_user/current_document.php?id=12082) EE2-1.6: Combination of spatial GPM tests [F. Wang, Y. Yu, H. Yu, D. Wang (OPPO), A.Natesan, J. N. Shingala, Jeeva Raj Arumugam, Vaibhav Valvaiker (Ittiam), [T. Lu](mailto:tlu@dolby.com), F. Pu, P. Yin, S. McCarthy (Dolby), K. Naser, Y. Chen, A. Robert, K. Reuzé (InterDigital)]

[JVET-AB0208](https://jvet-experts.org/doc_end_user/current_document.php?id=12135) Crosscheck of JVET-AB0155 (EE2-1.6: Combination of spatial GPM tests) [R.-L. Liao (Alibaba)] [late]

JVET-AB0212 is another crosscheck on JVET-AB0155

[JVET-AB0240](https://jvet-experts.org/doc_end_user/current_document.php?id=12169) Crosscheck of JVET-AB0155 (EE2-1.6: Combination of spatial GPM tests) [Y. Kidani (KDDI)] [late]

[JVET-AB0156](https://jvet-experts.org/doc_end_user/current_document.php?id=12083) EE2-1.10: Template-based multiple reference line intra prediction [L. Xu, Y. Yu, H. Yu, D. Wang (OPPO)]

[JVET-AB0199](https://jvet-experts.org/doc_end_user/current_document.php?id=12126) Crosscheck of JVET-AB0156 (EE2-1.10: Template-based multiple reference line intra prediction) [X. Li (Alibaba)] [late]

[JVET-AB0157](https://jvet-experts.org/doc_end_user/current_document.php?id=12084) EE2-1.12: Combination of EE2-1.10 and EE2-1.11 [L. Xu, Y. Yu, H. Yu, D. Wang (OPPO), H. Wang, [V. Seregin](mailto:vseregin@qti.qualcomm.com), B. Ray, M. Karczewicz (Qualcomm)]

[JVET-AB0200](https://jvet-experts.org/doc_end_user/current_document.php?id=12127) Crosscheck of JVET-AB0157 (EE2-1.12: Combination of EE2-1.10 and EE2-1.11) [X. Li (Alibaba)] [late]

[JVET-AB0163](https://jvet-experts.org/doc_end_user/current_document.php?id=12090) EE2-1.2: Test on Spatial GPM [A. Natesan, J. N. Shingala, J. R. Arumugam, V. Valvaiker (Ittiam), T. Lu, P. Yin, F. Pu, T. Shao, A. Arora, S. McCarthy (Dolby)]

[JVET-AB0218](https://jvet-experts.org/doc_end_user/current_document.php?id=12145) Crosscheck of JVET-AB0163 (EE2-1.2: Test on Spatial GPM [F. Wang (OPPO)] [late]

[JVET-AB0165](https://jvet-experts.org/doc_end_user/current_document.php?id=12092) EE2-1.16: Picture-level geometry transform [W. Jia, K. Zhang, Y. Wang, T. Fu, Y. Li, L. Zhang (Bytedance)]

[JVET-AB0207](https://jvet-experts.org/doc_end_user/current_document.php?id=12134) Cross-check of JVET-AB0165 (EE2-1.16: Picture-level geometry transform) [P. Andrivon (Xiaomi)] [late]

[JVET-AB0169](https://jvet-experts.org/doc_end_user/current_document.php?id=12096) EE2-1.9: Self-Aware Filter Estimation for CCLM [K. Zhang, L. Zhang, Z. Deng, T. Fu (Bytedance)]

[JVET-AB0243](https://jvet-experts.org/doc_end_user/current_document.php?id=12172) Crosscheck of JVET-AB0169 (EE2-1.9: Self-Aware Filter Estimation for CCLM) [X. Li (Alibaba)] [late]

[JVET-AB0184](https://jvet-experts.org/doc_end_user/current_document.php?id=12111) EE2-5.1: Extended Fixed-Filter-Output based Taps for ALF [W. Yin, K. Zhang, Z. Deng, L. Zhang (Bytedance)]

[JVET-AB0195](https://jvet-experts.org/doc_end_user/current_document.php?id=12122) Crosscheck of JVET-AB0184 (EE2-5.1: Extended Fixed-Filter-Output based Taps for ALF) [L. Xu (OPPO)] [late]

### EE2 related contributions (17)

Contributions in this area were discussed at 1915–2000 on Friday 21 October 2022 (chaired by JRO), and 1400-1640 on Saturday 22 October 2022 (chaired by JRO).

Also consider JVET-AB0150 in this category.

[JVET-AB0062](https://jvet-experts.org/doc_end_user/current_document.php?id=11977) EE2-related: Modifications of EE2-3.2 (Using block vector derived from IntraTMP for IBC) [W. Lim, D. Kim, J. Kim, S.-C. Lim, J. S. Choi (ETRI), [K. Naser](mailto:karam.naser@interdigital.com), T. Dumas, T. Poirier, F. Galpin, A. Robert (InterDigital)]

This contribution proposes two modifications related to EE2-3.2 and EE2-3.3. EE2-3.2 aims at using intra template matching prediction (IntraTMP) block vector (BV) as intra block copy (IBC) block vector predictor. EE2-3.3 is the combination of EE2-3.1, which applies IntraTMP to chroma components, and EE2-3.2. The proposed modifications for IBC block vector prediction are 1) spatial BV candidate ordering in the IBC BV candidate list and 2) adding IntraTMP BVs to history-based block vector prediction (HBVP) buffer. Compared to ECM-6.0, experimental results of modification 1, modification 2, and combination of both modifications on top of EE2-3.2 and EE2-3.3 are summarized below.

Modification 1 on EE2-3.2

* AI: {-0.09%, -0.13%, -0.05%; 100%, 98%}, RA: {-0.05%, -0.03%, -0.03%; 98%, 98%}

Modification 2 on EE2-3.2

* AI: {-0.10%, -0.14%, -0.02%; 100%, 99%}, RA: {-0.01%, -0.09%, -0.03%; 98%, 100%}

Combination of modification 1 and 2 on EE2-3.2

* AI: {-0.12%, -0.15%, -0.02%; 101%, 102%}, RA: {-0.04%, -0.09%, -0.07%; 100%, 99%}

Modification 1 on EE2-3.3

* AI: {-0.19%, -0.19%, -0.16%; 102%, 103%}, RA: {-0.08%, -0.05%, -0.11%; 100%, 100%}

Modification 2 on EE2-3.3

* AI: {-0.21%, -0.23%, -0.18%; 102%, 101%}, RA: {-0.07%, -0.06%, -0.15%; 98%, 99%}

Combination of modification 1 and 2 on EE2-3.3

* AI: {-0.21%, -0.23%, -0.21%; 102%, 103%}, RA: {-0.05%, -0.06%, -0.17%; 100%, 100%}

The results reported are relative to ECM6 anchor. The additional benefit compared to EE2-3.2 (which was adopted) appears very small (at most 0.05% which is extremely low for screen content, and it seems to be coming with an increase of run times.

No action.

[JVET-AB0194](https://jvet-experts.org/doc_end_user/current_document.php?id=12121) Cross-check of JVET-AB0062 (EE2-related: Modifications of EE2-3.2 and EE2-3.3) [H. Jang (LGE)] [late]

[JVET-AB0103](https://jvet-experts.org/doc_end_user/current_document.php?id=12030) EE2-1.16 related: Modifications of picture-level geometry transform [J. Choi, S. Yoo, J. Lim, S. Kim (LGE)]

This contribution introduces two methods based on picture-level geometry transform. Both methods allow only vertical flip geometry transform. In Method 1, a flag is signalled to indicate if the vertical flip geometry transform is applied or not after comparing RD costs in the encoder side. In Method 2, a picture-level edge feature analysis is applied based on the edge types of pixels to determine if the vertical flip of geometry transform is applied or not. Hence, the RD comparing process required in Method 1 is not necessary in Method 2.

The experimental results of the proposed methods are shown below:

* Method 1 (only vertical flip geometry transform is considered)
  + AI: -0.20% / -0.31% / -0.34% for Y/U/V with 100%/98% enc/dec complexity
  + RA: X.XX% / X.XX% / X.XX% for Y/U/V with XX%/ XX % enc/dec complexity
* Method 2 (the edge feature based geometry transform selection method on top of Method 1)
  + AI: -0.17% / -0.29% / -0.36% for Y/U/V with 100%/97% enc/dec complexity
  + RA: X.XX% / X.XX% / X.XX% for Y/U/V with XX%/ XX % enc/dec complexity

No need for review, as no action was taken on the EE proposal.

[JVET-AB0222](https://jvet-experts.org/doc_end_user/current_document.php?id=12149) Crosscheck of JVET-AB0103 (EE2-1.16 related: Modifications of picture-level geometry transform) [W. Jia (Bytedance)] [late] [miss]

[JVET-AB0104](https://jvet-experts.org/doc_end_user/current_document.php?id=12031) EE2-related: On directional planar prediction [S. Yoo, J. Choi, J. Nam, M. Hong, J. Lim, S. Kim (LGE)]

In this proposal, modified horizontal and vertical planar prediction methods are suggested. Specifically, the additional two modes are considered as horizontal and vertical prediction modes during the reference sample construction and transform processes. In addition, a context-coded flag to indicate the direction between horizontal and vertical is bypass-coded. Further, the directional flag is inferred either horizontal planar or vertical planar mode based on the DIMD mode. If the DIMD first mode is less than 34, the intra mode of the current block is inferred to be a horizontal planar mode. Otherwise, the intra mode of the current block is set to be a vertical planar mode. The experimental results of the proposed methods are shown below:

* Method 1 (To consider new prediction modes as either horizontal or vertical mode for reference sample construction and transform)
  + AI: -0.12% / -0.02% / -0.05% for Y/U/V, 116% / 100% for EncT, DecT
  + RA: X.XX% / X.XX% / X.XX% for Y/U/V, XXX% / XXX% for EncT, DecT
* Method 2 (Using the DIMD mode to specify the direction of the new prediction modes)
  + AI: -0.08% / -0.05% / -0.04% for Y/U/V, 108% / 101% for EncT, DecT
  + RA: -0.05% / 0.02% / -0.09% for Y/U/V, 101% / 100% for EncT, DecT

Compared to EE2-1.15 (which was not adopted), method 1 has better compression (0.12% instead of 0.08%) with same increase of encoder runtime as the original EE proposal, whereas method 2 has same performance with less increase of runtime (8% instead of 15%).

Investigate in EE, along with JVET-AB0110 and JVET-AB0162, but it would be expected to provide a significantly better tradeoff, i.e. further reducing encoder runtime while retaining or improving the compression.

[JVET-AB0253](https://jvet-experts.org/doc_end_user/current_document.php?id=12182) Cross-check of JVET-AB0104 on planar directional planar prediction [V. Seregin (Qualcomm)] [late] [miss]

[JVET-AB0110](https://jvet-experts.org/doc_end_user/current_document.php?id=12037) EE2-1.15-related: Improvements on planar horizontal and planar vertical mode [K. Kim, D. Kim, J.-H. Son, J.-S. Kwak (WILUS)]

This contribution proposes a derivation method of a transform kernel for planar horizontal and planar vertical mode. If an intra prediction mode of a current block is the planar horizontal/vertical mode, the vertical/horizontal intra prediction mode is used to derive a transform kernel in MTS set and LFNST set.

Additionally, this contribution proposes the block size restriction for planar horizontal & vertical mode to reduce the complexity.

On top of ECM-6.0, simulation results of the proposed method are reported as below:

Test 1:

AI: {-0.14%, -0.03%, -0.04%, 115%, 98%}

RA: {}

Test 2:

AI: {}

RA: {}

Investigate in EE, along with JVET-AB0104 and JVET-AB0162, but it would be expected to provide a significantly better tradeoff, i.e. further reducing encoder runtime while retaining or improving the compression.

[JVET-AB0202](https://jvet-experts.org/doc_end_user/current_document.php?id=12129) Crosscheck of JVET-AB0110 (EE2-1.15-related: Improvements on planar horizontal and planar vertical mode) [X. Li (Alibaba)] [late]

[JVET-AB0115](https://jvet-experts.org/doc_end_user/current_document.php?id=12042) EE2-1.14 related: Modifications of MTS and LFNST for IntraTMP coded block [D. Kim, K. Kim, J. Son, J. S Kwak (WILUS), K. Naser, T. Poirier, F. Galpin, A. Robert (InterDigital)]

In ECM-6.0, for a IntraTMP coded block, the LFNST transform set 0 or the block size-based implicit MTS kernel is applied. Planar mode is mapped for the block. In this contribution, two kinds of methods are proposed. Method 1 is modifications of MTS and LFNST for IntraTMP coded blocks. It is proposed to use DIMD to derive an intra prediction mode based on the neighouring samples of the IntraTMP block, then the MTS transform set or the LFNST transform set is respectively determined by the derived intra mode. Method 2 is intra prediction mode derivation using DIMD for all IntraTMP coded blocks. The derived intra prediction mode is stored and used for generating MPM list of the neighbor intra prediction blocks. The proposed method is implemented on top of EE2-TEST-1.14: IntraTMP Adaption for Camera Captured Contents.

On top of EE2-TEST-1.14, simulation results of the proposed method are reported as below:

Method 1:

AI: {-0.05%, -0.03%, -0.05%, 100%, 101%}

RA: {0.xx%, 0.xx%, 0.xx%, xxx%, xxx%}

Method1 + Method 2:

AI: {0.xx%, 0.xx%, 0.xx%, xxx%, xxx%}

RA: {0.xx%, 0.xx%, 0.xx%, xxx%, xxx%}

Concept is similar to EE2-4.1, where DIMD is used to determine the LFNST kernel; in this proposal, a similar concept is applied to IntraTMP blocks, results are reported to give additional gain on top of the new method of MV derivation (EE2-1.14). It is reported to provide some additional gain with almost no enc/dec run time increase.

Investigate in EE.

[JVET-AB0128](https://jvet-experts.org/doc_end_user/current_document.php?id=12055) EE2-related: CCCM template selection [P. Bordes, K. Naser, E. François, F. Galpin (InterDigital)]

In EE2 test 1.13 described in document JVET-AA2024, three types of templates for CCLM and CCCM to derive the CC-models are defined: full template, top-only template, and left-only template. It is proposed to extend the template selection to six.

It is reported that the results of the proposed method implemented on top of EE2-test-1.13b in AI are - 0.06% (Y), -1.40% (U), -1.39% (V) compared to ECM-6.0 anchor, and -0.02% (Y), -0.28% (U), -0.28% (V), 100% (EncT), 99% (DecT) compared to EE2-test-1.13b.

Not enough benefit to be considered.

[JVET-AB0220](https://jvet-experts.org/doc_end_user/current_document.php?id=12147) Crosscheck of JVET-AB0128 (EE2-related: CCCM template selection) [Y.-J. Chang (Qualcomm)] [late]

[JVET-AB0138](https://jvet-experts.org/doc_end_user/current_document.php?id=12065) EE2-related: MRL candidate list reordering [Y. Lee, B. Kim, B. Jeon (SKKU)]

This contribution proposes a reordering scheme of the MRL candidate list. The MRL candidate list is reordered according to the SAD between the predictor generated with adjoined reference line (i.e., *intra\_luma\_ref\_idx*=0) and the predictor generated with another reference line in MRL candidate list (i.e., *intra\_luma\_ref\_idx*0).

Proposed method achieves BDBR gain compared to ECM-6.0 as follows:

AI - -0.01% (Y) / 0.04% (Cb) / -0.03% (Cr), Enc: 106%, Dec: 101%.

Additional SAD computations (between predictors corresponding to reference line 0 and other reference lines) are necessary at the decoder. Benefit in compression is too small to justify to consider this proposal.

[JVET-AB0139](https://jvet-experts.org/doc_end_user/current_document.php?id=12066) EE2-related: On Chroma Fusion improvement [C. Zhou, Z. Lv, J. Zhang (vivo)]

In this contribution, a new chroma fusion method is proposed where the weights are derived by the adjacent template of both the reconstructed luma samples and the predicted chroma samples obtained by applying the non-LM mode. The derivation is based on the LDL decomposition method used in CCCM. The experimental results are summarized as follows:

On top of ECM-6.0

AI: -0.01% (Y), -0.40% (U), -0.34% (V), 102% (EncT), 101% (DecT)

AI-TGM: -0.16% (Y), -0.50% (U), -0.54% (V), 102% (EncT), 100% (DecT)

Question : Was DIMD chroma enabled ? Yes.

An additional block-level signaling is introduced for the switching between the fusion methods.

Several experts expressed interest.

Investigate in EE.

[JVET-AB0230](https://jvet-experts.org/doc_end_user/current_document.php?id=12159) Crosscheck of JVET-AB0139 (EE2-related: On Chroma Fusion improvement) [J. Chen (Alibaba)] [late]

[JVET-AB0144](https://jvet-experts.org/doc_end_user/current_document.php?id=12071) EE2 related: Extension of test EE2-3.3 [F. Le Léannec, P. Andrivon, M. Radosavljević, M. Blestel (Xiaomi)]

for chroma component, and the re-use of block vectors derived during ITMP predictor as candidate predictors for the prediction of block vectors of IBC coding units.

First aspect consists in storing ITMP-derived block vectors into the IBC HMVP table, that is the table used to history-based block vector prediction of IBC.

Second aspect increases the number of IBC merge candidates derived before reordering during the IBC candidate list construction for IBC merge and AMVP.

The proposed extension of test EE2-3.3 leads to the following gain over ECM-6.0:

* AI class TGM: -0.46% / -0.42% / -0.52% (Y / Cb / Cr)
* AI class F: -0.11% / -0.15% / 0.06%
* RA class TGM: -0.20% / -0.22% / -0.24%
* RA class F: -0.08% / -0.10% / -0.01%

The encoder and decoder runtimes are reportedly unchanged compared to ECM-6.0 anchor.

No need for presentation, as it is extending EE2-3.3 which was not considered for adoption.

[JVET-AB0237](https://jvet-experts.org/doc_end_user/current_document.php?id=12166) Crosscheck of JVET-AB0144 (EE2 related: Extension of test EE2-3.3)) [H. Wang (Qualcomm)] [late] [miss]

[JVET-AB0145](https://jvet-experts.org/doc_end_user/current_document.php?id=12072) EE2-2.6-related: On Decoder-side Affine Model Refinement (DAMR) [J. Chen, R.-L. Liao, X. Li, Y. Ye (Alibaba)]

This contribution proposes to refine the affine model for affine merge coded blocks in decoder sider without additional signaling. In the proposed method, both the base MV and non-translation parameters of affine model are refined to improve the accuracy of the affine model inherited from the previously coded blocks in merge mode. It is reported that on top of ECM-6.0, by applying the proposed method, it achieves {-0.26%(Y), -0.16%(U), -0.19%(V)} coding gain in RA on average with luma coding gain reaching -0.36% on class A1 and -0.50% on classA2.

It was pointed out that the refinement may be simpler for a 4-parameter model.

Interest was expressed to investigate in EE, along with JVET-AB0178. It was pointed out that encoding/decoding run times are not giving a good tradeoff with the compression benefit in the current version, and would need to be significantly reduced. The proposal for EE2-2.6 that was adopted had a much better tradeoff.

[JVET-AB0151](https://jvet-experts.org/doc_end_user/current_document.php?id=12078) EE2-2.1 related: ARMC merge candidate list reordering for AMVP-merge mode for low-delay pictures [K. Cui, C. S. Coban, Z. Zhang, V. Seregin, H. Huang, M. Karczewicz (Qualcomm), H. Jang (LGE)]

This contribution proposes to use ARMC for merge candidate list reordering of AMVP-merge mode for low-delay pictures. In this mode, a merge candidate with the smallest cost after reordering is paired with the first AMVP candidate to derive the prediction. The proposed method was implemented on top of ECM-6.0. It reports -0.26% (Y), -0.23% (U), -0.14% (V) BD-rate reduction for low-delay-B configuration with 104% encoding and 99% decoding run time compared to ECM-6.0.

The decision of using one or the other AMVP merge mode is determined at slice level. If all pictures in the RPL for the current slice are from the past, then the modified method is used.

Investigate in EE.

[JVET-AB0249](https://jvet-experts.org/doc_end_user/current_document.php?id=12178) Cross-check of JVET-AB0151 "EE2-2.1 related: ARMC merge candidate list reordering for AMVP-merge mode for low-delay pictures" [F. Le Léannec (Xiaomi)] [late]

[JVET-AB0161](https://jvet-experts.org/doc_end_user/current_document.php?id=12088) EE2-1.16 related: Encoder optimization for picture-level geometry transform [W. Jia, K. Zhang, Y. Wang, T. Fu, Y. Li, L. Zhang (Bytedance)]

This proposal proposes an optimized encoder for picture-level geometry transform proposed in EE2-1.16, specifically for RA testing conditions. On top of ECM-6.0, simulation results of the proposed optimized method are reported as below:

RA: {-0.09%, -0.21%, -0.28%, 100%, 100%}.

No need for presentation, the corresponding EE2-1.16 was not adopted.

[JVET-AB0162](https://jvet-experts.org/doc_end_user/current_document.php?id=12089) EE2-related: On horizontal and vertical planar modes [X. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba)]

In this contribution, two modifications to the horizontal and vertical planar modes are proposed. First, a gradient based decoder side derivation method is proposed to decide which of the two new planar modes is used. Second, the propagated modes of the horizontal planar mode and vertical planar mode are modified into vertical mode and horizontal mode respectively when deriving transform kernel. It is reported that on top of ECM-6.0, the overall coding performance impact for {Y, U, V, EncT, DecT} are {x.xx%, x.xx%, x.xx%, xxx%, xxx%} for AI and {x.xx%, x.xx%, x.xx%, xxx%, xxx%} for RA.

For AI, performance is expected to be 0.06% bit rate reduction with 8% encoding time increase.

Investigate in EE, along with JVET-AB0110 and JVET-AB0104, but it would be expected to provide a significantly better tradeoff, i.e. further reducing encoder runtime while retaining or improving the compression.

[JVET-AB0246](https://jvet-experts.org/doc_end_user/current_document.php?id=12175) Crosscheck of JVET-AB0162 (EE2-ralated: On horizontal and vertical planar modes) [Z. Fan, Y. Yasugi, T. Ikai (Sharp)] [late]

[JVET-AB0177](https://jvet-experts.org/doc_end_user/current_document.php?id=12104) EE2-related: Sub-block processing for affine DMVR [H. Huang, Y. Zhang, Z. Zhang, C.-C. Chen, V. Seregin, M. Karczewicz (Qualcomm), J. Chen, R.-L. Liao, X. Li, Y. Ye (Alibaba)]

This contribution proposes sub-block processing for affine DMVR, where bilateral matching is performed per subblock. The refined subblock MVs and accumulated subblock bilateral matching costs are used to derive the refined control point motion vectors. The simulation result reports -0.19%, -0.14%, -0.16% Y, U, V BD-rate reduction with 100% encoder time and 101% decoder time.

Method without regression would give 0.15% bitrate reduction.

How much gain would be retained when used on top of EE2-2.6? Likely not much, if at all. The method without regression has same bit rate reduction as EE2-2.6, but it is claimed to have less complexity (not necessarily in run time).

Not enough understanding about the potential benefit over the EE2-2.6 proposal, such that JVET would not revert the decision made on that proposal.

It is not known whether the proposal gives additional benefit on top of EE2-2.6, but that could be investigated in EE, and the proponents are willing to participate. The purpose of that EE would not be to replace EE2-2.6

[JVET-AB0225](https://jvet-experts.org/doc_end_user/current_document.php?id=12153) Cross-check of JVET-AB0177 on Sub-block processing for affine DMVR [X. Li (Google)] [late]

[JVET-AB0231](https://jvet-experts.org/doc_end_user/current_document.php?id=12160) Cross-check of JVET-AB0177 on Sub-block processing for affine DMVR [F. Galpin (InterDigital)] [late]

[JVET-AB0251](https://jvet-experts.org/doc_end_user/current_document.php?id=12180) Crosscheck of JVET-AB0177 (EE2-related: Sub-block processing for affine DMVR) [F. Wang, Y. Yu (OPPO)] [late]

[JVET-AB0178](https://jvet-experts.org/doc_end_user/current_document.php?id=12105) EE2-related: Control-point motion vector refinement for affine DMVR [H. Huang, C.-C. Chen, Y. Zhang, Z. Zhang, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes control point motion vector refinement for affine DMVR. The control-point motion vector is independently refined by bilateral matching of a block centered by the control-point, then it is iteratively refined to minimize the bilateral matching cost of the coding block. The proposed method is also extended to be applied in affine MMVD mode. The simulation result reports -0.39%, -0.23%, -0.29% Y, U, V, BD-rate reduction.

A significant amount of bilateral matching steps is applied.

Interest was expressed to investigate in EE along with JVET-AB0145. It was pointed out that encoding/decoding run times are much too high to give a reasonable tradeoff with the compression benefit in the current version, these would need to be significantly reduced. Report results with and without using the method in MMVD mode

[JVET-AB0186](https://jvet-experts.org/doc_end_user/current_document.php?id=12113) EE2-related: Modification of extended offline-filter taps for ALF [I. Jumakulyyev, N. Hu, V. Seregin, M. Karczewicz (Qualcomm)]

In ECM-6.0, an online-trained adaptive loop filter contains 2 offline-filtered taps. In EE2-5.1, the number of offline-filtered taps is extended to 8. In this contribution, this number is further extended to 12.

On top of ECM-6.0, simulation results are reported as below:

AI: -0.11%, 0.02%, 0.02%, 104%, 104%,

RA: -0.18%, 0.02%, -0.02%, 101%, 102%.

On top of EE-2.5.1, simulation results are reported as below:

AI: -0.02%, 0.01%, 0.01%, 102%, 102%,

RA: -0.04%, 0.01%, 0.01%, 100%, 101%.

In terms of encoder run time for AI, the tradeoff with compression benefit is not as good. Also according to crosscheckers, the runtime increase was found.

No action on this proposal.

[JVET-AB0229](https://jvet-experts.org/doc_end_user/current_document.php?id=12158) Cross-check of JVET-AB0186 (EE2-related: Modification of extended offline-filter taps for ALF) [K. Andersson (Ericsson)] [late]

[JVET-AB0257](https://jvet-experts.org/doc_end_user/current_document.php?id=12186) EE2 related: Improved directional planar prediction [S. Yoo, J. Choi, J. Nam, M. Hong, J. Lim, S. Kim (LGE), K. Kim, D. Kim, J. Son, J. Kwak (WILUS), X. Li, R. Liao, J. Chen, Y. Ye (Alibaba)] [late]

In this proposal, improved horizontal and vertical planar prediction method is suggested. Firstly, a derivation method of a transform kernel for planar horizontal and planar vertical mode is proposed. If an intra prediction mode of a current block is the planar horizontal/vertical mode, the vertical/horizontal intra prediction mode is used to derive a transform kernel in MTS set and LFNST set. Further, the directional flag is inferred either horizontal planar or vertical planar mode based on the DIMD mode. If the DIMD first mode is less than 34, the intra mode of the current block is inferred to be a horizontal planar mode. Otherwise, the intra mode of the current block is set to be a vertical planar mode. The experimental results of the proposed methods are shown below:

* + AI: X.XX% / X.XX% / X.XX% for Y/U/V, XXX% / XXX% for EncT, DecT
  + RA: X.XX% / X.XX% / X.XX% for Y/U/V, XXX% / XXX% for EncT, DecT

Investigate in EE, along with JVET-AB0110, JVET-AB0162, and JVET-AB0104, but it would be expected to provide a significantly better tradeoff, i.e. further reducing encoder runtime while retaining or improving the compression.

[JVET-AB0262](https://jvet-experts.org/doc_end_user/current_document.php?id=12191) Crosscheck of JVET-AB0257 (EE2-related: Improved directional planar prediction) [D. Kim, S.-C. Lim (ETRI)] [late]

### ECM modifications beyond EE2 (28)

Contributions in this area were discussed at 1700–2010 on Saturday 22 October 2022 (chaired by JRO), and 1600–2000 on Tuesday 25 October 2022 (chaired by JRO until 1800, Y. Ye afterwards, contributions JVET-AB0181 … JVET-AB0191).

[JVET-AB0065](https://jvet-experts.org/doc_end_user/current_document.php?id=11980) Non-EE2: Adaptive reference region DIMD [Z. Fan, Y. Yasugi, T. Ikai (Sharp)]

This contribution proposes to introduce new mode for DIMD, named adaptive reference region DIMD. The proposed DIMD consists of DIMD\_TL (top and left), DIMD\_L (left) and DIMD\_T (top) where a new syntax element is signalled to select neighboring reconstructed pixels (reference region) used to derive a prediction direction (DIMD intra mode). The experimental results reportedly show that the average luma BDrate gain of the proposal is 0.1 % and 0.06 %, under AI and RA, respectively.

Signaling and switching between the different DIMD modes is performed at block level.

Encoder runtime increased by 15% in AI, due to additional RD decisions.

The different DIMD modes have different number of samples.

Several experts expressed interest to investigate in EE. One primary goal should be to achieve a muchg better tradeoff between encoder runtime and coding gain.

[JVET-AB0201](https://jvet-experts.org/doc_end_user/current_document.php?id=12128) Crosscheck of JVET-AB0065 (Non-EE2: Adaptive reference region DIMD) [X. Li (Alibaba)] [late]

[JVET-AB0082](https://jvet-experts.org/doc_end_user/current_document.php?id=12009) AHG12: Fixes for RPR [K. Andersson, R. Yu (Ericsson)]

RPR can be used to make use of a reference pictures with different resolution than the current picture. It was discovered that RPR CTC could not be run without encoder and decoder crash. It was also noticed that the 12-tap ECM luma filter was not used for upscaling luma before outputting pictures when upscaling to full resolution (UpscaledOutput=2) and not either for upscaling luma pictures for ME on encoder side.

The encoder/decoder issue was fixed by increasing the code size from 9 to 10 bit for encoding lambda in ARMC.

The fix of the luma filter length from 8-tap to 12-tap gives compared to ECM-6.0 including the first fix for RPR CTC LDB (with switch of resolution every 0.5s):

2x: Y/U/V: -1,70%/0,04%/-0,12%

1.5x: Y/U/V: -2,03%/-0,16%/-0,05%

It is suggested to consider the two fixes for next version of ECM.

It was reported that potentially further problems in interaction between RPR and other parts of ECM implementation.

It was confirmed by original proponents of ARMC that the change of lambda encoding is necessary (this would need to be asserted as a normative change, but is fixing a bug for the non-CTC case of low QP). Its interaction with the software crash with RPR seems to be somewhat random, as it should not cause problems in CTC.

The modification of filter length is intended for improving the performance for out-of-loop post processing in case of low-resolution coding and upscaling. The RPR filters (used in-loop) were already changed to 12 tap in ECM (6-tap for chroma), whereas the upscaling filters for luma were still 8-tap.

Decision(BF): Increase the syntax element for signaling lambda in ARMC from 9 to 10 bits.

Decision(SW): Modify upscaling filters (post processing for upscaling into original resolution) such that they are identical with motion compensation filters.

[JVET-AB0252](https://jvet-experts.org/doc_end_user/current_document.php?id=12181) Crosscheck of JVET-AB0082 (AHG12: Fixes for RPR) [C. S. Coban (Qualcomm)] [late] [miss]

[JVET-AB0094](https://jvet-experts.org/doc_end_user/current_document.php?id=12021) Non-EE2: Direct block vector (DBV) mode for chroma prediction [J.-Y. Huo, X. Hao, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.), J. Ren, M. Li (OPPO)]

In this contribution, a new chroma intra prediction mode, direct block vector (DBV) mode, is proposed for screen content. This mode is used for chroma components when dual tree is activated in intra slice. After sampling format scaling, the BV at the co-located luma position is used to apply block prediction for chroma. It is reported that the coding performance of ECM 6.0 is as below:

For AI configuration:

F:-0.27%, -0.38%, and -0.31%, with xxx % EncT, 99% DecT.

TGM:-1.81%, -1.93%, and -2.48%, with xxx % EncT, 99% DecT.

For RA configuration:

F: -0.24%, -0.10%, and -0.26 %, with xxx % EncT, xxx% DecT.

TGM: -0.58%, -0.57%, and -0.74%, with xxx% EncT, xxx% DecT.

Effect on camera-captured content is very small, as it is only effective when IBC is used for luma.

No encoder run time results available. It is asserted to be small, as only one additional mode is checked for chroma (cross checker estimates this to be 1-2%, but results not finished yet).

Several experts expressed interest to investigate in EE. The EE should also compare against the method of VVC/VTM, where just the co-located luma IBC vector is used for chroma as well (without additional signaling). G. Li (Tencent) volunteers performing this test.

[JVET-AB0260](https://jvet-experts.org/doc_end_user/current_document.php?id=12189) Crosscheck of JVET-AB0094 (Non-EE2: Direct block vector (DBV) mode for chroma prediction) [X. Li (Alibaba)] [late]

[JVET-AB0095](https://jvet-experts.org/doc_end_user/current_document.php?id=12022) Non-EE2: Block Vector Difference Sign Prediction (BVDSP) for IBC blocks [J.-Y. Huo, X. Hao, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.), J. Ren, M. Li (OPPO)]

In ECM6.0, signs of block vector difference (BVD) in IBC blocks are transmitted with equal probability (EP) and signs of motion vector difference (MVD) in INTER blocks are coded by MVD sign prediction. In this contribution, the block vector difference sign prediction (BVDSP) is proposed to improve coding efficiency. It is reported that the coding performance to ECM 6.0 is as below:

For AI configuration:

F: -0.21%, -0.25%, and -0.04%, with xxx % EncT, 100 % DecT.

TGM: -0.43%, -0.40%, and -0.43%, with xxx % EncT, 101 % DecT.

For RA configuration:

F: -0.06 %, -0.04%, and -0.11 %, with xxx % EncT, xxx% DecT.

TGM: -0.27 %, -0.36 %, and -0.28 %, with xxx% EncT, xxx% DecT.

The method is different (but conceptually similar) to the method of sign prediction used in MV coding. Proponents claim it is simpler. Also the coding of sign difference is different.

Several experts expressed interest to investigate in EE.

[JVET-AB0261](https://jvet-experts.org/doc_end_user/current_document.php?id=12190) Crosscheck of JVET-AB0095 (Non-EE2: Block Vector Difference Sign Prediction (BVDSP) for IBC blocks) [X. Li (Alibaba)] [late]

[JVET-AB0099](https://jvet-experts.org/doc_end_user/current_document.php?id=12026) Non-EE2: CCCM with Multi-shape filters [C. Fang, S. Peng, D. Jiang, J. Lin, X. Zhang (Dahua)]

This contribution proposes an improved CCCM method with multi-shape filters. In this method, five additional filters with different shapes are provided. The proposed method was implemented on top of ECM-6.0, it reportedly provides {Y, U, V} BD-rate reduction with encoder and decoder runtimes as follows:

AI: -0.01%, -0.85%, -0.83%, 110% (EncT), 98% (DecT)

In terms of tradeoff encoder runtime vs. compression benefit, this does not show advantage over the previous proposals from EE2-1.7.

No action.

[JVET-AB0100](https://jvet-experts.org/doc_end_user/current_document.php?id=12027) Non-EE2: Separable KLT for intra coding [M. Koo, J. Zhao, J. Lim, S. Kim (LGE)]

In this contribution, KLT kernels of length 8, 16, and 32 are proposed, which can be applied as separable primary transforms on top of adaptive MTS. These KLTs are commonly applied to every block shape of which width and height ranges from 4 to 32 for only intra-coded CU. And, since they can also be applied to only one direction with DCT-2 coupled in the other, up to three pairs can be candidates as follows: (KLT, KLT), (KLT, DCT-2), and (DCT-2, KLT). These pairs are activated by setting a newly employed flag and also an additional index is followed to indicate one of the pairs. The experimental results for the proposed KLTs are reportedly observed that the BD-rates are reduced by 0.10% with 113%/100% encoding/decoding time changes for AI.

Explicit signalling, one more primary transform – likely reason for the increased encoder run time.

Higher gain for higher resolutions – how was the training set compiled? It was a mixture of various resolutions.

Proponents believe they could reduce the encoding time.

Transforms are not mode dependent, just different kernels for the different block sizes. All transforms are integer.

With the current version, the tradeoff encoder complexity vs. compression gain is not attractive.

Questions were asked how much benefit comes by the combination with length-4 DCT.

Further study recommended, better tradeoff should be demonstrated.

[JVET-AB0111](https://jvet-experts.org/doc_end_user/current_document.php?id=12038) Non-EE2: Intra prediction fusion with PMPM list [G. Moon, D. Park, Y.-U. Yoon, J.-G. Kim (KAU)]

This contribution proposes a intra prediction fusion method that blends the prediction of a primary MPM mode and the prediction of its proceeding mode for all intra modes in the primary MPM list except the planar mode. The proposed method was implemented on top of ECM-6.0 software, and the experimental results are summarized as follows:

* AI: x.x%, x.x%, x.x%, x% (EncT), x% (DecT)
* RA: x.x%, x.x%, x.x%, x% (EncT), x% (DecT)

Preliminary results indicate 0.12% gain with 15% encoder run time increase. Currently, a full RD comparison of different modes (see slide 5 of PPT presentation version 1) is conducted.

Further study recommended to reduce the encoder run time without losing much of the coding gain.

[JVET-AB0116](https://jvet-experts.org/doc_end_user/current_document.php?id=12043) AHG12 - Location-dependent Decoder-side Intra Mode Derivation [S. Blasi, J. Lainema (Nokia)]

This contribution proposes to use sample-based weights when blending the intra predictors determined using Decoder-side Intra Mode Derivation (DIMD). The type of blending and the weights to use for a given predictor are inferred during the DIMD derivation process, by analysing the directionality of the samples in the templates above and on the left.

It is reported that adding the proposed location-dependent DIMD to ECM provides -0.08%, -0.10% and -0.11% BD-rate impact for Y, U and V, respectively. In RA configuration the impact is reportedly -0.04%, 0.07% and -0.09% for Y, U and V, respectively. The encoder and decoder runtimes are reportedly unchanged with respect to ECM.

Histograms from templates above and left are used to determine the weights. Integer approximations of weight are used.

Encoder/decoder run times are changing only by minimum (confirmed by cross-check).

It was suggested to study the combination with JVET-AB0065.

Study in EE.

[JVET-AB0255](https://jvet-experts.org/doc_end_user/current_document.php?id=12184) Cross-check of JVET-AB0116: “AHG12 - Location-dependent Decoder-side Intra Mode Derivation [T. Dumas (InterDigital)] [late]

[JVET-AB0117](https://jvet-experts.org/doc_end_user/current_document.php?id=12044) AHG12 - Template-based Intra Mode Derivation with Directional blending [S. Blasi, J. Lainema, I. Zupancic, D. Bugdayci Sansli (Nokia)]

This contribution proposes to use directional blending to fuse predictors determined using Template-based Intra Mode Derivation (TIMD). When using directional TIMD, two separate TIMD modes are derived using the template above and the template on the left, respectively. The two modes are blended using sample-based weights.

It is reported that adding directional TIMD to ECM provides -0.19%, -0.08% and 0.09% BD-rate impact in AI configuration for Y, U and V, respectively. The encoder runtime is reported to increase by 8% on average and the decoder runtime is reported to increase by 2% on average with respect to ECM. In RA configuration the impact is reportedly -0.10%, 0.01% and -0.11% for Y, U and V, respectively, with no change in encoder and decoder runtime with respect to ECM.

Study in EE, also in combination with JVET-AB0116 and JVET-AB0065.

[JVET-AB0119](https://jvet-experts.org/doc_end_user/current_document.php?id=12046) Non-EE2: Gradient and location based convolutional cross-component model (GL-CCCM) for intra prediction [R. G. Youvalari, P. Astola, J. Lainema (Nokia)]

This contribution proposes a gradient and location based convolutional cross-component model (GL-CCCM) for chroma prediction to improve compression efficiency of ECM. The proposed 7-tap convolutional filter maps luma values into chroma values when a GL-CCCM prediction mode is activated by a PU level flag. The filter input consists of one spatial luma sample, two gradient values, two location information, a nonlinear term, and a bias term. Filter coefficients are derived for each chroma block separately using regression based MSE minimization (i.e., the same solver as CCCM) on reference samples in the PU’s neighborhood. The impact on coding efficiency and runtimes over ECM-6.0 is reportedly {for Y, U, V, EncT, DecT}:

AI { -0.06%, -0.77%, -0.80%, 102%, 100%}, RA { -0.03%, -0.37%, -0.47%, 99%, 100%}

Furthermore, the tool has been extended to support different templates as in EE2-1.13 using full template, top-only template, and left-only template. The results on top of EE software show further coding performance increase of the EE proposals.

Results on top of EE2-1.13. A:

AI { -0.08%, -1.53%, -1.58%, 104%, 98%}, RA { -0.04%, -0.94%, -1.20%, 101%, 96%}

Results on top of EE2-1.13. B:

AI { -0.09%, -1.73%, -1.69%, 102%, 100%}, RA { -0.xx%, -0.xx%, -0.xx%, 10x%, 10x%}

Investigate in EE. It should be also investigated how much gain comes from gradient alone, and how much gain comes by additionally using location information.

[JVET-AB0133](https://jvet-experts.org/doc_end_user/current_document.php?id=12060) AHG12: Inter-RPL and 1-byte NAL unit headers [R. Sjöberg, M. Pettersson, J. Ström (Ericsson)]

This contribution proposes two modifications to ECM-6.0.

The first proposed modification is to add prediction to the signaling of the RPLs in the SPS. The method utilizes that the L0 and L1 picture references of an entry in the SPS list of RPLs are limited to the L0 and L1 picture references of the previous entry plus a reference to the previous picture.

The second proposed modification is to make the NAL unit header in ECM 1 byte rather than 2 bytes for the 16 most common NAL unit types when the layer\_id value is equal to 0.

For both modifications, the proponents report a CTC average luma BD-rate of −0.01/−0.11%/−0.13% for AI/RA/LDB.

The slice data and therefore also all decoded sample values is asserted to be unaffected by the proposal. The proponents claim that the decoding run-time is reduced by the proposal but that the difference is very small.

The proponents report that for RA CTC, the method gives −0.63% on Class D and −0.37% on Class F (not part of the CTC average).

With only the NAL unit header modification, the proponents report a CTC average luma BD-rate of −0.01/−0.03%/−0.07% for AI/RA/LDB. The proponents claim that the savings are expected to scale with the number of slices used per picture.

Saving bits in high-level syntax is not of primary importance for this exploration – this would rather be interesting when developing a standard.

No action.

Cross-check JVET-AB0273 (add header)

[JVET-AB0273](https://jvet-experts.org/doc_end_user/current_document.php?id=12202) Cross-check for JVET-AB0133 [K. Sühring (HHI)] [late]

[JVET-AB0142](https://jvet-experts.org/doc_end_user/current_document.php?id=12069) Non-EE2: Optimizing the use of available decoded reference samples [T. Dumas, K. Reuzé, Y. Chen, K. Naser (InterDigital)]

This contribution proposes to optimize the use of the available decoded reference samples around the current CU during the intra prediction step. This encompasses three changes: the order of a few MPMs of the current luma CB now depends on the availability of the above-right and bottom-left decoded reference samples, the DIMD region of decoded reference samples of the current luma CB is extended towards above-right and bottom-left if available, and a few candidate wide-angle intra prediction modes are added during the TIMD derivation step if they intensively use available above-right/bottom-left decoded reference samples. The experimental results reportedly show that the average BD-rate gains of the proposal are -0.05%, -0.04%, -0.08% and xxx%, -xxx%, -xxx% under AI and RA, respectively.

No change in encoder and decoder run time.

Study in EE, also in combination with JVET-AB0116, JVET-AB0117 and JVET-AB0065

[JVET-AB0233](https://jvet-experts.org/doc_end_user/current_document.php?id=12162) Cross-check of JVET-AB0142 (Non-EE2: optimizing the use of available decoded reference samples) [M. Abdoli (IRT b-com)] [late]

[JVET-AB0166](https://jvet-experts.org/doc_end_user/current_document.php?id=12093) Non-EE2: Unified pruning of affine merge candidates derivation [Z. Deng, K. Zhang, L. Zhang (Bytedance)]

This contribution proposes to unify pruning for affine merge candidates derivation. In ECM-6.0, two different pruning logics (denoted as A and B) are used for affine merge candidates derivation, which is asserted to complicate the affine prediction design. Two solutions are tested in this proposal. In solution #1, pruning in affine merge are unified to pruning logic A. In solution #2, pruning in affine merge are unified to logic B. Simulation results based on ECM-6.0 are reported as below:

Solution #1:

RA: 0.00%, 0.05%, -0.01%, 100%, 100%

LB: -0.02%, 0.10%, 0.21%, 100%, 100%

Solution #2:

RA: -0.01%, 0.01%, -0.03%, 100%, 100%

LB: 0.04%, 0.11%, 0.42%, 100%, 100%

This proposal suggests a cleanup of the design. This comes with a very small loss in performance, but changes like this are not of primary importance at this stage.

[JVET-AB0265](https://jvet-experts.org/doc_end_user/current_document.php?id=12194) Crosscheck of JVET-AB0166 (Non-EE2: Unified pruning of affine merge candidates derivation) [W. Chen (kwai)] [late] [miss]

[JVET-AB0168](https://jvet-experts.org/doc_end_user/current_document.php?id=12095) Non-EE2: Pixel based affine motion compensation [Z. Zhang, H. Huang, Y. Zhang, P. Garus, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes to set the minimum affine subblock size to be 1×1 for luma channel of an affine coded block. It proposes to adaptively apply pixel based affine motion compensation when OBMC is not applied. It also proposes to only signal OBMC flag when the current block is an affine AMVP coded block without Multi-Hypothesis Predictor.

The proposed method was implemented on top of ECM-6.0 and the test results are as follows:

RA: -0.11% (Y), -0.02% (U), -0.05 % (V), 102% (EncT), 100% (DecT).

LDB: -0.02% (Y), 0.01% (U), 0.14% (V), 100% (EncT), 100% (DecT)

Investigate in EE. It was suggested to also investigate always using 1x1, and avoiding to switch between 1x1 and 4x4 (this might however lead to a significant increase in encoding time). It was asked if this would still work with SIMD? Should be investigated.

[JVET-AB0170](https://jvet-experts.org/doc_end_user/current_document.php?id=12097) Non-EE2: Block Vector Difference Prediction for IBC blocks [A. Filippov, V. Rufitskiy (Ofinno)]

This contribution presents a technique to predict signs and magnitudes of x- and y-components of block vector differences (BVDs) for IBC blocks. In particular, BVD signs and suffix bins of exponential Golomb code used to represent BVD magnitudes are predicted by estimating template matching cost of candidate blocks to take advantage of regular CABAC mode rather than its by-pass mode at the entropy coding stage.

On top of ECM-6.0, the following maximum BD-rate gain results are reportedly obtained:

* AI configuration: {-0.50%, -0.62%, -0.32%} for class F and {-1.52%, -1.32%, -1.38} for class TGM.

On top of ECM-6.0, the following BD-rate gain results are reportedly obtained for the case when up to 4 bins of BVD suffix magnitudes are predicted:

* AI configuration: {-0.31%, -0.44%, -0.15%, 99%, 101%} for class F and {-0.85%, -0.82%, -0.86%, 100%, 100%} for class TGM;
* RA configuration: {-0.22%, -0.13%, 0.05%, 101%, 100%} for class F and {-0.63%, -0.64%, -0.62%, 100%, 101%} for class TGM;

Investigate in EE, also in combination with JVET-AB0095.

[JVET-AB0173](https://jvet-experts.org/doc_end_user/current_document.php?id=12100) AHG12: BVP candidates clustering and BVD sign derivation for Reconstruction-Reordered IBC mode [D. Ruiz Coll, V. Warudkar, J.-K. Lee (Ofinno)]

This contribution proposes modifying the IBC AMVP list construction based on clustering the block vector predictors according to the L2 distance between the candidates and selecting one candidate per cluster based on the TM cost. In addition, if the Reconstruction-Reordered IBC (RRIBC) mode is selected for IBC AMVP, two new BVP candidates are determined, and the sign of the non-null BVD component is inferred at the decoder side. On top of ECM-6.0, the test results for the proposed method for class F and class TGM are reported as follows:

* AI: -0.26%, -0.22%, and -0.21%, with 99% EncT, 95% DecT.
* RA: -0.22%, -0.26%, and -0.18%, with 100% EncT, 100% DecT.

On top of the test reported in EE2 3.3, the following gain is obtained over ECM-6.0:

* AI Class F: -0.23%, -0.36%, and -0.12%, with 96% EncT, 97% DecT.
* AI Class TGM: 0.62%, -0.59%, and -0.70%, with 93% EncT, 87% DecT.

Investigate in EE, also in combination with JVET-AB0095 and JVET-AB0170.

[JVET-AB0264](https://jvet-experts.org/doc_end_user/current_document.php?id=12193) Crosscheck of JVET-AB0173: AHG12: BVP candidates clustering and BVD sign derivation for Reconstruction-Reordered IBC mode [K. Naser (InterDigital)] [late] [miss]

[JVET-AB0174](https://jvet-experts.org/doc_end_user/current_document.php?id=12101) AHG12: Division-free operation and dynamic range reduction for convolutional cross-component model (CCCM) [A. Aminlou, J. Lainema, R. G. Youvalari, P. Astola (Nokia)]

This contribution proposes two modifications to convolutional cross-component model (CCCM) based chroma prediction with an intention to simplify the implementation. The first modification removes an offset from reference luma and chroma samples. Reference sample values at the top-left corner of the PU are used as the offsets for simplicity. Using such offset is asserted to reduce the dynamic range of data and the bitdepth required for model derivation. The second modification replaces the division operation in CCCM by a piece-wise polynomial (power of 2) function.

The impact on coding efficiency and runtimes over ECM-6.0 is reportedly {for Y, U, V, EncT, DecT}:

AI {0.01%, 0.04%, 0.05%, 101%, 101%} RA { -0.01%, 0.03%, 0.13%, 102%, 100%}

Only replacing the division would lead to 0.19% bitrate increase. The change using the offset presents that. Has it been observed that the offset might need to consider clipping to stay within 16bit range, e.g. if the top-left sample would be zero? It was argued that if this would happen in extreme cases, the prediction would be wrong and the CCCM mode would not be selected.

Was confirmed by crosscheck.

Decision: Adopt JVET-AB0174 to ECM7.

[JVET-AB0219](https://jvet-experts.org/doc_end_user/current_document.php?id=12146) Crosscheck of JVET-AB0174 (AHG12: Division-free operation and dynamic range reduction for convolutional cross-component model (CCCM)) [Y.-J. Chang (Qualcomm)] [late]

[JVET-AB0175](https://jvet-experts.org/doc_end_user/current_document.php?id=12102) Non-EE2: Non-Separable Primary Transform for Intra Coding [P. Garus, M. Coban, B. Ray, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes to utilize Non-Separable Primary Transforms (NSPTs) for certain block sizes instead of separable DCT-II + LFNST. In the proposed design, DCT-II+LFNST is replaced by NSPT for TB sizes 4x4, 4x8, 8x4 and 8x8. The NSPT follow the design of LFNST for each block size, *i.e.* 35 transform sets depending on the intra mode and 3 candidates.

The proposed method was implemented on top of ECM-6.0 and provide these results following the common test conditions [2]:

AI: -0.19% (Y), -0.16% (U), -0.17 % (V), 100% (EncT), 99% (DecT).

For 4x8, 8x4, and 8x8, coefficients are zeroed out.

Gain is higher for low resolutions, and almost no gain for classes A1 and A2.

How was training performed? Same set as used for LFNST extended, plus additional sequences outside of CTC, more low-resolution content. CTC QPs were used. It was asked if using the extended set for re-training LFNST might improve LFNST performance?

The proponents believe that the smaller block sizes are more frequently used in the low resolutions, which is the reason for the better performance there.

Memory size for the mode dependent kernels is roughly 4 MB in current version.

Extension for larger block sizes? Was not considered yet.

RA results not ready yet. It is also applied to intra blocks in inter coded slices.

Investigate in EE

Percentages of usage should be reported.

Results when training the transforms with only the LFNST training set should be reported

[JVET-AB0180](https://jvet-experts.org/doc_end_user/current_document.php?id=12107) Non-EE2: CCCM using non-downsampled luma samples [H.-J. Jhu, C.-W. Kuo, X. Xiu, W. Chen, N. Yan, C. Ma, X. Wang (Kwai), V. Seregin, Y.-J. Chang, B. Ray, M. Karczewicz (Qualcomm)]

In ECM-6.0, convolutional cross-component model (CCCM) is applied to predict the chroma samples from reconstructed luma samples with downsampling. Moreover, the CCCM extends the 2-parameter linear model in the CCLM to 7-tap non-linear model. In this proposal, CCCM using non-downsampled luma samples is proposed where the chroma samples are directly predicted from the reconstructed luma samples, i.e., without downsampling. The proposed method was implemented on top of ECM-6.0, and the simulation results are summarized as follows:

For natural-sequences:

AI { -0.02%, -0.31%, -0.31%, 101%, 100% }, RA { -0.05%, -0.13%, -0.26%, 100%, 99% }

For TGM sequences:

AI { -1.89%, -3.62%, -3.36%, 101%, 100% }, RA { -0.70%, -1.83%, -1.52%, 100%, 100% }

Downsampling filtering has impact that in case of sharp edges (which is more noticeable in case of screen content) CCCM would not be used. A block-level flag is used to disable it.

It was suggested that the high-level flag might be used to disable the block-level signalling (JVET-AB0187 is proposing something like this)

Investigate in EE.

[JVET-AB0203](https://jvet-experts.org/doc_end_user/current_document.php?id=12130) Crosscheck of JVET-AB0180 (Non-EE2: CCCM using non-downsampled luma samples) [X. Li (Alibaba)] [late]

[JVET-AB0181](https://jvet-experts.org/doc_end_user/current_document.php?id=12108) Non-EE2: Using prediction samples or residual samples for adaptive loop filter [C. Ma, X. Xiu, C.-W. Kuo, W. Chen, H.-J. Jhu, N. Yan, X. Wang (Kwai)]

In ECM-6.0, ALF uses the reconstructed samples after SAO and before deblocking filter as input. In this contribution, it is proposed to additionally utilize the prediction samples or the residual samples as the input to the ALF. Two filter shapes, 1x1 and 3x3 diamond shapes, are tested to be applied on the additional ALF input. The simulation results on top of ECM-6.0 are summarized as follows.

Test 1: prediction samples, 1x1 filter shape

AI: {-0.07%, 0.00 %, 0.00%, 101%, 101%}; RA: {-0.06%, 0.03 %, 0.01 %, 100%, 100%}; LDB: {0.03%, 0.08 %, 0.29%, 100%, 100%};

Test 2: prediction samples, 3x3 filter shape

AI: {-0.08%, 0.01%, 0.01%, 102%, 102%}; RA: {-0.06%, 0.10 %, 0.00 %, 101%, 101%}; LDB: {0.04%, 0.01 %, 0.31%, 101%, 101%};

Test 3: residual samples, 1x1 filter shape

AI: {-0.07%, 0.01 %, 0.01 %, 101%, 101%}; RA: {-0.07%, 0.02 %, -0.01 %, 100%, 100%}; LDB: {-0.01%, 0.10 %, 0.40 %, 100%, 100%}.

Test 4: residual samples, 3x3 filter shape

AI: {-0.07%, 0.01%, 0.01 %, 102%, 102%}; RA: {-0.07%, 0.07 %, 0.09 %, 101%, 101%}; LDB: {-0.01%, 0.01 %, 0.30%, 101%, 101%}.

The four tests are performed independently of each other.

There is one additional flag in the ALF APS, which is always set to true in the proposal, which means effectively one of the proposed method is always used.

Cross checkers confirmed the results, expressed support for investigation in EE because of the performance vs. complexity tradeoff.

The gains are not large as currently proposed, but are claimed to be comparable to some other proposals recently investigated in EE.

Investigate in EE.

[JVET-AB0234](https://jvet-experts.org/doc_end_user/current_document.php?id=12163) Cross-check of JVET-AB0181 on Using prediction samples or residual samples for adaptive loop filter [X. Li (Google)] [late] [miss]

JVET-AB0268 Crosscheck of JVET-AB0181 Test 3 (Non-EE2: Using prediction samples or residual samples for adaptive loop filter) [J. Chen (Alibaba)] [late]

[JVET-AB0182](https://jvet-experts.org/doc_end_user/current_document.php?id=12109) Non-EE2: Bi-predictive local illumination compensation [X. Xiu, N. Yan, H.-J. Jhu, W. Chen, C.-W. Kuo, C. Ma, X. Wang (Kwai)]

In ECM-6.0, local illumination compensation (LIC) is one inter tool to address the illumination variation between one current block and its prediction block. The tool is only enabled to uni-predicted blocks. This contribution proposes to extend the existing LIC design to bi-prediction. Specifically, when applying the LIC to bi-predicted blocks, two sets of LIC parameters (i.e., scale and offset) are calculated and applied to L0 and L1 prediction blocks separately, which are then combined to form the final bi-prediction. One iterative approach is used to calculate the LIC parameters in L0 and L1 by minimizing the difference between the template samples of the current block and their prediction samples.

The proposed method is tested based on ECM-6.0 software with tool-on configuration (BASE\_ENCODER = 1, BASE\_NORMATIVE = 1 and TOOLS = 0). It reportedly shows that the proposed bi-predictive LIC scheme can provide 0.33% and 0.29% BD-rate savings for RA and LDB configurations, respectively, over the existing uni-predictive LIC.

Reported results are not CTC. Performance results under CTC are currently not available.

Multiple experts commented that the gains are interesting, especially if they could be retained under CTC.

Further study is encouraged, the proponents are asked to implement the proposed method on top of the latest version of ECM and bring results under CTC.

[JVET-AB0245](https://jvet-experts.org/doc_end_user/current_document.php?id=12174) Crosscheck of JVET-AB0182 (Non-EE2: Bi-predictive local illumination compensation) [G. Li (Tencent)] [late]

[JVET-AB0185](https://jvet-experts.org/doc_end_user/current_document.php?id=12112) Non-EE2: ALF with Diversified Extended Taps [W. Yin, K. Zhang, L. Zhang (Bytedance)]

In the current adaptive loop filter (ALF) design, online-trained filters contain 3 kinds of taps: spatial taps, reconstruction-before-DBF based taps and fixed-filter-output based taps. In this contribution, several diversified extended taps are introduced to provide additional texture information for ALF. The extended taps can take either reference samples or the filtered samples generated by feeding reconstruction-before-DBF into fixed filters as input sources.

On top of ECM-6.0, simulation results are reported as below:

RA: -0.15%, 0.10%, 0.05%, 103%, 106%.

LB: -0.11%, -0.11%, 0.43%, 106%, 106%.

Current design does not apply to all intra configuration, and hence has no performance impact.

Two inputs may be used by the proposed method, and one APS flag is used to indicate which one is used.

For intra coded pictures in RA and LB, the second kind of input (offline filter output generated by reconstruction before deblocking) is used.

The ALF design in ECM-6.0 is extended by up to 14 taps.

This proposal uses multiple pictures in filtering, both L0 and L1 reference pictures are used.

It was commented that the performance vs. runtime tradeoff looks interesting.

Investigate in EE

[JVET-AB0187](https://jvet-experts.org/doc_end_user/current_document.php?id=12114) Non-EE2: No luma subsampling for CCCM [V. Seregin, Y.-J. Chang, B. Ray, M. Karczewicz (Qualcomm)]

In ECM, convolutional cross-component intra prediction model derives the model between luma and chroma. In 4:2:0, luma component is subsampled first before deriving and applying the prediction model. In this contribution, it is proposed to disable the subsampling stage for CCCM.

Test results reportedly show BD-rate reduction for Y, U, V components as follows.

AI configuration:

Class F: -0.20%, -0.45%, 0.06%; 101% enc time, 98% dec time

Class TGM: -1.35%, -2.48%, -2.17%; 100% enc time, 102% dec time

RA configuration:

Class F: -0.44%, -0.74%, -0.25%%; 100% enc time, 100% dec time

Class TGM: -0.58%, -1.31%, -1.14%; 99% enc time, 102% dec time

This is related to JVET-AB0180. Two main differences from JVET-AB0180:

* Sequence level decision on whether to downsample luma, this proposal is only applied to SCC sequences
* CCCM filter shape is different, this proposal uses 5x5, and JVET-AB0180 uses 3x2

Investigate in EE together with JVET-AB0180, block-level decision vs. sequence-level decision, as well as performance impact from filter shapes.

[JVET-AB0188](https://jvet-experts.org/doc_end_user/current_document.php?id=12115) Non-EE2: Extensions of intra block copy [Y. Wang, K. Zhang, L. Zhang, N. Zhang (Bytedance)]

This contribution presents three aspects to extend the use of intra block copy (IBC). In aspect #1, combined IBC and intra prediction (IBC-CIIP) is proposed, in which the prediction of a CU is obtained by blending two predictions generated by IBC and intra prediction. In aspect #2, IBC with geometry partitioning (IBC-GPM) is proposed, in which a CU is divided into two sub-partitions geometrically, predicted by IBC and intra prediction individually. In aspect #3, IBC with local illumination compensation (IBC-LIC) is proposed. On top of ECM-6.0, simulation results of the proposed method are reported as below:

AI: Class F { -2.50%, -2.56%, -2.55%; 160%, 101%}, Class TGM { -1.27%, -1.14%, -1.13%; 157%, 101%};

RA: Class F {-1.60%, -1.58%, -1.63%; 109%, 100%}, Class TGM { -0.36%, -0.33%, -0.28%; 104%, 99%}.

Test 2: A simplified version:

AI: Class F { -2.31%, -2.26%, -2.14%; 121%, 99%}, Class TGM { -1.08%, -0.98%, -1.03%; 122%, 101%};

RA: Class F {-1.44%, -1.38%, -1.56 %; 104%, 100%}, Class TGM { -0.31%, -0.38%, -0.30%; 102%, 101%}.

It was commented that aspect #3 (IBC-LIC) gives additional gain on top of aspects #1 (IBC-CIIP) and #2 (IBC-GPM) without additional run time increase. The runtime increase in the proposal mainly comes from aspects #1 and #2.

The simplifications are encoder only by reducing number of RDO checks.

It would be desirable to futher reduce encoding time by other fast encoding algorithms as well as code optimization.

Investigate in EE, including encoder speedup methods.

[JVET-AB0189](https://jvet-experts.org/doc_end_user/current_document.php?id=12116) AHG12: On bit length control of regression based affine merge candidate derivation [Y. Zhang, H. Huang, V. Seregin, C.-C. Chen, M. Karczewicz (Qualcomm)]

This contribution provides a solution for bit length control of regression based affine merge candidate derivation method. Certain thresholds are defined to reduce the bit length of the input data to the linear regression process. With the proposed changes, simulation result is as following:

RA: {-0.05%, 0.03%, -0.07% }; LB: {-0.05%, -0.01%, 0.16%}.

In ECM-6.0, temporary fixes are introduced to control the dynamic range of determinant of some correlation matrices used in regression based affine. But these fixes would bring small performance loss. This proposal suggests to control the dynamic range of input data along with some intermediate right shifting when computing determinants of some matrices. Small performance gain can be achieved.

Cross checker confirmed that the results are matched. Further, the runtime from crosschecker is reported to be reliable, and shows 98% for both encoding and decoding.

Decision: adopt JVET-AB0189 to ECM-7.0

[JVET-AB0217](https://jvet-experts.org/doc_end_user/current_document.php?id=12144) Cross-check of JVET-AB0189: "AHG12: On bit length control of regression based affine merge candidate derivation" [F. Le Léannec (Xiaomi)] [late]

[JVET-AB0190](https://jvet-experts.org/doc_end_user/current_document.php?id=12117) Non-EE2: Combination of JVET-AB0094 and JVET-AB0095 for screen content [J.-Y. Huo, X. Hao, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.), J. Ren, M. Li (OPPO)]

Proposed is a combination of the methods proposed in JVET-AB0094 and JVET-AB0095.

The experimental results are as below:

For AI configuration:

F: -0.47%, -0.66 %, and -0.52 %, with xxx % EncT, 100 % DecT.

TGM: -2.20 %, -2.38 %, and -2.91 %, with xxx % EncT, 100 % DecT.

No need to present, already discussed along with JVET-AB0094 and JVET-AB0095, and those will be further investigated in EE.

[JVET-AB0191](https://jvet-experts.org/doc_end_user/current_document.php?id=12118) Non-EE2: Combined intra block copy and intra mode [C. Ma, X. Xiu, W. Chen, J.-H. Jhu, C.-W. Kuo, N. Yan, X. Wang (Kwai)]

This contribution proposes to combine IBC prediction with intra prediction. The design of the proposed method is similar to the current CIIP mode except that the inter part in CIIP is replaced with IBC mode. The method was tested on top of ECM6.0 and the simulation results are summarized as follows

Results on natural content compared with ECM6.0 with IBC on

AI: {-0.08%, -0.20%, -0.16%, 100%, 100%};

Results on Class F

AI: {-0.09%, -0.10 %, 0.04%, 100%, 101%}; RA: {0.00%, -0.05%, 0.10%, 99%, 100%};

A revised contribution to be uploaded with updated results on class F, additinoal results on class TGM, and updated slide deck.

Further encoder only improvements have been made since the last version of the contribution was uploaded, and presentation was made based on these numbers not yet available in the archive.

Results on Class F and Class TGM as presented:

Class F :

AI: {-0.12%, -0.14 %, 0.15%, 100%, 101%}; RA: {-0.05%, -0.07%, -0.21%, 99%, 100%};

Class TGM :

AI: {-0.15%, -0.02 %, -0.01%, 100%, 100%}; RA: {-0.10%, -0.02%, -0.03%, 100%, 100%};

It was commented that this is similar in concept to aspect #1 in JVET-AB0188. It is a subset of that aspect in JVET-AB0188, it does not increase runtime and has lower performance improvement.

It was commented that the results on natural content use non-CTC (by turning IBC on), and turning on IBC for natural content has previously been decided to not give good performance vs. runtime tradeoff.

It was commented that the gains in the SCC category are also relatively small.

Investigate in EE as a subtest of the tests based on JVET-AB0188.

[JVET-AB0192](https://jvet-experts.org/doc_end_user/current_document.php?id=12119) Non-EE2: Extended partitioning mode for the inter/intra prediction [Y. Kidani, H. Kato, K. Kawamura (KDDI)]

This contribution proposes two extended partitioning modes for the inter/intra prediction to further enhance coding performance of ECM. The first is an extended geometric partitioning mode (GPM) that increases the discreate angles of the existing GPM from 10 to 16. The second is L-shaped partitioning mode (LSPM) that separates a coding block into two areas with two orthogonal straight lines (i.e., L-shaped partitioning), different from the existing GPM. The extended GPM and LSPM is signaled at a coding block level by extending the existing GPM partition index. The signalling overhead for them is suppressed by the template-based reordering for the GPM partition index adopted in the current ECM. Therefore, the extend GPM and LSPM can improve the prediction accuracy in more complicate object boundaries in addition to the existing GPM.

Simulation test results of over all classes on the top of ECM-6.0 in RA and LB are reported as follows:

* Extend GPM only enabled
  + RA (Overall): BD-rate Y/U/V: -0.03%/0.00%/-0.01%; EncT: 101%; DecT: 100%
  + LB (Overall): BD-rate Y/U/V: -0.02%/0.12%/0.19%; EncT: 102%; DecT: 101%
* LSPM only enabled
  + RA (Overall): BD-rate Y/U/V: -0.03%/0.00%/-0.04%; EncT: 102%; DecT: 100%
  + LB (Overall): BD-rate Y/U/V: -0.10%/-0.05%/0.25%; EncT: 102%; DecT: 100%
* Extended GPM + LSPM enabled
  + RA (Overall): BD-rate Y/U/V: -0.07%/0.03%/-0.02%; EncT: 104%; DecT: 100%
  + LB (Overall): BD-rate Y/U/V: -0.08%/-0.04%/0.23%; EncT: 106%; DecT: 101%

The increase in complexity (both in terms of number of additional partitioning modes, and increase in decoder run time) is non-negligible, while the benefit in compression is relatively low.

Spatial GPM is not applied so far. One expert suggests to test that, as it might increase the compression benefit.

Further study (not EE yet).

# High-level syntax (HLS) and related proposals (27)

## AHG9: SEI messages on neural-network post filter (15 + 58 ballot comments)

Contributions in this area were discussed at 1200–1330 and 1500-1930 on Thursday 20 October 2022 (chaired by JRO). Follow-up review in BoG JVET-AB0244 on Friday 21 October 2022 (chaired by S. Deshpande), and in JVET plenary at 0900-1100 on Sunday 23 October (chaired by JRO).

[JVET-AB0193](https://jvet-experts.org/doc_end_user/current_document.php?id=12120) AHG9: A summary of proposals on NNPF SEI messages [Y.-K. Wang (Bytedance)]

The following design questions relating to ballot comments and input documents below are summarized in this contribution ("BC#nnn" refers to ballot comment #nnn from MPEG input document m60678).

1. On NNPF input and output format
   1. Is the syntax for specification of input and output formatting really necessary? Doesn’t the NNR standard itself determine its input and output format behaviour? Can this information conflict or require a conversion process that may be unnecessary? Remove any unnecessary input/output formatting specification. (BC#035)
   2. Replace nnpfc\_inp\_format\_flag and nnpfc\_out\_format\_flag with ue(v)-coded nnpfc\_inp\_format\_idc and nnpfc\_out\_format\_idc and reserve values greater than 1. (BC#009)

From the discussion:

It was reported that NNR specifies the network in a certain expression of precision and number format of weights, and the topology. There is no specification of input and output ranges and precision. There is also no requirement of implementing the network in a certain precision.

In case of floating point, the current SEI message is specifying input and output ranges of 0..1. In case of integer, it is assumed to unsigned in the range of the given bit depth. Clipping is also to be performed at the output.

It is pointed out that the specific conversion necessary could also be highly dependent on the capabilities of a processing device available. Further, if a decoder delivers 12 bit output, it would be strange if the input is truncating it to less bits. Also, the output format depends on the subsequent device.

It was suggested that it would be sufficient fo the output to specify the semantics of how to interpret the output in terms of ranges, such that the subsequent device can deal with it as appropriate.

The input format also specifies the arrangement of tensors, how to construct patches (which could also be overlapping) from the decoder output picture. The network may have been optimized for a given bit depth at the input, therefore a conversion for a bit depth (or precision) different from the decoder output is justified.

Decision:

- It was agreed on item a) that the current specification input formatting specification is appropriate and can be justified, whereas for the output format it is sufficient to specify the interpretation of value ranges semantically. The syntax for signaling integer or float would be kept. (M. Hannuksela volunteered to provide appropriate text for the output format part – see further discussion under JVET-AB0258.

- Item b) was agreed (both for input and output) – ballot comment #009 to be accepted.

1. (On the external-means mode and the URI mode of NNPF) Keep the external-means mode and the URI mode (the modes with nnpfc\_mode\_idc equal to 0 and 2)?
   1. Yes (BC#011)
      1. Move nnpfc\_uri\_tag[ i ] and nnprf\_uri[ i ] to be outside the conditional if statement based on nnpfc\_purpose\_and\_formatting\_flag? – agreed
         1. Yes, move the gated syntax that contains nnpfc\_uri\_tag and nnpfc\_uri just after nnpfc\_mode\_idc. (BC#010 item 1, JVET-AB0047 item 1) – agreed
         2. Yes, keep the syntax elements just after the nnpfc\_complexity\_element( ) syntax structure. (JVET-AB0059 item 1) (it was commented that thisis conceptionally identical to 1.)
      2. Add a requirement to have nnpfc\_purpose\_and\_formatting\_flag equal to 1 when nnpfc\_mode\_idc is equal to 0 or 2. (BC#010 item 2, JVET-AB0047 item 2) – agreed for the URI mode, and in correspondence with v. it should only be mandatory for an SEI that is not an update of one that was sent previously.
      3. Consider allowing the use of Rec. ITU-T T.35 data as an additional alternative way of identifying processing to be applied for this SEI message, since T.35 data is already supported in the standard and can fulfil a similar purpose. (BC#020) – agreed in principle that this could be useful, requires clarification if this should become a separate mode, and text needs to be developed (S. McCarthy will provide first draft – see further discussion under JVET-AB0266).
      4. Specify a tag URI value (e.g., "tag:iso.org,2023:15938-17") indicating that the neural network identified by nnpfc\_uri conforms to ISO/IEC 15938-17. (BC#024, JVET-AB0047 item 3) – agreed. It is noted that this would be followed by a second URI which contains the specific NNR instantiation. It would refer to the newest edition of 15938-17.
      5. Add and modify constraints for the case when nnpfc\_mode\_idc is equal to 2 to allow an update of a previously signalled NNPF. (JVET-AB0059 item 2) – agreed to allow an update also in URI mode, however it was commented in general that the meaning of “update” needs to be improved in the spec. (JVET-AB0047 has some editorial suggestions on that aspect)
   2. No, remove the external-means mode and the URI mode (the modes with nnpfc\_mode\_idc equal to 0 and 2). (BC#019)

Discussion:

- It was mentioned that the case of receiving the NN description out-of-band is beneficial (this could include the case of NNR being specified out of band

- It was argued that it is unknown with external means if it is actually a neural network post processing that is invoked. However, there are elements in the syntax which describe what the processing is about, which are still invoked in modes 0 and 2 (see item ii) above, which makes signaling the purpose and the formatting mandatory).

- The tag URI mechanism proposed in BC#011 for mode 2 could resolve the majority of concerns from the previous bullet point.

- It was agreed to retain only the NNR mode (to become mode 0) and the URI mode (to become mode 1, and potentially a new T.35 mode, see iii. above). The signaling mechanism shall be kept extensible.

1. On NNPF complexity
   1. Change the syntax of nnpfc\_complexity\_element( ) to get rid of the part that does not make sense. It was commented that the syntax specification for nnpfc\_complexity\_element( ) does not make much sense – e.g. the syntax structure is sent whenever nnpfc\_complexity\_idc is greater than 0, but it doesn’t have anything in it unless nnpfc\_complexity\_idc is equal to 1. (BC#017)
   2. It is asserted that the current extension mechanism in the syntax structure for the complexity elements does not allow for signalling additional complexity syntax elements while also signaling the first set of complexity syntax elements. Modify the nnpfc\_complexity\_element( ) syntax structure such that each bit in the binary representation of nnpfc\_complexity\_idc corresponds to one group of syntax elements. (JVET-AB0135 item 1)
   3. Add two additional syntax elements (nnpfc\_total\_kilobyte\_size and nnpfc\_num\_layers) descriptive of the complexity of the neural network signalled by the NNPFC SEI message. (JVET-AB0135 item 1)

Discussion:

- SEI messages have usually not been extended so far according to general policies. A device can ignore an SEI message if it cannot interpret it by not knowing the message type, but it should not start decoding it to find out if it can interpret it or not. SEI messages do not have profiles.

- Several experts expressed an opinion that providing some kind of extension mechanism might be beneficial for this specific SEI message, and it was already agreed that potentially an extension for signalling of the NNPF mode could be foreseen. The complexity syntax elements currently are located at the very end of the SEI message before the NNR payload data. For extensibility, more information could be added after NNR data, or by defining a new NNPF mode.

- It was agreed to replace the current nnpfc\_complexity\_idc syntax element by a flag. This resolves a. and b.

- For item c, it was agreed that the nnpfc\_total\_kilobyte\_size is useful (memory needs), whereas the number of layers may not be too beneficial for a decoding device to decide about its ability to process. It was suggested that more study might be useful to better express the complexity, e.g. regularity of an architecture might also have impact on the complexity.

1. On NNPF purposes
   1. Define and signal different types of visual quality improvements (note that visual quality improvement is one of the specified NNPF purposes), including the following (BC#021, JVET-AB0049 item 2):
      1. General visual quality improvement, targeting at increasing either the fidelity or the subjective visual quality of the reconstructed picture after applying the neural-network post-processing filter. The improvement may be measured by any objective or subjective metric.
      2. Objective-oriented/fidelity-oriented visual quality improvement, targeting at increasing the fidelity of the reconstructed picture after applying the neural-network post-processing filter. The fidelity may be measured by PSNR, MS-SSIM etc.
      3. Subjective-oriented visual quality improvement, targeting at increasing the subjective visual quality of the reconstructed picture after applying the neural-network post-processing filter. The subjective visual quality may be measured by LPIPS or MOS.
      4. Film grain-oriented visual quality improvement, with synthesizing filter grain on the reconstructed picture after applying the neural-network post-processing filter.
   2. Specify the value 0 for nnpfc\_purpose to be a specific purpose instead of "unknown or unspecified". (BC#034, JVET-AB0049 item 3)
   3. Specify the value range of nnpfc\_purpose to be 0 to 255, inclusive (instead of 0 to 232 − 2, inclusive). (JVET-AB0049 item 4)
   4. Add frame rate upsampling as an additional purpose. (JVET-AB0058)
      1. Conditionally signal two syntax elements related to frame rate upsampling information.
      2. Add processes for deriving input tensors and storing output tensors.

Discussion:

- It was pointed out that definition of sub-purposes might be too granular; sometimes the boundaries between sub-purposes might be vague (e.g. different variants of visual quality improvement)

- There are many other potential purposes of post processing, e.g. artistic effects (style generation), de-hazing, view synthesis, object tracking, annotation, … all these would currently fall under the “unspecified” category.

- it was argued that depending on the purpose (e.g. surveillance), a decoder might decide not doing some specific post processing such as subjective quality improvement. Film grain insertion is likely only be done for certain materials.

- Further study appears necessary to come up with a list of purposes that is useful for the current version of this SEI message. The current purpose are very much related to the question if the output of the network in terms of spatial/color resolution is identical to the input. The frame rate upsampling would be similar in terms of adding the temporal dimension to that.

- About item b., It is generally agreed that the definition of an unspecific purpose may not be desirable, however as long as we don’t have a more complete set of potential purposes, it is the only way to use it for purposes that don’t fall into one of the specified categories.

- About item c., restricting to a maximum of 255 purposes appears too low from the discussion above. Agreed to define a range of 0..1023, with a value reserved for potential future extensions.

- About item d., a presentation of JVET-AB0058 was given (including a demo of an implementation). It was commented that the syntax specification of input tensors would require an update ([0][idx][0]). No limitation on number of input/output pictures. It was commented that in its current version the SEI message does not restrict to pictures that are in terms of decoder timing available in the DPB. The output pictures typically require additional memory. It was further commented that in terms of timing it would be required to associate the activation SEI with the last input picture, and the other input pictures are the corresponding preceding images in output order with temporal ID less than or equal to the TID of the last input picture. Agreed to add this as purpose 5, adopt JVET-AB0058 with the changes suggested. It was further suggested that it might be considered to be used in temporal sublayers as well (further study recommended on that aspect, and if that is relevant at all).

1. (On NNPF output picture width and height signalling) Make the following changes (BC#022, JVET-AB0049 item 6):
   1. Instead of directly signal the filter output picture width and height values, the delta values of them compared to the filter input picture width and height values, respectively, are signalled using ue(v)-coded syntax elements, e.g., named nnpfc\_delta\_pic\_width and nnpfc\_delta\_pic\_height, respectively.
   2. The value ranges of nnpfc\_delta\_pic\_width and nnpfc\_delta\_pic\_height are specified, e.g., as follows: The value of nnpfc\_delta\_pic\_width shall be in the range of 0 to CroppedWidth \* 15 − 1, inclusive. The value of nnpfc\_delta\_pic\_height shall be in the range of 0 to CroppedHeight \* 15 − 1, inclusive.

Discussion:

- It was pointed out that saving some bits (per a.) is not overly important in SEI messages. No action on this aspect.

– It was agreed that specifying a value range is needed. However, if the SEI message retains signaling the actual width and height of the output, the range should be specified such that the CroppedWidth|Height is the minimum, and CroppedWidth|Height \* 16 – 1 is the maximum.

BoG (S. Deshpande) to continue review of the open items from here. Meeting in parallel with EE2 activities in JVET plenary, starting from 1400 CEST on Friday 21 Oct. See further disposition under JVET-AB0244 below.

[JVET-AB0244](https://jvet-experts.org/doc_end_user/current_document.php?id=12173) AHG9: BoG Report on Neural-network Post-filter Characteristics SEI Message [S. Deshpande] [late]

Was presented at 0900 on Sunday 23 October.

BoG discussed following items (please note the numbering below starts at 6 to match the corresponding identically numbered items in JVET-AB\_Notes-d\* document:

1. On NNPF auxilary inputs
   1. Move the normalization of the quantization strength input variable StrengthControlVal from the VSEI text to the use of NNPFC SEI message in the VVC text. (JVET-AB0046 item 1)
   2. Change the normalization of StrengthControlVal from 2(StrengthControlVal − 42) / 6 to a division (÷) by 63. With this change, it is asserted that StrengthControlVal becomes a variable that has been normalized to the range 0..1. In Table 23, insert the given StrengthControlVal directly into the input tensor: inputTensor[ ][ ][ ][ ] = StrengthControlVal. (JVET-AB0046 item 2, BC#027)
   3. Define and specify the use of more auxiliary inputs to the neural network post-processing filter, e.g., prediction samples. (BC#040, JVET-AB0074 item 1)
   4. Make the following change (BC#045, JVET-AB0058 and JVET-AB0074 item 2):

When nnpfc\_inp\_order\_idc is equal to 3, nnpfc\_component\_last\_flag is equal to 0, and nnpfc\_auxiliary\_inp\_idc is equal to 0, no auxiliary input matrix is used. In other words, remove the assignment operation "inputTensor[ 0 ][ 6 ][ yP + overlapSize ][ xP + overlapSize ] = 2(StrengthControlVal – 42)/6'' for this case from Table 23.

Discussion:

Regarding items a, b, the goal is to move VVC specific initialization to VVC text instead of it being in VSEI text (where it manifests as 2(StrengthControlVal − 42) / 6 – especially the values 42 and 6 in that equation).

BoG recommends to adopt items 6a and 6b, with additionally specifying in VSEI text that the range of StrengthControlVal is in the range of 0 to 1, inclusive.

Regarding item c: It was commented that we should consider which information the post processing filter has access to. For e.g. it was commented that information from slice layer above is easily accessible. Previously there was boundary strength information as an additional auxiliary input which was gotten rid of due to consideration that all post filters may not easily have access to that information. It was commented that such auxiliary information may be easily accessible to post filter via a software decoder. It was commented by one participant that in their opinion the prediction samples information may not easily be accessible if the decoder is not software decoder.

It was commented that the quoted references 1-2 in the proposal use this information via a modified VTM software.

It was also asked what is CroppedYPred, CroppedCbPred, CroppedCrPred and where it is defined. It was agreed by the proponents that maybe it is not clearly defined (and also not defined in VVC spec).

It was asked by proponents if our design principle is to have the auxiliary input information easily obtainable from all types of decoder.

It was decided to raise if this should be our design principle as a question to the larger group.

Recommendation: Discuss this design question in JVET.

From the discussion in JVET: Conceptually, decoding and SEI message processing have been separated so far and are considered as separate stacks. If specific decoding-internal information is accessed, the SEI message should rather be in VVC than VSEI. For the StrengthControlVal, the ongoing VVC amendment provides the corresponding interface. The syntax element nnpfc\_auxiliary\_input\_idc should be renamed. The input tensor is initialized with all its elements to StrengthControlVal; this might be used for other purposes with more local adaptation capability, but not in this version.

As a general design rule, only information that can easily be accessed from the video decoding process (e.g., slice header or above) should be used as additional input to an nnpf network.

Regarding item d: (Chaired by M. M. Hannuksela): It was agreed that this was an oversight bug in the current specification.

BoG recommends to adopt item 6d.

1. (On syntax elements specified as UTF-8 character strings UTF-8 ) There are two syntax elements that are specified as UTF-8 character strings. One of them says it shall be as specified in ISO/IEC 10646 and the other does not. Shouldn’t they both be as specified in ISO/IEC 10646? Clarify. (BC#028)
   1. Clarified by removing explicit reference to ISO/IEC 10646 from the semantics of nnpfc\_uri. Both of the syntax elements are of type st(v), which is specified in VSEI to be null-terminated string encoded as universal coded character set (UCS) transmission format-8 (UTF-8) characters as specified in ISO/IEC 10646. (JVET-AB0047, one of the editorial changes)

BoG recommends to agree to this editorial change (adopt 7a)

1. On value ranges for nnpfc\_id and nnpfa\_id and decoder handling of their reserved values
   1. When encountering an NNPFC SEI messages with a reserved value of nnpfc\_id, the decoder shall ignore the NNPFC SEI message instead of ignoring nnpfc\_id value. (BC#030, JVET-AB0049 item 1)
   2. Specify the value range and decoder handling of reserved values for nnpfa\_id similarly as for nnpfc\_id. (BC#055, JVET-AB0049 item 10)

Discussion:

About 8a: current spec says ignore the reserved values of nnpfc\_id. It was asserted by the proponent that this is not sufficient. It was commented that the current language is motivated by the corresponding language in frame packing SEI. It was agreed that this case is different than that.

BoG recommends to adopt item 8a.

About 8b: nnpfa\_id in the activation SEI message is similar to nnpfc\_id and so this is a bug-fix.

BoG recommends to adopt item 8b.

1. (On array convention) The standard generally uses an array convention in which the horizontal index is first and the vertical index is second. In several of these expressions, the indexing is swapped relative to that usual convention, e.g. for InpSampleVal( ) indexing. Is the order swapping necessary? Use the same convention consistently. If using an inconsistent convention is necessary (e.g. for compatibility with the NNR standard), highlight the difference in a NOTE to minimize reader confusion. (BC#036)

It was commented that this was intentionally done. This is because in inference engines that the commenter had checked (i.e. SADL, Tensorflow, PyTorch, MXNet) the vertical index was first and horizontal index was second. It was also commented that NNR does not have a particular convention.

It was commented that specifying this in our specification maybe just a matter of editorial convention.

It was commented by a few participants that since this is about deriving input tensors and storing output tensors, we should specify an order and that the tensors should use that order.

BoG recommends to keep the current order and add a NOTE.

1. On the value of nnpfc\_out\_sub\_c\_flag (BC#037, JVET-AB0049 item 5)

Replace the following sentence:

If SubWidthC is equal to 2 and SubHeightC is equal to 1, nnpfc\_out\_sub\_c\_flag shall not be equal to 0.

with the following:

When ChromaFormatIdc is equal to 2 and nnpfc\_out\_sub\_c\_flag is present, the value of nnpfc\_out\_sub\_c\_flag shall be equal to 1.

This discussion was chaired by M. M. Hannuksela.

It was commented that the current phrasing in the draft is equivalent to the proposed phrasing. It was commented that however the proposed phrasing is easier to understand.

BoG recommends to adopt item 10.

1. (On coupled or separate filtering of chroma components) Enable separate processing of chroma components, including processing only one of the chroma components by one NNPFC SEI message. (BC#044, JVET-AB0075)

Discussion: It was commented by proponent that current specification operates in the same manner on both chroma components and that same filtering is applied to the two components.

It was commented by participant that currently actual neural-network could decide to operate only on Cb or Cr (and ignore the other component).

It was also commented that maybe the same process could be applied if we change inp\_order\_idc to mean one input matrix (not just luma). To that it was commented that then we may need some other signaling which specifies whether the component was luma or Cb or Cr.

It was commented that this adds certain amount of additional specification text and cases which would need to be checked carefully.

BoG recommends Further study.

1. (On NNPF patch size) Specify the following: When nnpfc\_constant\_patch\_size\_flag is equal to 0, the patch size width, denoted by inpPatchWidth, and the patch size height, denoted by inpPatchHeight, are provided by external means. An example of such external means is an API that passes the values of inpPatchWidth and inpPatchHeight to the decoder and render entity in a video application system, and the values may be configured by a user through a user interface of the application. (BC#046, JVET-AB0049 item 6)

Discussion: It was commented by two participants that some clarification is indeed useful for nnpfc\_constant\_patch\_size\_flag equal to 0. The text in JVET-AB0049 item 6 was preferred. An additional suggestion was to add the quoted text:

The values of inpPatchWidth and inpPatchHeight are provided by external means not specified in this Specification “or determined by the post-processing unit”.

BoG recommends to adopt JVET-AB0049 item 6 with additional change noted above.

1. On NNPF activation, on/off control, and filter selection
   1. Enable activating an NNPF for a set of consecutive pictures in output order in a CLVS by using one NNPFA SEI message. (BC#049, JVET-AB0050, JVET-AB0060)
      1. Yes, enable it by one of the following (JVET-AB0050):
         1. Add a flag: nnpfa\_on\_flag.
         2. Add a new SEI message for deactivating a filter.
         3. Add two flags: nnpfa\_cancel\_flag and nnpfa\_persistence\_flag.
      2. Yes, enable it by not changing the NNPFA SEI message but adding an activation type to the NNPFC SEI message, with the following specified activation types: 0 - the NNPFA SEI message is applied; 1 – all pictures in the current CVS use this NNPF; 2 – all pictures in the current CVS with the same nuh\_layer\_id and TempralId, respectively, as the NAL unit containing the NNPFC SEI message use this NNPF. (JVET-AB0060)
   2. Replace nnpfa\_id with nnpfa\_activation\_id\_plus1 and specify that nnpfa\_activation\_id\_plus1 equal to 0 in the NNPFA SEI message indicates that no NNPF is activated for the current picture, to be able to explicitly indicate that no NNPF is activated for the current picture. (JVET-AB0134 item 1)

Item 13-a and 13-b was chaired by M. M. Hannuksela.

Discussion: The main usage scenario considered is activating same NNPF for multiple pictures.

It was commented that the current method only uses 1 byte per activation and if there is motivation for these proposals besides the bit savings. It was pointed out that rather it is one SEI message per activation which will be more than 1 byte so it could be about 6 bytes or so. It was also argued that if it is common usage to have same NNPF for multiple pictures then it may be a good design to support signaling as proposed in these proposals.

It was commented that the motivation for proposal b was to make the signaling robust to loss of SEI message(s).

It seemed there is some inclination to take an action based on above discussion about bit savings.

It was commented by a proponent of 13a that solution in 13b does not support signaling activating a NNPF for multiple pictures. Although proponent commented that item 13b can be combined with item 13a.

Thus, it was suggested to first decide regarding item 13a.

It was asked if it is a common use case to have different NNPF for different temporal sublayers and/or for different layers. It was answered by a proponent that for RA case since base QP is typically different for different temporal sublayer this may be the common. It was commented that this aspect may also be handled by using auxiliary input.

It was asked if the following case is supported by the proposals in item 13a: that a first NNPF if first activated and then a second NNPF filter is activated without deactivating the first NNPF. Is this allowed or forbidden. It was commented that this is already supported in design in current spec. For example, current specification allows activating multiple NNPFs for same picture.

It was commented by proponent of 13a-i that the aspect of 13a-ii which adds an activation type to the NNPFC SEI message is mixing the concept of NNPF activation and NNPF filter signaling which may not be desirable.

BoG recommends to adopt 13-a-i-3.

Discussion of b (and other aspects) after above recommendation:

A comment was made that we should consistently use same type of persistence across different SEIs. And that other SEI messages don’t have id associated with cancellation. It was commented that for NNPF you maybe have multiple NNPFs which can apply to same picture which is different aspect than those SEIs. It was commented that there may be a robustness issue for a splicer if cancelling is associated with an id. The counter argument was that typically a splicer starts a new CVS which cancels the active filter.

The aspect of changing the name nnpfa\_id to further distinguish it from nnpfc\_id is left to the editors. BoG recommends no further action on item b.

* 1. Enable activating different NNPFs for different regions (e.g., subpictures, slices, tiles, CTUs) of a picture? (BC#050, JVET-AB0152 option 2)
     1. Yes, activate one or more NNPFs in an NNPFA SEI message, and a selected NNPF can be used for the entire current picture, some slices of the current picture, or some CTUs of the current picture. (JVET-AB0152 option 2)

Discussion: It was commented CTU wise signaling and activation of different NN post filters was tested by one company. It was observed that the gain from this was somewhat cancelled by additional overhead (e.g of using CABAC coded CTU wise signaling).

Separately it was commented that such an idea has shown gain for NN in-loop filter.

It was asked if this would be useful for artistic use (different than coding gain aspect) – e.g. bokeh effect. There seemed some interest in that aspect.

There was a comment that JVET should first conclude on benefit of region-wise NN PF and then the syntax/ HLS aspect of SEI signaling should be considered.

It was commented that there could be a region-wise NNPF activation message in future and so we are not closing door on this.

BoG recommends further discussion in JVET.

In the discussion in JVET, this was recommended for further study. It would be straightforward to design a region-wise activation message, but the SEI message should define the regions by itself and not directly refer to partitioning of a video codec. Switching between different filters per region might also be conflicting with patch-based processing. A simple region-wise on/off switching would be straightforward (item d below would be doing that).

It was also commented that in the NNVC exploration some experimentation in AI setting (where the NN loop filter is actually a post filter) did show some benefit of switching on/off at CTU level. It is however unclear if these were designs which used many other informations from the video decoder, such that they cannot be really be classified as post.

* 1. Enable regional on/off control of NNPF? (JVET-AB0134 item 2, JVET-AB0152)
     1. Yes, activate one NNPF in an NNPFA SEI message and the NNPF can be used for the entire current picture or a list of explicitly signalled regions (with region width, height, top and left positions signalled) of the current picture. (JVET-AB0134 item 2a)
        1. Alternatively, same functionality, but add four more syntax elements targeting at reducing the signalling overhead. (JVET-AB0134 item 2b)
     2. Yes, activate one NNPF in an NNPFA SEI message and the NNPF can be used for the entire current picture, some slices of the current picture, or some CTUs of the current picture. (JVET-AB0152 option 1)
     3. Yes, also enable activating different NNPFs for different slices or CTUs of a picture. (JVET-AB0152 option 2)

It seems item 13-d-iii is covered by discussion in c above.

Regarding items 13-d i-ii, the high-level question is if we want to enable regional on/off control of NNPF.

Currently an NNPF active for a picture is applied to the entire picture. Do we want to allow only applying it to part of that picture (and turn it off for the other part of the picture).

There was no apparent example of doing such on-off control for NN PF, but there has been example of doing it for NN in-loop filter.

It was commented that doing NN PF on part of the picture would have less complexity that doing it on entire picture.

There was a comment that JVET should first conclude on benefit (including simplicity) of region-wise NN PF being turned on/ off and then the syntax/ HLS aspect of SEI signaling should be considered.

It was commented that there could be a region-wise NNPF on/ off message in future and so we are not closing door on this.

BoG recommends further discussion in JVET.

From the discussion in JVET, JVET-AB0134 item 2a was identified as a straightforward approach, however it would introduce significant overhead, and it is unknown if that would still provide benefit. An alternative approach might be a regular subdivision of a picture into a block structure, and signaling on/off for each block (JVET-AB0134 item 2b would be somewhat like that), but no evidence about potential benefit for a plain post filter is available. This is left for future study, and could easily be done by extending the activation message, or creating a new one.

1. On presence and decoding order of NNPFC and NNPFA SEI messages
   1. Specify the following constraint ("no NNPFC, no NNPFA"): An NNPFA SEI message with a particular value of nnpfa\_id shall not be present in a current PU unless within the current CLVS there is an NNPFC SEI message with nnpfc\_id equal to the particular value of nnpfa\_id present in a PU preceding the current PU in decoding order or in the current PU. (BC#053, JVET-AB0049 item 8)
   2. Specify the following constraint ("NNPFC precedes NNPFA"): When a PU contains both an NNPFC SEI message with a particular value of nnpfc\_id and an NNPFA SEI message with nnpfa\_id equal to the particular value of nnpfc\_id, the NNPFC SEI message shall precede the NNPFA SEI message in decoding order. (BC#054, JVET-AB0049 item 9)
   3. Is it allowed for more than one NNPFC SEI message to be present for the same picture? If so, what does this mean?

Is there some limit on the number of them? If more than one is present, what happens if the nnpfc\_purpose values are the same or if they are different? What happens if the nnpfc\_mode\_idc values are the same or if they are different? Add appropriate clarification, explanation and/or constraints. (BC#056)

Discussion 14-a: It was commented by multiple participants that “no NNPFC means no NNPFA” makes sense.

BoG recommends to adopt 14-a.

Discussion 14-b: It was commented by multiple participants that in that case NNPFC should precede NNPFA and the proposal makes sense.

BoG recommends to adopt 14-b.

Discussion 14-c: About “Is it allowed for more than one NNPFC SEI message to be present for the same picture”- the answer is the current spec allows it. For the rest of the questions the group agrees that this should be answered and clarified and that some sensibility constraints maybe needed. It was suggested that offline discussion maybe useful to arrive at how to clarify the answers to these questions and what (if any) specification text is needed to clarify this.

It was agreed that all these questions should be answered. Such a discussion is encouraged and if it results in a new input, that will be discussed.

1. (On NNR base and update) It is our understanding that a neural network representation (NNR) can either be “intra” (coded without reference to some other NNR) or predictively/incrementally coded. Is there a way to identify which is which? In the case of incremental NNR, what is the reference? For example, is the reference the “base” representation, or could it be the accumulated result of several updates relative to a base representation? In what order are the updates applied? Clarify. (BC#058)

It was commented that JVET-AB0047 provides editorial clarifications on this. These were described verbally.

The current understanding of the group:

It is our understanding that a neural network representation (NNR) can either be “intra” (coded without reference to some other NNR) or predictively/incrementally coded.

Is there a way to identify which is which? For NNR, profile 0 is intra. For NNR profile 1 can be intra or differential update.

Also, the answer to the next question defines “base”, which is the “intra” NNR and the other NNR is update of the base.

In the case of incremental NNR, what is the reference? The reference is the “base”, and the “base” is the NNR corresponding to first NNPFC SEI with a particular nnpfc\_id.

For example, is the reference the “base” representation, or could it be the accumulated result of several updates relative to a base representation? Answered above.

In what order are the updates applied? The updates are not cumulative (due to robustness issue if some updates are lost and that it may make SEI message design complicated), so an update is always applied on top of a base. Also, the order is the bitstream order.

BoG recommendation: Delegated to the editors to clarify based on JVET-AB0047 and above answers.

1. Editorial changes not covered above:
   1. BC#008, 012~016, 018, 023, 025, 026, 029, 031~033, 038, 039, 041~043, 047, 048, 051, 052, 057, 059~065
   2. JVET-AB0047, excluding the technical changes in items 1 to 3

Editorial aspects were not discussed in the BoG.

All BoG recommendations which are indicated not requiring additional discussion in JVET were agreed in the JVET plenary on Sunday 23 Oct.; for the items that required additional discussion, the corresponding agreements are inserted above.

Ballot comments in MPEG input document [m60678](https://dms.mpeg.expert/doc_end_user/current_document.php?id=84173&id_meeting=192)

Comments 008 to 065

[JVET-AB0046](https://jvet-experts.org/doc_end_user/current_document.php?id=11961) AHG9: On StrengthControlVal of the NNPFC SEI message [M. M. Hannuksela, M. Santamaria, F. Cricri (Nokia)]

[JVET-AB0047](https://jvet-experts.org/doc_end_user/current_document.php?id=11962) AHG9: nnpfc\_mode\_idc related changes to the NNPFC SEI message [M. M. Hannuksela, F. Cricri, M. Santamaria (Nokia)]

[JVET-AB0049](https://jvet-experts.org/doc_end_user/current_document.php?id=11964) AHG9: Miscellaneous aspects of the two neural-network post-filtering SEI messages [Y.-K. Wang, Y. Li, C. Lin, J. Li, K. Zhang, L. Zhang (Bytedance)]

[JVET-AB0050](https://jvet-experts.org/doc_end_user/current_document.php?id=11965) AHG9: Activation of a neural-network post-processing filter for multiple pictures [Y.-K. Wang, K. Zhang, L. Zhang, C. Lin, J. Li, Y. Li (Bytedance)]

[JVET-AB0058](https://jvet-experts.org/doc_end_user/current_document.php?id=11973) AHG9: Frame Rate Upsampling Information in Neural-network Post-filter Characteristics SEI Message [S. Deshpande, A. Sidiya (Sharp)]

[JVET-AB0059](https://jvet-experts.org/doc_end_user/current_document.php?id=11974) AHG9: Comments on Neural-network Post-filter Characteristics SEI Message [S. Deshpande (Sharp)]

[JVET-AB0060](https://jvet-experts.org/doc_end_user/current_document.php?id=11975) AHG9: On activation of the neural-network post-filter characteristics SEI message [T. Chujoh, Y. Yasugi, T. Ikai (Sharp)]

[JVET-AB0074](https://jvet-experts.org/doc_end_user/current_document.php?id=11989) AHG9: Auxiliary input for neural-network post-processing filter [Y. Li, J. Li, C. Lin, K. Zhang, L. Zhang, Y.-K. Wang (Bytedance)]

[JVET-AB0075](https://jvet-experts.org/doc_end_user/current_document.php?id=11990) AHG9: Separate processing of chroma components for neural-network post-processing filter [Y. Li, J. Li, C. Lin, K. Zhang, L. Zhang, Y.-K. Wang (Bytedance)]

[JVET-AB0134](https://jvet-experts.org/doc_end_user/current_document.php?id=12061) AHG9: On NN post-filter activation SEI message [M. Pettersson, R. Sjöberg, J. Ström (Ericsson)]

[JVET-AB0135](https://jvet-experts.org/doc_end_user/current_document.php?id=12062) AHG9: On complexity metrics for NN post-filter characteristics SEI message [M. Pettersson, R. Sjöberg, J. Ström (Ericsson)]

[JVET-AB0152](https://jvet-experts.org/doc_end_user/current_document.php?id=12079) AHG9: Regional on/off control and selection of NNPFs [J. Li, C. Lin, K. Zhang, L. Zhang, Y.-K Wang, Y. Li (Bytedance)]

[JVET-AB0258](https://jvet-experts.org/doc_end_user/current_document.php?id=12187) AHG9: Specification text for interpretation of value ranges of output tensors of neural-network post-filtering semantically [M. M. Hannuksela (Nokia)] [late]

Was reviewed offline by relevant experts, who confirmed that the text is appropriate, and no request was made to review the text in the JVET session on Sunday 23 October.

Decision: Adopt JVET-AB0258

[JVET-AB0266](https://jvet-experts.org/doc_end_user/current_document.php?id=12195) AHG9: Specification text for use of Rec. ITU-T T.35 syntax structure to specify a neural network in the neural-network post filter characteristics SEI message [Sean McCarthy (Dolby), Sachin Deshpande (Sharp)] [late]

Topic of T.35 mode for NNPFC was discussed in the BoG JVET-AB0244 at 6:10 CEST on Friday 21 October:

A draft for signaling was presented by S. McCarthy.

It was commented that the T.35 syntax structure should be copied into the NNPFC message when the nnpfc\_mode\_idc indicates signaling via T.35 syntax structure to resolve any ambiguity of which T.35 message has the NN data.

Further review of an updated text was presented in JVET on Sunday 23 October. It was commented that there was no good reason to enable two different ways of invoking non-in-band non-NNR signalling of NNPF strutures, and it might even be undesirable if they would be mixed in a bit stream. It was agreed not taking action.

## AHG9: SEI messages on topics other than NNPF (9 + 7 ballot comments)

Contributions in this area were discussed at 1115–1300 on Sunday 23 October 2022 (chaired by JRO), and at 1230-1400 on Monday 24 October 2022 (chaired by JRO).

Ballot comments in MPEG input document [m60678](https://dms.mpeg.expert/doc_end_user/current_document.php?id=84173&id_meeting=192)

Comments 001 to 006, 066

[JVET-AB0051](https://jvet-experts.org/doc_end_user/current_document.php?id=11966) AHG9: On the SEI processing order SEI message [Y.-K. Wang (Bytedance), Hendry (LGE)]

The following changes for the SEI processing order SEI message are proposed:

1. It is required that an SEI processing order SEI message shall contain at least two pairs of po\_sei\_payload\_type[ i ] and po\_sei\_payload\_type[ i ]. – agreed.
2. The value 0 for po\_sei\_processing\_order[ i ] is treated the same as other values of the syntax element. – agreed after offline discussion with original proponents.
3. For any two different integer values of m and n that are greater than or equal to 0, po\_sei\_processing\_order[ m ] greater than 0 and equal to po\_sei\_processing\_order[ n ] indicates that there is no preferred order of processing between the SEI messages with payloadTypes equal to po\_sei\_payload\_type[ m ] and po\_sei\_payload\_type[ n ]. –agreed.

[JVET-AB0069](https://jvet-experts.org/doc_end_user/current_document.php?id=11984) AHG9: On the SEI processing order SEI message [Y. He, M. Coban, M. Karczewicz (Qualcomm)]

This contribution proposed two independent modifications to the SEI processing order SEI message.

1. It is proposed to remove syntax element, po\_sei\_processing\_order[ i ], from the SEI message, and infer the preferred processing order to be the same as the signaling order of payload type, po\_sei\_payload\_type[ i ]. – no action on this.
2. It is proposed to change the data length of po\_sei\_processing\_order to u(16). – agreed.

It was commented that by removing the syntax element po\_sei\_processing\_order, some of the intended flexibility of the decoder would be given up. Further, this syntax element is deemed to be important (by the original proponents) where observing the intended processing order is important in generating a useful output.

Offline clarification was requested between the original proponents of the SEI processing order SEI message and the contributors of JVET-AB0051 and JVET-AB0069. It was reported in a follow-up discussion at 0900 on Thursday 27 Oct. that no agreement had been reached about removing that syntax element. As non consensus on this aspect was reached, the syntax element should be retained.

New version of draft JVET-AB2027, including the changes agreed from JVET-AB0051 and JVET-AB0069.

[JVET-AB0055](https://jvet-experts.org/doc_end_user/current_document.php?id=11970) AHG9: On leading pictures design in DRAP SEI Message [Hendry, S. Kim (LGE)]

It is asserted that there is text description disagreement regarding leading pictures associated with a DRAP picture in VVC and in VSEI specification. In Annex C of VVC specification, the leading pictures are not discarded in the event of random accessing from the DRAP and consequently it is reasonable to assume that they should be decodable leading pictures. On the other hand, in the semantics of DRAP SEI message in VSEI specification, there is no constraint that prohibits leading pictures of a DRAP picture to refer to pictures that precede the DRAP picture that is not the associated IRAP picture which means the leading pictures can be non-decodable / skipped leading pictures.

To overcome the asserted problem, it is proposed that text related to the leading pictures of a DRAP picture must be consistent in both VVC and VSEI specification. Specifically, it is proposed that all leading pictures associated with a DRAP picture are decodable leading pictures; hence, appropriate modification so that the needed constraint (i.e,. the leading pictures shall not refer to pictures that precedes the DRAP picture in decoding order that is not the associated IRAP picture) is reflected the semantics of DRAP SEI message.

The DRAP SEI message was originally designed in the HEVC context, and when carried over to VSEI for usage with VVC, due to a different design of HRD, this inconsistency occurs (which was known at the time when it was specified). Several experts expressed opinion that it would be too late to make the suggested change. It was suggested to add a note in annex C of VVC that a decoder might face a problem by not being able to decode some of the leading pictures.

Decision: Put a note into the errata report JVET-AB1004 (editor action), such that it might be included in the next version of VVC (DAM currently under ballot).

[JVET-AB0056](https://jvet-experts.org/doc_end_user/current_document.php?id=11971) AHG9: On leading pictures design in EDRAP SEI Message [Hendry, S. Kim (LGE)]

The current EDRAP SEI message has a flag called *edrap\_leading\_pictures\_decodable\_flag* that specifies whether or not leading pictures associated with the EDRAP picture are decodable leading pictures. This design is asserted to have the following problems:

1. In the random access from the EDRAP picture event, if the EDRAP picture has associated leading pictures and they are not decodable, they need to be removed / discarded from the bitstream to be decoded. However, currently there is no way to identify which pictures that follow the EDRAP picture in decoding order are leading picture and whether it is decodable leading or non-decodable leading picture without first parsing and decoding the POC value of the picture.
2. In the current design, it is a binary indication such that all leading pictures associated with an EDRAP picture are either decodable or unknown. It may be desirable to allow mix of non-decodable and decodable leading pictures such as the design of RASL and RADL pictures associated with a CRA picture. Particularly, if a mechanism is to be specified to allow identification of leading pictures associated with an EDRAP picture without having to derive the POC values, it may not cause additional complexity to allow mix of non-decodable and decodable leading pictures associated with an EDRAP picture.

To overcome the asserted problem, the following is proposed:

1. Introduce two new SEI messages to indicate leading pictures associated with an EDRAP picture. The SEI messages are:
   1. DADL indication SEI message to indicate DRAP-Decodable Leading (DADL) picture, that is, leading picture associated with an EDRAP picture that is decodable when random accessing from the EDRAP picture.
   2. DASL indication SEI message to indicate DRAP-Skipped Leading (DASL) picture, that is, leading picture associated with an EDRAP picture that is not decodable (and need to be skipped) when random accessing from the EDRAP picture.
2. Define necessary constraints to ensure DADL and DASL pictures work.
3. Replace the current *edrap\_leading\_pictures\_decodable\_flag* in the EDRAP SEI Message with *edrap\_leading\_pictures\_present\_flag* to specify whether the EDRAP picture may or may not be associated with leading pictures.

The suggested action #3 would create a compatibility issue with existing implementations of EDRAP SEI.

Currently, if a mix of decodable and non-decodable leading pictures would be present, it would be indicated as “non-decodable” in the EDRAP SEI. It was pointed out that a decoder could find out by analysis RPL and POC in slice header if a picture is decodable or not.

In the discussion, no other experts supported to take action on defining the two proposed SEI messages DADL and DASL.

[JVET-AB0057](https://jvet-experts.org/doc_end_user/current_document.php?id=11972) AHG9: On the associated IRAP for DRAP and EDRAP pictures [Hendry, S. Kim (LGE)]

It is asserted that in addition to having DRAP and EDRAP pictures to be associated with an IRAP, it should be possible to allow them to be associated with a GDR picture with ph\_recovery\_poc\_cnt equal to 0.

This contribution proposes the following:

1. In additional to DRAP / EDRAP pictures is associated with an IRAP, allow it to be associated with a GDR picture with ph\_recovery\_poc\_cnt equal to 0.
2. Update the text editorially as follows:
   * For DRAP and EDRAP pictures, define an anchor picture that is associated with a DRAP and / or an EDRAP picture. An anchor picture is a picture in which random access can be performed from that anchor picture and the anchor picture can be correctly decoded without referring to any other picture.
   * The list of pictures that can be used / considered as anchor picture for DRAP and / or EDRAP pictures includes: IRAP picture and GDR picture with ph\_recovery\_poc\_cnt equal to 0.

It was commented that the syntax element ph\_recovery\_poc\_cnt is not available in the VSEI spec. It would be necessary to express this differently.

From VSEI perspective, the definition of IRAP already includes GDR pictures with ph\_recovery\_poc\_cnt equal to 0. It was suggested to put an explanatory note in the VVC spec that the definition of IRAP in VSEI subsumes the case of GDR with ph\_recovery\_poc\_cnt equal to 0. Such a note should be put in places of VVC that refer to VSEI.

Decision: Put such a note into the errata report JVET-AB1004 (editor action), such that it might be included in the next version of VVC (DAM currently under ballot).

[JVET-AB0070](https://jvet-experts.org/doc_end_user/current_document.php?id=11985) AHG9: On inclusion of post-filter hint SEI message into VSEI [Hendry, J. Nam, S. Kim, J. Lim (LGE)]

This contribution proposes to include post-filter hint SEI message in the VSEI specification. It is asserted that application of filters to decoded pictures is desirable and useful for some use-case scenarios. Furthermore, one of the use-case scenarios in which the SEI message can be used is to signal the filters for resampling decoded pictures, post decoding and outputting process but prior to displaying, when reference picture resampling (RPR) feature is enabled so that the resolution of the pictures to be displayed can be the same.

It is also asserted that currently the post-filter hint SEI message has parsing dependency to the underlying codec (i.e., the need to get the chroma format information). This contribution suggests solution for it by signalling whether filters from chroma component is present or not.

Finally, it is also suggested that for the post-filter hint SEI message, the persistence of the SEI message is modified so that it may persist for more than one picture.

The following is proposed:

1. Include post-filter hint SEI message in VSEI specification.
2. Remove the parsing dependency related to the chroma format information by explicitly signal whether or not filter coefficients for chroma components are present.
3. Modify the persistent of post-filter hint SEI message to allow its scope to cover one or more pictures by using the cancel flag and persistent flag.

The inclusion of this well-known SEI message in VSEI was agreed to be useful.

During the discussion, it was suggested to check if the problem pointed out in terms of the parsing dependency from VUI might also exist in other SEI messages.

Regarding the persistency, this appears useful. It was reported that in AVC its persistence was just the current picture. At that time, it was mainly intended to be used as Wiener-based post filter, which was optimized for each picture. Being able to use the same filter for several pictures increases the flexibility.

Decision: Adopt JVET-AB0070. Beyond that (editor action item): Check if the VUI dependency exists in other SEI messages of VSEI, and if so, take appropriate action either in JVET-AB2006 or JVET-AB1004.

[JVET-AB0096](https://jvet-experts.org/doc_end_user/current_document.php?id=12023) AHG9: Resolution Change Information SEI message [V. Drugeon, K. Abe, T. Toma (Panasonic)]

Dynamic resolution encoding can be used for dynamically adapting the resolution of a video on a scene-by-scene basis to reach the desired bitrate while preserving the quality. However, many receivers have been developed with the assumption that the picture resolution within a bitstream stays constant, for example within an MPEG-DASH Representation.

It is suggested to consider signalling of dynamic resolution encoding in bitstreams, together with the resolutions used. The authors of the present document believe that such a signalling would simplify implementation of receivers in environments where bitstreams may use dynamic resolution encoding.

This input is a follow-up of JVET-AA0091 attempting to address the comments made at the last meeting not to make any promises and guarantees regarding the bitstream within an SEI message.

Persistence scope is the whole CVS (or could even be beyond, from this aspect conceptually similar to manifest and prefix SEIs). The SEI message is intended to inform a decoder in the beginning about what it can expect in terms of resolution changes.

Unclear if VSEI would support persistence across several CVSs. Same with HEVC.

It was discussed if this could be put as a DCI extension into VVC. However, compatibility with existing devices could be a problem in that case, e.g. with splicing resolutions might come into the stream which are not in the list.

It is pointed out that, if used in an application standard, the list of resolutions communicated in the SEI message could be exhaustive/binding. However, it was pointed out that in an application standard this might lead to a situation where always the same values are signalled.

Several experts expressed opinion that the usefulness of this SEI message is not obvious.

[JVET-AB0267](https://jvet-experts.org/doc_end_user/current_document.php?id=12196) AHG9: On phase indication SEI message persistence [[J. Samuelsson-Allendes](mailto:samuelssonj@sharplabs.com), S. Deshpande (Sharp)] [late]

This contribution presents a proposed modification to the semantics of the phase indication SEI message in response to ballot comment number 066 in m60678. It is proposed to remove the dependency on the VVC-specific syntax element ph\_pic\_parameter\_set\_id and replace it with dependency on the variables CroppedWidth and CroppedHeight. The variables CroppedWidth and CroppedHeight are not VVC-specific and are already used in other parts of the VSEI specification.

Decision: Adopt JVET-AB0267

[JVET-AB0269](https://jvet-experts.org/doc_end_user/current_document.php?id=12198) AHG9: Status of SEI descriptions in JVET-Z2002 [S. McCarthy, F. Pu (Dolby)] [late]

This contribution provides information on the status of descriptions of SEI message implementations in JVET-Z2002, as agreed during review of the JVET AHG report: SEI message studies (AHG9) (JVET-AB0009) during the JVET meeting on Thursday 20 October 2022.

JVET-Z2002-v2 provides the most recent algorithm descriptions for VVC and VTM, including SEI message implementations in VTM.

Information on the status of descriptions of SEI message implementations in VTM was requested during review of JVET AHG report: SEI message studies (AHG9) (JVET-AB0009) during the JVET meeting on Thursday 20 October 2022.

This contribution provides such information.

Rec. ITU-T H.266 (V2) | ISO/IEC 23090-3:2022 specify payloadType values for 37 SEI messages. An additional 7 SEI messages are included in amendments and drafts related to future anticipated versions of the VVC specification.

The following SEI message implementations in VTM are described in JVET-Z2002:

• Implementation related to film grain characteristics SEI message

• Parsing of Green Metadata SEI Messages

• Implementation related to shutter interval information SEI message

• Implementation related to neural-network post-filter characteristics and activation SEI messages

In addition, draft text on implementation related to SEI processing order SEI message has been provided. This requires a description in JVET-AB2002. It needs to be clarified if all other SEI messages that were recently implemented in VTM software 17.1, 17.2, 18.0, 18.1 need some description in the document. K. Sühring provided the following list of those which were most recently integrated:

**VTM 18.1**

-Add filler payload SEI message

**VTM 18.0**

New:

- JVET-A0110: Adding support for Phase indication SEI message

Modifications:

JVET-AA0056: AHG9: On syntax gating in the neural-network post-filter characteristics SEI message

JVET-AA0067\_NNPFC\_SEI\_FIX

JVET-AA0102 and JVET-AA2027: SEI Processing Order

JVET-AA0054\_PROPOSAL2: Signal 1 flag instead of 2 flags to specify output chroma format information

JVET-AA0055 Proposal A and B (NNPF SEI)

JVET-AA0054 Proposal 1: Signaling of external URI information in neural network post-filter characteristics SEI message

**VTM 17.1**

- JVET-Z0047: Improved film grain analysis

**VTM 17.0**

**-**JVET-Z0244: NN post-filter SEI

- JVET-Z0120: SII SEI support and illustration of use case "Backwards-compatible HFR video”

- JVET-Z0046: Green Metadata SEI

**VTM 16.1**

- JVET-Y0044 Signal Green metadata and VDI SEI messages (message types only)

- JVET-T0055 aspect4: SEI consistency

It is agreed that it is up to the discretion of editors of JVET-AB2002 to identify any missing descriptions and request original proponents of those SEI messages to provide these.

## Film grain synthesis (1)

Contributions in this area were discussed at 1445–1515 on Monday 24 October 2022 (chaired by JRO).

JVET-AB0122 is also related

[JVET-AB0042](https://jvet-experts.org/doc_end_user/current_document.php?id=11955) Proposed text: Film grain synthesis technology for video applications (toward Draft 3) [D. Grois, Y. He, W. Husak, P. de Lagrange, A. Norkin, M. Radosavljević, A. Tourapis, W. Wan]

This draft technical report provides guidance on the use of film grain synthesis technology for video applications. Such technology may be used in conjunction with metadata signalling mechanisms, such as the supplemental enhancement information messages available in several video coding standards. The purpose of this document is to provide a publicly referenceable overview of the end-to-end processing steps for film grain synthesis, which may include content analysis, noise/film grain removal and film grain model parameter estimation, parameter encoding, encapsulation, and decoding, and film grain synthesis and blending for consumer distribution applications.

This version reorganized the technical report in the following order: overview, synthesis, analysis, parameter descriptions and examples while also adding text in areas that were thinly populated. Several figures and pictures were added to provide visual examples to better instruct the reader. Examples have been moved to an Annex.

The question was raised if there was a rule of referencing to a URL in a technical report (as there are suc references to AOM and VVC software).

J. Ohm mentions that he might be able to find some literature references that could be used for mapping 2D autoregressive models of reduced order.

Decision: To be converted into output documents JVET-AB2020 (further improvements during editing period).

## Non-SEI HLS aspects (2)

Contributions in this area were discussed at 1515–1600 on Monday 24 October 2022 (chaired by JRO).

[JVET-AB0071](https://jvet-experts.org/doc_end_user/current_document.php?id=11986) On the selected schedule index for conformance test when RPR is enabled [Hendry, S. Kim (LGE)]

When reference picture resampling (RPR) feature is enabled, it is expected that the bitstream bitrate changes significantly during the period where picture resolution is different (i.e,. either bigger or smaller) than pictures before and after that period. Despite this, the selected schedule index for conformance test cannot be change for this event which means the channel bitrate (i.e., the one associated with the selected schedule index) stays the same during such picture resolution change.

It is asserted using the assumed channel bitrate that was selected from the beginning of the test during resolution change may cause coded picture buffer (CPB) underflow. Therefore, it is proposed that Hypothetical Stream Scheduler (HSS) is allowed to change the selected schedule index in the event that picture resolution changes.

The following is proposed:

* When RPR is enabled, for HRD operations and conformance test, the selected schedule index may be changed / updated at the AU where the resolution changes.
* The change of schedule index is done at AU with temporal id equal to 0.

It was asked if the change in bit rate is actually more problematic in RPR compared to other cases, e.g. when switching to a different resolution with an IDR picture?

It was pointed out that also in case of RPR the encoder has to observe the CPB/DPB buffer constraints, and that the number of bits currently in the buffer would be known.

It was pointed out that the change of scheduling parameters in the middle of a CVS is undesirable and might cause problems.

No evidence is shown that the current way of HRD specification has problems with RPR.

It is not obvious that any problem exists with the current design. Further, the suggested change might be in conflict with some existing implementations. No action on the proposal.

[JVET-AB0120](https://jvet-experts.org/doc_end_user/current_document.php?id=12047) On HRD delivery schedule interpolation [Y. Sanchez, R. Skupin, T. Schierl (HHI)]

In this document a proposal is described that suggests to include a formula for an interpolated value for the initial CPB removal delay offset in section C.5, in a similar way, as done for the interpolated value of initial CPB removal delay. It is asserted that this parameter is needed for conformance and is currently missing.

The history of these definitions can be traced back by approximately 20 years. It was suggested to trace back the actual reasoning behind the introduction of the definitions, and by which occasion changes were made when carrying over to subsequent standards.

It was asked at which point in the scheduling of the buffer the removal delay offset is actually used.

Further study on this issue is recommended.

Agreed to include this in JVET-AB1004, but clearly point out that it is not yet confirmed to require a change. This would apply to AVC, HEVC and VVC.

# Plenary meetings, joint meetings, BoG reports, and liaison communications

## JVET plenaries

No intermediate plenaries were held, as document review and decisions were made in single-track mode at this meeting (with some BoG activity as noted). Further detail on scheduling is recorded in section ‎2.13.

General plenary wrap-up discussions are recorded under sections 8, 9, 10, and 11.

It was suggested to identify and propose editors of standards published in 2021 to receive Excellence Awards from ISO/IEC.

## Information sharing meetings

Information sharing sessions with other WGs and AGs of the MPEG community were held on Monday 24 Oct. 0900–1130, Wednesday 26 Oct. 0900–1000, and Friday 28 Oct. 1300–1700 UTC. The status of the work in the MPEG WGs and AGs was reviewed at these information sharing sessions.

## Joint meeting with Q6/16, WG 2 and WG 4 on Tuesday 25 Oct. 11:15-12:30 on Video Coding for Machines

(notes taken by Q6/16 rapporteur)

~180+90+2 people (some dual)

1. Use cases - CfE (feature coding) & CfP (video coding): Two tracks
2. Tests were done on images and video
3. Summarization of results
4. Candidate architectures
5. What & when to standardize
6. Potential joint interest and next steps

Specific to the video coding track:

end-to-end, block-based, or a combination

Stages: pre-analysis, preproc / pre-transformation (e.g. repacking, resolution resampling, possibly spatially varying, restoration filtering), side-channel data, specialized encode (possibly with ordinary block-based coding syntax and semantics and decoding process), ordinary block-based or specialized decode, post-proc / inverse transformation.

gains ~50%

“normative”

possibly two conformance points (one for core decoding process, another for what happens after that)

V-PCC and MIV are analogous

lossless coding trick for input to conventional coder - may violate bit rate constraint of the core coder

Igor Curcio

Lu Yu

Gary Sullivan

Jens-Rainer Ohm for WG 5

Yuan Zhang onsite

JO AG 2

(block diagram)

No conclusion yet about basic architecture choice.

Choice would need to be made based on tradeoffs.

Gauge interest from VCEG.

In C139, it was suggested that VCEG consider the following actions:

* study the proposed VVC-based technologies that are responded to the CfP of VCM;
* follow the evaluation of the responses to the CfP of VCM and compare VVC-based technologies with non-VVC-based technologies;
* evaluate and determine whether VCEG is interested in collaborating with MPEG to jointly develop a VCM standard.

Therefore, it is suggested for VCEG to seek further information from MPEG regarding the responses to the CfP of the VCM, in particular, the VVC-based technologies. It is suggested that a joint meeting with MPEG be scheduled during the 140th MPEG meeting.

Comment re potential architecture dependency.

Comment re dependency on what will be done with the decoded result.

Metadata could be part of MPEG-7, for example.

Post-proc could be part of the “consuming machine”.

Consuming machine could receive intermediate data - e.g. MVs, modes, segmentations

Suggested that a tech WG should make the architecture decision.

Comment on value of generalization for different machine tasks. e.g. by having ordinary video input to the receiving machine.

resilience to different machine-based tasks on the receiving side.

Some think can do well with just encoder non-normative optimization and preprocessing of a conventional core codec - which could be the subject of a TR on how to do that.

Architecture choice could be affected by cost-benefit tradeoffs - e.g. more effectiveness for some architectures.

Certainly, ordinary video will be a common input to machine vision analysis.

Possible multi-track consideration:

* Machine analysis (and optimization for) of ordinarily decoded video (e.g. subject of a TR)
* Improved capability for some alternative architecture
  + “wrapper” technology around an ordinary core technology
  + modified core architecture

Real-time and non-realtime usage.

Comment: “Track 2” would produce watchable video

Suggestion:

* Plan a TR on “non-normative” approach (like TR on HDR)? JVET to work on this (AHG, …).
* In parallel further study of other approaches in MPEG with exchange of information.

## BoGs (0)

One break-out group was established at this meeting to conduct discussion and develop recommendations on neural network-based video coding. See section 5.2.1 and the report JVET-AA0247 for details.

## Liaison communications

The JVET received a liaison letter from SC 29/WG 1 JPEG ([m60532](https://dms.mpeg.expert/doc_end_user/current_document.php?id=83778&id_meeting=191)) related to its JPEG AI project. The evaluation of responses to the JPEG AI Call for Proposals was to take place at the upcoming JPEG meeting. The intended overall timeline for the standardization process was as follows: 2022-10 WD, 2023-04 CD, 2023-10 DIS, 2023-10 IS. (Additional information on the state of the JPEG AI project was provided in JVET-AA0047 by key JPEG AI experts.) Initial text for a reply was drafted in session 24 and finalized in the closing sessions, and issued as document WG 5 N 152.

[m60701](https://dms.mpeg.expert/doc_end_user/current_document.php?id=84196&id_meeting=192) Liaison statement from SC 29/WG 1 to WG 5 on JPEG AI [WG 1 via SC 29 Secretariat]

TBP

[m61406](https://dms.mpeg.expert/doc_end_user/current_document.php?id=84901&id_meeting=192) Liaison statement from SMPTE to SC 29/WG 4 and 5 regarding Reserving VUI values to signal the SMPTE ST 2128 color space [SMPTE via SC 29 Secretariat]

TBP

Liaison statement to SMPTE on ITP-PQ-C2 colour space

A response was reviewed at 1130 on Thursday 27 Oct. It was suggested to add a paragraph requesting more information from them, also about the current publication status, such that ST 2128 could be referenced

# Project planning

## Software timeline

ECM7 software (including all adoptions) was planned to be available 3 weeks after the meeting.

The NNVC 3.0 codebase software was planned to be available 3 weeks after the meeting.

VTM19.0 software was planned to be available on 2022-11-30. (Note that further updates may be released later)

HM17.0 software was planned to be available on 2022-12-16.

## Core experiment and exploration experiment planning

An EE on neural network-based video coding was established, as recorded in output document JVET-AB2023.

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-AB2024.

Initial versions of these documents were presented and approved in the first plenary session on Friday 22 July.

## Drafting of specification text, encoder algorithm descriptions, and software

The following agreement has been established: the editorial team has the discretion to not integrate recorded adoptions for which the available text is grossly inadequate (and cannot be fixed with a reasonable degree of effort), if such a situation hypothetically arises. In such an event, the text would record the intent expressed by the committee without including a full integration of the available inadequate text.

## Plans for improved efficiency and contribution consideration

The group considered it important to have the full design of proposals documented to enable proper study.

Adoptions need to be based on properly drafted working draft text (on normative elements) and HM/VTM encoder algorithm descriptions – relative to the existing drafts. Proposal contributions should also provide a software implementation (or at least such software should be made available for study and testing by other participants at the meeting, and software must be made available to cross-checkers in EEs).

Suggestions for future meetings included the following generally-supported principles:

* No review of normative contributions without draft specification text
* VTM algorithm description text is strongly encouraged for non-normative contributions
* Early upload deadline to enable substantial study prior to the meeting
* Using a clock timer to ensure efficient proposal presentations (5 min) and discussions

As general guidance, it was suggested to avoid usage of company names in document titles, software modules etc., and not to describe a technology by using a company name.

## General issues for experiments

It was emphasized that those rules which had been set up or refined during the 12th JVET meeting should be observed. In particular, for some CEs of some previous meetings, results were available late, and some changes in the experimental setup had not been sufficiently discussed on the JVET reflector.

Group coordinated experiments have been planned as follows:

* “Core experiments” (CEs) are the coordinated experiments on coding tools which are deemed to be interesting but require more investigation and could potentially become part of a draft standard by the next meeting or in the near future.
* “Exploration experiments” (EEs) are also coordinated experiments. These are conducted on technology which is not foreseen to become part of a draft standard in the near future. The investigating methodology for assessment of such technology can also be an important part of an EE. (Further general rules for EEs, as far as deviating from the CE rules below, should be discussed in a future meeting. For the current meeting, procedures as described in the EE description document are deemed to be sufficient.)
* A CE is a test of a specific fully described technology in a specific agreed way. It is not a forum for thinking of new ideas (like an AHG). The CE coordinators are responsible for making sure that the CE description is complete and correct and has adequate detail. Reflector discussions about CE description clarity and other aspects of CE plans are encouraged.
* A description of each experiment is to be approved at the meeting at which the experiment plan is established. This should include the issues that were raised by other experts when the tool was presented, e.g., interference with other tools, contribution of different elements that are part of a package, etc. The experiment description document should provide the names of individual people, not just company names.
* Software for tools investigated in a CE will be provided in one or more separate branches of the software repository. Each CE will have a “fork” of the software, and within the CE there may be multiple branches established by the CE coordinator. The software coordinator will help coordinate the creation of these forks and branches and their naming. All JVET members will have read access to the CE software branches (using shared read-only credentials as described below).
* During the experiment, revisions of the experiment plans can be made, but not substantial changes to the proposed technology.
* The CE description must match the CE testing that is done. The CE description needs to be revised if there has been some change of plans.
* The CE summary report must describe any changes that were made in the process of finalizing the CE.
* By the next meeting it is expected that at least one independent cross-checker will report a detailed analysis of each proposed feature that has been tested and confirm that the implementation is correct. Commentary on the potential benefits and disadvantages of the proposed technology in cross-checking reports is highly encouraged. Having multiple cross-checking reports is also highly encouraged (especially if the cross-checking involves more than confirmation of correct test results). The reports of cross-checking activities may (and generally should) be integrated into the CE report rather than submitted as separate documents.
* It is mandatory to report encoder optimizations made for the benefit of a tool, and if an equivalent optimization could be applied on the anchor, a comparison against the improved anchor shall be provided.
* A new proposal can be included in a CE based on group decision, regardless if an independent party has already performed a cross-check in the meeting when it was first proposed.

It is possible to define sub-experiments within particular CEs, for example designated as CEX.a, CEX.b, etc., where X is the basic CE number.

As a general rule, it was agreed that each CE should be run under the same testing conditions using one software codebase, which should be based on the group test model software codebase. An experiment is not to be established as a CE unless there is access given to the participants in (any part of) the CE to the software used to perform the experiments.

The general agreed common conditions for single-layer coding efficiency experiments for SDR video are described in the prior output document JVET-T2010.

Experiment descriptions should be written in a way such that it is understood as a JVET output document (written from an objective “third party perspective”, not a proponent perspective – e.g. not referring to methods as “improved”, “optimized”, “enhanced”, etc.). The experiment descriptions should generally not express opinions or suggest conclusions – rather, they should just describe what technology will be tested, how it will be tested, who will participate, etc. Responsibilities for contributions to CE work should identify individuals in addition to company names.

CE descriptions contain a basic description of the technology under test, but should not contain excessively verbose descriptions of a technology (at least not unless the technology is not adequately documented elsewhere). Instead, the CE descriptions should refer to the relevant proposal contributions for any necessary further detail. However, the complete detail of what technology will be tested must be available – either in the CE description itself or in documents that are referenced in the CE description that are also available in the JVET document archive.

Any technology must have at least one cross-check partner to establish a CE – a single proponent is not enough. It is highly desirable have more than just one proponent and one cross-checker.

The CE development workflow is described at:

<https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/wikis/Core-experiment-development-workflow>

CE read access is available using shared accounts: One account exists for MPEG members, which uses the usual MPEG account data. A second account exists for VCEG members with account information available in the TIES informal ftp area (IFA) system at:

<https://www.itu.int/ifa/t/2017/sg16/exchange/wp3/q06/vceg_account.txt>

Some agreements relating to CE activities were established as follows:

* Only qualified JVET members can participate in a CE.
* Participation in a CE is possible without a commitment of submitting an input document to the next meeting. Participation is requested by contacting the CE coordinator.
* All software, results, and documents produced in the CE should be announced and made available to JVET in a timely manner.
* A JVET CE reflector will be established and announced on the main JVET reflector. Discussion of logistics arrangements, exchange of data, minor refinement of the test plans, and preparation of documents shall be conducted on the JVET CE reflector, with subject lines prefixed by “[CEx: ]”, where “x” is the number of the CE. All substantial communications about a CE other than such details shall take place on main JVET reflector. In the case that large amounts of data are to be distributed, it is recommended to send a link to the data rather than the data itself, or upload the data as an input contribution to the next meeting.

General timeline for CEs

T1= 3 weeks after the JVET meeting: To revise the CE description and refine questions to be answered. Questions should be discussed and agreed on JVET reflector. Any changes of planned tests after this time need to be announced and discussed on the JVET reflector. Initially assigned description numbers shall not be changed later. If a test is skipped, it is to be marked as “withdrawn”.

T2 = Test model software release + 2 weeks: Integration of all tools into a separate CE branch of the VTM is completed and announced to JVET reflector.

* Initial study by cross-checkers can begin.
* Proponents may continue to modify the software in this branch until T3.
* 3rd parties are encouraged to study and make contributions to the next meeting with proposed changes

T3: 3 weeks before the next JVET meeting or T2 + 1 week, whichever is later: Any changes to the CE test branches of the software must be frozen, so the cross-checkers can know exactly what they are cross-checking. A software version tag should be created at this time. The name of the cross-checkers and list of specific tests for each tool under study in the CE plan description shall be documented in an updated CE description by this time.

T4: Regular document deadline minus 1 week: CE contribution documents including specification text and complete test results shall be uploaded to the JVET document repository (particularly for proposals targeting to be promoted to the draft standard at the next meeting).

The CE summary reports shall be available by the regular contribution deadline. This shall include documentation about crosscheck of software, matching of CE description and confirmation of the appropriateness of the text change, as well as sufficient crosscheck results to create evidence about correctness (crosscheckers must send this information to the CE coordinator at least 3 days ahead of the document deadline). Furthermore, any deviations from the timelines above shall be documented. The numbers used in the summary report shall not be changed relative to the description document.

CE reports may contain additional information about tests of straightforward combinations of the identified technologies. Such supplemental testing needs to be clearly identified in the report if it was not part of the CE plan.

New branches may be created which combine two or more tools included in the CE document or the VTM (as applicable).

It is not necessary to formally name cross-checkers in the initial version of the CE description document. To adopt a proposed feature at the next meeting, JVET would like to see comprehensive cross-checking done, with analysis of whether the description matches the software, and a recommendation of the value of the tool and given tradeoffs.

The establishment of a CE does not indicate that a proposed technology is mature for adoption or that the testing conducted in the CE is fully adequate for assessing the merits of the technology, and a favourable outcome of CE does not indicate a need for adoption of the technology into a standard or test model.

Availability of specification text is important to have a detailed understanding of the technology and also to judge what its impact on the complexity of the specification will be. There must also be sufficient time to study this in detail. CE contributions without sufficiently mature draft specification text in the CE input document should not be considered for adoption.

Lists of participants in CE documents should be pruned to include only the active participants. Read access to software will be available to all members.

# Establishment of ad hoc groups

The ad hoc groups established to progress work on particular subject areas until the next meeting are described in the table below. The discussion list for all of these ad hoc groups was agreed to be the main JVET reflector ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de)). The previously approved rules for MPEG ad hoc groups established in document [SC29/AG2 N 46](https://www.mpegstandards.org/wp-content/uploads/2022/01/ISO-IECJTC1-SC29-AG2_N0046_AhG.pdf) were agreed to apply to these ad hoc groups.

Review of AHG plans was conducted during the plenary on Thursday 27 October 2022 at 0510 UTC.

|  |  |  |
| --- | --- | --- |
| **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-AB1008, JVET-AB2002, JVET-AB2006, and JVET-AB2027). * Collect reports of errata for the VVC, VSEI, HEVC, AVC, CICP, the codepoint usage TR specification and the published HDR-related technical reports and produce the JVET-AB1004 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, J. Boyce, A. Browne, S. Kim, S. Liu, J.‑R. Ohm, G. J. Sullivan, A. Tourapis, Y.-K. Wang, Y. Ye (vice-chairs) | N |
| **Test model software development (AHG3)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Coordinate development of test models (VTM, HM, SCM, SHM, HTM, MFC, MFCD, JM, JSVM, JMVM, 3DV-ATM, 360Lib, and HDRTools) software and associated configuration files. * Produce documentation of software usage for distribution with the software. * Enable software support for recently standardized additional SEI messages. * Discuss and make recommendations on the software development process. * Propose improvements to the guideline document for developments of the test model software. * Perform comparative tests of test model behaviour using common test conditions. * 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. * Update the CTC document JVET-AB2010, and prepare drafts of merged 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 establishing a test plan for VVC scalability features by the next meeting. * 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. * 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 remote expert viewing for purposes of subjective quality evaluation. * Coordinate with AG 5 in studying and developing further methods of subjective quality evaluation, e.g. based on crowd sourcing. * Prepare availability of viewing equipment and facilities arrangements for future meetings. | V. Baroncini, T. Suzuki, M. Wien (co-chairs), S. Liu, S. Puri, A. Segall, P. Topiwala, S. Wenger, J. Xu, Y. Ye (vice-chairs) | Y (tel., 2 weeks notice) |
| **Conformance testing (AHG5)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the draft conformance testing for operation range extensions JVET-Y2026 and the additional conformance bitstreams for VVC multilayer configurations JVET-AB2028, and investigate the need for improvements. * Study the requirements of VVC, HEVC, and AVC conformance testing to ensure interoperability. * Maintain and update the conformance bitstream database, and contribute to report problems, and suggest actions to resolve these. * Study additional testing methodologies to fulfil the needs for VVC conformance testing. | D. Rusanovskyy, I. Moccagatta (co-chairs), F. Bossen, K. Kawamura, T. Ikai, H.-J. Jhu, K. Sühring, Y. Yu (vice-chairs) | N |
| **ECM software development (AHG6)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Coordinate development of the ECM software and associated configuration files. * Produce documentation of software usage for distribution with the software. * Prepare and deliver ECM-7.0 software version and the reference configuration encodings according to the ECMcommon test conditions. * Investigate encoder speedup and other encoder software optimization. * Coordinate with ECM algorithm description editors to identify any mismatches between software and text, make further updates and cleanups to the software as appropriate. | V. Seregin (chair), J. Chen, F. Le Léannec, K. Zhang (vice-chairs) | Y (tel., 2 weeks notice) |
| **Low latency and constrained complexity (AHG7)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Identify additional application scenarios and their requirements for low latency and constrained complexity, taking into account aspects of real-time encoding and decoding. * Discuss requirements of already identified scenario such as cloud gaming, game casting, video conferencing, video surveillance and remote control of systems. * Evaluate and refine new CTC for low latency and constrained complexity application scenarios, including the option of enabling GDR according to requirements of gaming applications. * Investigate a set of tools that provide a reasonable tradeoff regarding complexity vs. compression, as well as latency constraints. * Conduct tests with ECM and VTM to determine the impact of discussed configurations on coding efficiency and run time. * Refine the set of test sequences for gaming applications, and if necessary collect new test materials that are suitable for the intended application domains, and establish an applicable dataset in coordination with AHG4. * Coordinate with AHG3 and AHG12 to discuss and recommend configuration(s) applicable to ECM and VTM, taking into account complementarity with existing CTCs. | A. Duenas, T. Poirier, S. Liu (co-chairs), L. Wang, J. Xu (vice-chairs) | N |
| **High bit depth, high bit rate, and high frame rate coding (AHG8)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the benefits and characteristics of VVC coding tools for high bit depth, high bit rate, and high frame rate coding. * Produce test results with JVET-AA2018 testing conditions for high bit depth and high bit rate coding in coordination with AHG3. * Contribute to the development of software and conformance testing for operation range extensions in coordination with AHG3 and AHG5. | A. Browne, T. Ikai (co-chairs), D. Rusanovskyy, X. Xiu, Y. Yu (vice-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. * 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. * Coordinate with AHG3 for software support of SEI messages. | S. McCarthy, Y.-K. Wang (co-chairs), T. Chujoh, S. Deshpande, C. Fogg, 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 pre processing 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. * Consider neural network-based encoding optimization technologies for video coding standards. * Investigate other methods of improving objective and/or subjective quality, including adaptive coding structures and multi-pass encoding. * Study methods of rate control and rate-distortion optimization and their impact on performance, subjective and objective quality. * Study the potential of defining default or alternate software configuration settings and test conditions optimized for either subjective quality, or higher objective quality, and coordinate such efforts with AHG3 and AHG6. * Study the effect of varying configuration parameters depending on temporal layer, such as those related to deblocking, partitioning, chroma QP. | P. de Lagrange, A. Duenas, R. Sjöberg, A. Tourapis (AHG chairs) | N |
| **Neural network-based video coding (AHG11)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Evaluate and quantify the performance improvement potential of NN-based video coding technologies compared to existing video coding standards such as VVC, including both individual coding tools and novel architectures. * Refine the test conditions for NN-based video coding in JVET-AB2016. Generate and distribute anchor encoding, and develop supporting software as needed. * Study the impact of training (including the impact of loss functions) on the performance of candidate technologies, and identify suitable material for testing and training. * Analyse complexity characteristics, perform complexity analysis, and develop complexity reductions of candidate technology. * Finalize and discuss the EE on neural network-based video coding. * Coordinate with other relevant groups, including SC29/AG5 on the evaluation and assessment of visual quality and AHG12 on the interaction with ECM coding tools. In the case of conducting subjective viewing, provide support to AHG4 for the timely availability of test materials and test subjects. * Coordinate with AHG14 on items related to NNVC software development. | E. Alshina, S. Liu, A. Segall (co‑chairs), F. Galpin, J. Li, T. Shao, H. Wang, Z. Wang, M. Wien, P. Wu (vice‑chairs) | Y (tel., 2 weeks notice) |
| **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 ECM7 algorithm description JVET-AB2025. * 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-Y2017. * Analyse the results of exploration experiments described in JVET-AB2024 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, G. Li, X. Li, K. Naser, H. Yang (vice-chairs) | Y (tel., 2 weeks notice) |
| **Film grain technologies (AHG13)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the benefits and characteristics of film grain technologies, including autoregressive and frequency-filtering technologies. * Discuss the JVET-AB2020 draft of the Technical Report on Film grain synthesis technology for video applications and suggest improvements as necessary. * Study alternative film grain models and their associated documentation. * Study preprocessing and encoder technologies for determining values for FGC (Film Grain Characteristics) SEI message syntax elements. * Identify potential need for additional film grain technology and signalling, if needed. * Coordinate development of film grain technology software and configuration files. * Coordinate with AG 5 on finalizing the draft plan for subjective quality testing of the FGC SEI message JVET-AB2022, and conduct preparations for such testing. * Coordinate with AHG3 for software support of the FGC SEI message. | W. Husak, M. Radosavljević, W. Wan (co-chairs), D. Grois, Y. He, P. de Lagrange, A. Segall, 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-3.0 software version and the reference configuration encodings according to the NNVC common test conditions JVET-AB2016. Prepare and release anchor data for all configurations of the software. * 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-AB2019) 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. | S. Eadie, F. Galpin, Y. Li, L. Wang, Z. Xie (co-chairs) |  |
| **Optimization of encoders and receiving systems for machine analysis of coded video content (AHG15)**  ([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. * Define test conditions and generate anchors. * Define evaluation framework, procedure and methodologies. * Coordinate software development and experiments. * Evaluate proposed technologies and their suitability for machine analysis applications. * Prepare a draft technical report on optimization of encoders and receiving systems for machine analysis of coded video content. * Coordinate with MPEG VCM AHG on common interests and activities such as test materials, studying characteristics and requirements of targeted machine analysis tasks, etc. | C. Hollmann, S. Liu, S. Wang, M. Zhou (co-chairs) |  |

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 170](https://dms.mpeg.expert/doc_end_user/current_document.php?id=82006&id_meeting=189)) 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 151](https://dms.mpeg.expert/doc_end_user/current_document.php?id=82006&id_meeting=189), as noted in section 9.

[JVET-AB1000](https://jvet-experts.org/doc_end_user/current_document.php?id=11942) Meeting Report of the 28th JVET Meeting [J.-R. Ohm] [WG 5 N 156] (2022-11-25)

Initial versions of the meeting notes (d0 … d9) were made available on a daily basis during the meeting.

Remains valid – not updated: [JCTVC-H1001](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=5095) HEVC software guidelines [K. Sühring, D. Flynn, F. Bossen (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-Z1003](https://jvet-experts.org/doc_end_user/current_document.php?id=11705) Coding-independent code points for video signal type identification (Draft 1 of 3rd edition) [WG 5 DIS N 162] [G. J. Sullivan, A. Tourapis] (2022-11-18)

The drafted revision was issued without changes as WG 5 N 162 Text of ISO/IEC DIS 23091-2:202x Coding-independent code points – Part 2: Video (3rd edition), with due date 2022-11-09. As no changes were suggested in the CD ballot, a DoCR was not issued.

[JVET-AB1004](https://jvet-experts.org/doc_end_user/current_document.php?id=11943) Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR [B. Bross, I. Moccagatta, C. Rosewarne, G. J. Sullivan, Y. Syed, Y.-K. Wang] (2022-12-17, near next meeting)

New aspects included the issues pointed out in JVET-AB0223 on the Timing / DU information SEI message in HEVC and VVC. Furthermore, the change discussed in context of JVET-AB0267 (cropped width and height) should be included.

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] (2022-05-06)

Remains valid – not updated [JVET-T1006](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10538) Annotated regions and shutter interval information SEI messages for AVC (Draft 2) [J. Boyce, S. McCarthy, Y.-K. Wang]

(This can be removed after publication of the new edition of ISO/IEC 14496-10.)

Remains valid – not updated: [JCTVC-V1007](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=10312) SHVC Test Model 11 (SHM 11) Introduction and Encoder Description [G. Barroux, J. Boyce, J. Chen, M. M. Hannuksela, Y. Ye] [WG 11 N 15778]

[JVET](https://jvet-experts.org/doc_end_user/current_document.php?id=11708)-AB1008 Additional colour type identifiers for AVC, HEVC and Video CICP (Draft 2) [G. J. Sullivan, W. Husak, A. Tourapis] (2022-11-25) [WG 5 Preliminary WD N163]

This includes identifiers for ITP-PQ-C2 colour representation proposed in JVET-AB0172. It will be assigned the new codepoint number 17, to be integrated with the codepoint tables currently in JVET-Z1008 for AVC and HEVC (currently for YCoCg-R), and another table for Video CICP (only ITP-PQ-C2).

Remains valid – not updated: [JCTVC-X1009](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=10572) Common Test Conditions for SHVC [V. Seregin, Y. He]

Remains valid – not updated [JCTVC-O1010](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=8511) Guidelines for Conformance Testing Bitstream Preparation [T. Suzuki, W. Wan]

Remains valid – not updated [JVET-AA1011](https://jvet-experts.org/doc_end_user/current_document.php?id=11944) HEVC multiview profiles supporting extended bit depth (draft 1) [A. Tourapis, W. Husak] [WG 5 preliminary WD N 143] (2022-08-26)

This output arises from JVET-AA0239.

Proponents of the new profiles were requested to develop conformance bitstreams, and contribute to HTM software changes if necessary.

JVET-T1012 Overview of IT systems used in JVET [J. Ohm, I. Moccagatta, K. Sühring, M. Wien] (2022-12-23)

No output: JVET-T1013

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 for HM – not updated: [JCTVC-Z1015](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=10689) Common Test Conditions for Screen Content Coding [H. Yu, R. Cohen, K. Rapaka, J. Xu]

No output: JVET-X1016 through JVET-X1019

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] (2022-08-19)

This specifies only the CTC for non-4:2:0 colour formats. The corresponding document for VVC is JVET-T2013, with no unification yet.

Remains valid – not updated: [JVET-T2001](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10540) Versatile Video Coding Draft 10 [B. Bross, J. Chen, S. Liu, Y.-K. Wang]

[JVET-AB2002](https://jvet-experts.org/doc_end_user/current_document.php?id=11709) Algorithm description for Versatile Video Coding and Test Model 18 (VTM 18) [A. Browne, Y. Ye, S. Kim] [WG 5 N 161] (2022-01-06, near next meeting)

Updates include descriptions of new SEI message processing, in particular for the SEI processing order SEI), Green metadata, and potentially some other SEI messages as listed under JVET-AB0269. Further non-normative modifications adopted for VTM are the contributions on RPR encoder, and upsanpling filters (JVET-AB0080 and JVET-AB0081).

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-N1003](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=6638) Guidelines for VVC reference software development [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-AA2005](https://jvet-experts.org/doc_end_user/current_document.php?id=11946) New level and systems-related supplemental enhancement information for VVC (Draft 3) [B. Bross, E. François, A. Tourapis, Y.-K. Wang] [WG 5 DAM 1 N 145] (2022-07-29)

[JVET-AB2006](https://jvet-experts.org/doc_end_user/current_document.php?id=11947) Additional SEI messages for VSEI (Draft 2) [S. McCarthy, T. Chujoh, M. M. Hannuksela, G. J. Sullivan, Y.-K. Wang] [WG 5 DIS N 158] (2022-11-09)

A DoCR for ballot responses on CDAM 1 (WG 5 N 157) was reviewed Thursday 27 October at 0920.

The aspect of having several NNPFC messages activated in one picture was discussed (comment #056). It was agreed to allow this, as it is the choice of the decoding device to operate those which are fitting the purpose of the local output or its capabilities. Examples are SEI messages with different purposes, or SEI messages with the same purpose and different degrees of complexity.

It was agreed to convert this into the FDIS of a new edition after DAM ballot (comment #70). Due to short editing period, it is not practical to do this at the current stage.

Elements from (list from the adoptions in 6.1and 6.2) are added.

Remains valid – not updated: [JVET-S2007](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9679) Versatile supplemental enhancement information messages for coded video bitstreams Draft 5 [J. Boyce, V. Drugeon, G. J. Sullivan, Y.-K. Wang]

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]

This had been issued as ISO/IEC FDIS 23090-16 as WG 5 [N 112](https://dms.mpeg.expert/doc_end_user/current_document.php?id=82000&id_meeting=189), and submitted for ITU-T consent.

[JVET-AB2010](https://jvet-experts.org/doc_end_user/current_document.php?id=11471) 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] (2022-11-11)

Remains valid – not updated: [JVET-Z2011](https://jvet-experts.org/doc_end_user/current_document.php?id=11712) VTM and HM common test conditions and evaluation procedures for HDR/WCG video [A. Segall, E. François, W. Husak, S. Iwamura, D. Rusanovskyy] (2022-05-13)

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] (2021-03-31)

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-AB2016](https://jvet-experts.org/doc_end_user/current_document.php?id=11948) Common Test Conditions and evaluation procedures for neural network-based video coding technology [E. Alshina, R.-L. Liao, S. Liu, A. Segall] (2022-11-11)

Update in the methodology of kMAC/pixel computation (recommendation from JVET-AB0023).

Remains valid – not updated: [JVET-Y2017](https://jvet-experts.org/doc_end_user/current_document.php?id=11473) Common Test Conditions and evaluation procedures for enhanced compression tool testing [M. Karczewicz and Y. Ye]

Only a change of config file for JVET-AA0098 was noted (and probably some other CTC related decision noted elsewhere). It was further decided to enable palette mode for classes F and TGM in the CTC both for VTM anchor and ECM. As these affected only the config files rather than the document, no need to produce a revised document was identified.

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] (2022-08-19)

JVET-AB2019 Description of algorithms and software in neural network-based video coding (NNVC) [Y. Li, H. Wang, L. Wang, F. Galpin, J. Ström] [WG 5 N 165] (2022-12-02)

Developed from JVET-AB0183 (see notes there about what else to added).

[JVET-AB2020](https://jvet-experts.org/doc_end_user/current_document.php?id=11950) Film grain synthesis technology for video applications (Draft 3) [D. Grois, Y. He, W. Husak, P. de Lagrange, A. Norkin, M. Radosavljević, A. Tourapis, W. Wan] [WG 5 [N 159](https://dms.mpeg.expert/doc_end_user/current_document.php?id=82207&id_meeting=189)] (2022-09-30)

Developed from JVET-AB0042.

JVET-AB2021 Draft verification test plan for VVC multilayer coding [S. Iwamura, P. de Lagrange, M. Wien] (2022-12-16)

A draft of this document was presented in JVET at 1610 on Thursday 27 Oct. This document had been developed in the joint BoG with AG 5 (meeting at 1600-1730 on Wednesday 26 Oct., no separate BoG report, as the document reflects the recommendations of the BoG).

JVET-AB2022 Draft plan for subjective quality testing of FGC SEI message [P. de Lagrange, W. Husak, M. Radosavljević, M. Wien] (2022-12-16)

A draft of this document was presented in JVET at 1630 on Thursday 27 Oct. This document had been developed in the joint BoG with AG 5 (meeting at 1600-1730 on Wednesday 26 Oct., no separate BoG report, as the document reflects the recommendations of the BoG).

It was asked if also tests are planned using HM. This aspect is left open for now; it might be done if it is demonstrated that such an additional comparison would have additional benefit in terms of what the test is intending to study. Is it about showing the advantages of the SEI message (which could be used for both VVC and HEVC), or could it also demonstrate that VVC has more advantage than HEVC if the SEI message is used.

It was commented that testing to quality of the synthesized film grain is very important (goal 2 of the plan).

(it was mentioned that a merge request on the SEI message implementation on top of HM was already made, and it is waiting for release of a new version)

[JVET-AB2023](https://jvet-experts.org/doc_end_user/current_document.php?id=11940) Exploration Experiment on Neural Network-based Video Coding (EE1) [E. Alshina, F. Galpin, Y. Li, M. Santamaria, J. Ström, H. Wang, L. Wang, Z. Xie] [WG 5 N 164] (2022-11-11)

An initial draft of this document was reviewed and approved at 0905-0945 on Friday 28 October.

Categories are loop filters, post filters, superresolution upsampling, intra coding (AE integerization), and intra prediction.

It was suggested that it would be interesting to perform subjective quality investigation with some of the super-resolution proposals in a future meeting (provided that close enough rate/quality matching with anchor is achieved).

[JVET-AB2024](https://jvet-experts.org/doc_end_user/current_document.php?id=11941) Exploration Experiment on Enhanced Compression beyond VVC capability (EE2) [V. Seregin, J. Chen, G. Li, K. Naser, J. Ström, F. Wang, M. Winken, X. Xiu, K. Zhang] [WG 5 N 166] (2022-11-25)

An initial draft of this document was reviewed and approved at 0945-1015 on Friday 28 October. JVET-AB0150 was added as test 2.7b. Clarification is necessary about the training set to be used in test 4.2b (referred to as “LFNST set”).

Categories are intra prediction, inter prediction, screen content coding, transforms, and in-loop filters.

[JVET-AB2025](https://jvet-experts.org/doc_end_user/current_document.php?id=11951) Algorithm description of Enhanced Compression Model 7 (ECM 7) [M. Coban, F. Le Léannec, R.-L. Liao, K. Naser, J. Ström, L. Zhang] [WG 5 N 167] (2022-12-31)

New elements from notes elsewhere in this report:

* … to be added …

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-Y2026](https://jvet-experts.org/doc_end_user/current_document.php?id=11477) Conformance testing for VVC operation range extensions (Draft 3) [WG 5 DAM [N 110](https://dms.mpeg.expert/doc_end_user/current_document.php?id=81998&id_meeting=189)] [D. Rusanovskyy, T. Hashimoto, H.-J. Jhu, I. Moccagatta, Y. Yu]

Editorial improvements are recorded in JVET-AA0109, no need for a new output at the current meeting was identified.

[JVET-AB2027](https://jvet-experts.org/doc_end_user/current_document.php?id=11952) SEI processing order SEI message in VVC (draft 2) [S. McCarthy, M. Hannuksela, Y.-K. Wang] [WG 5 preliminary WD 2 N 160] (2022-11-18)

New version (see discussion under JVET-AB0051 and JVET-AB0069)

JVET-AB2028 Additional conformance bitstreams for VVC multilayer configurations [S. Iwamura, I. Moccagatta] (2022-12-31)

Developed from JVET-AB0085.

JVET-AB2029 Visual quality comparison of ECM/VTM encoding [V. Baroncini, J.-R. Ohm, M. Wien] [AG 5 N75] (2022-12-09)

Initial version was reviewed Friday 28 October 1030.

# 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 SG 16 auspices when it meets (ordinarily starting meetings on the Wednesday of the first week and closing it on the Wednesday of the second week of the SG 16 meeting – a total of 8 meeting days), and
* Otherwise meeting under ISO/IEC JTC 1/‌SC 29 auspices when its MPEG WGs meet (ordinarily starting meetings on the Friday prior to the main week of such meetings and closing it on the same day as other MPEG WGs – a total of 8 meeting days).

In cases where an exceptionally high workload is expected for a meeting, an earlier starting date may be defined. In cases of online meetings, no sessions should be held on weekend days. This may imply an earlier starting date as well.

Some specific future meeting plans (to be confirmed) were established as follows:

* Wed. 11 – Fri. 13 and Mon. 16 – Fri. 20 January 2023, 29th meeting under ISO/IEC JTC 1/‌SC 29 auspices, to be held as teleconference meeting.
* During Fri. 21 – Fri. 28 April 2023, 30th meeting under ISO/IEC JTC 1/‌SC 29 auspices in Antalya, TR
* During Fri. 14 – Fri. 21 July 2023, 31st meeting under ITU-T SG16 auspices in Geneva, CH
* During October 2023, 32nd meeting under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.
* During January 2024, 33rd meeting under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.
* During April 2024, 34th meeting under ITU-T SG16 auspices, date and location t.b.d.
* During Fri. 12 – Fri. 19 July 2024, 35th meeting under ISO/IEC JTC 1/‌SC 29 auspices in Sapporo, JP
* During October 2024, 36th meeting under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.

The agreed document deadline for the 29th JVET meeting was planned to be Wednesday 4 January 2023.

Kenzler Conference Management and the entire team were thanked for the excellent hosting and organization of the 28th meeting of the JVET. Qualcomm was thanked for financially supporting the social event. Vittorio Baroncini and Mathias Wien were thanked for conducting subjective quality tests related to the ECM. VABTech was thanked for providing resources to run tests with non-expert viewers. HHI was thanked for providing equipment used in the experts viewing in Mainz. The following companies were thanked for preparing encoded bitstreams: Alibaba, Bytedance, Ericsson, InterDigital, Qualcomm, and Xiaomi. Experts who volunteered to participate in the viewing were also thanked.

It was suggested that in the next meeting, perspectives should be discussed for the ongoing JVET explorations, in terms of potentially developing standardization projects and realistic timelines. In particular, from the current status of the exploration on enhanced compression capability beyond current VVC profiles, evidence has been shown about visual quality improvement. Market needs and related requirements could be considered at the level of the parent bodies.

The 28th JVET meeting was closed at approximately XXXX hours on Friday 28 October 2022.

# Annex A to JVET report: List of documents

# Annex B1 to JVET report: List of meeting participants attending in person

The personally present participants of the twenty-eighth meeting of the JVET, according to an attendance sheet circulated during the meeting sessions (approximately XXX people in total), were as follows:

1. …

# Annex B2 to JVET report: List of meeting participants attending remotely

The remote participants of the twenty-eighth meeting of the JVET, according to the participation records from the Zoom teleconferencing tool used for the meeting sessions (approximately XXX people in total, not including those who attended only the joint sessions with other groups), were as follows:

1. …

# Annex C to JVET report: Recommendations of the 9th meeting of ISO/IEC JTC 1/SC 29/WG 5 MPEG Joint Video Coding Team(s) with ITU-T SG 16

**ISO/IEC JTC 1/SC 29/WG 5 N 155**

**Recommendations of the 9th WG 5 meeting**