|  |  |
| --- | --- |
| **Joint Video Experts Team (JVET)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29**  29th Meeting, by teleconference, 11–20 January 2023 | Document: JVET-AC\_Notes\_d9 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Meeting Report of the 29th Meeting of the Joint Video Experts Team (JVET), by teleconference, 11–20 January 2023** | | |
| *Status:* | Report document from the chair of JVET | | |
| *Purpose:* | Report | | |
| *Author(s) or Contact(s):* | **Jens-Rainer Ohm** Institute of Communication Engineering RWTH Aachen Melatener Straße 23 D-52074 Aachen | Tel: Email: | +49 241 80 27671 [ohm@ient.rwth-aachen.de](mailto:ohm@ient.rwth-aachen.de) |
| *Source:* | Chair of JVET | | |

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

# Summary

The Joint Video Experts Team (JVET) of ITU-T WP3/16 and ISO/IEC JTC 1/‌SC 29 held its twenty-ninth meeting during 11–20 January 2023 as an online-only meeting. For ISO/IEC purposes, JVET is alternatively designated ISO/IEC JTC 1/‌SC 29/‌WG 5, and this was the tenth 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 1300 hours UTC on Wednesday 11 January 2023. Meeting sessions were held on all days except the weekend days of Saturday and Sunday 14 and 15 January 2023, until the meeting was closed at approximately 00XX hours UTC on Saturday 21 January 2023. Approximately XXX people attended the JVET meeting, and approximately XXX input documents (not counting crosschecks), 15 AHG reports, 2 EE summary reports, 2 BoG reports, and 1incoming liaison document 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-eighth JVET meeting in producing the following documents:

* JVET-AB1004 Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR
* JVET-AB1008 Additional colour type identifiers for AVC, HEVC and Video CICP (Draft 2), also issued as WG 5 preliminary WD
* JVET-AB1012 Overview of IT systems used in JVET
* JVET-AB2002 Algorithm description for Versatile Video Coding and Test Model 18 (VTM 18)
* JVET-AB2006 Additional SEI messages for VSEI (Draft 3), also issued as WG 5 DAM
* JVET-AB2010 VTM and HM common test conditions and software reference configurations for SDR 4:2:0 10 bit video
* JVET-AB2016 Common Test Conditions and evaluation procedures for neural network-based video coding technology
* JVET-AB2019 Description of algorithms and software in neural network-based video coding (NNVC)
* JVET-AB2020 Film grain synthesis technology for video applications (Draft 3), also issued as WG 5 WD
* JVET-AB2021 Draft verification test plan for VVC multilayer coding
* JVET-AB2022 Draft plan for subjective quality testing of FGC SEI message
* JVET-AB2023 Exploration Experiment on neural network-based video coding (EE1)
* JVET-AB2024 Exploration Experiment on enhanced compression beyond VVC capability (EE2)
* JVET-AB2025 Algorithm description of Enhanced Compression Model 7 (ECM 7)
* JVET-AB2027 SEI processing order SEI message in VVC (draft 2), also issued as WG 5 preliminary WD
* JVET-AB2028 Additional conformance bitstreams for VVC multilayer configurations
* JVET-AB2029 Visual quality comparison of ECM/VTM encoding

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

* JVET-AB1004 Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR
* JVET-AB1008 Additional colour type identifiers for AVC, HEVC and Video CICP (Draft 2), also issued as WG 5 preliminary WD
* JVET-AB1012 Overview of IT systems used in JVET
* JVET-AB2002 Algorithm description for Versatile Video Coding and Test Model 18 (VTM 18)
* JVET-AB2006 Additional SEI messages for VSEI (Draft 3), also issued as WG 5 DAM
* JVET-AB2010 VTM and HM common test conditions and software reference configurations for SDR 4:2:0 10 bit video
* JVET-AB2016 Common Test Conditions and evaluation procedures for neural network-based video coding technology
* JVET-AB2019 Description of algorithms and software in neural network-based video coding (NNVC)
* JVET-AB2020 Film grain synthesis technology for video applications (Draft 3), also issued as WG 5 WD
* JVET-AB2021 Draft verification test plan for VVC multilayer coding
* JVET-AB2022 Draft plan for subjective quality testing of FGC SEI message
* JVET-AB2023 Exploration Experiment on neural network-based video coding (EE1)
* JVET-AB2024 Exploration Experiment on enhanced compression beyond VVC capability (EE2)
* JVET-AB2025 Algorithm description of Enhanced Compression Model 7 (ECM 7)
* JVET-AB2027 SEI processing order SEI message in VVC (draft 2), also issued as WG 5 preliminary WD
* JVET-AB2028 Additional conformance bitstreams for VVC multilayer configurations
* JVET-AB2029 Visual quality comparison of ECM/VTM encoding

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 ten JVET meetings were planned for 21 – 28 April 2023 under ISO/IEC JTC 1/‌SC 29 auspices, in Antalya, TR; during 11 – 19 July 2023 under ITU-T SG16 auspices in Geneva, CH; during 13 – 20 October 2023 under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.; during 17 – 26 January 2024 under ISO/IEC JTC 1/‌SC 29 auspices, to be conducted as teleconference meeting; during April 2024 under ITU-T SG16 auspices, date and location t.b.d.; during 12 – 19 July 2024 under ISO/IEC JTC 1/‌SC 29 auspices in Sapporo, JP; during October 2024 under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.; during January 2025 under ITU-T SG16 auspices, date and location t.b.d.; during April 2025 under ITU-T SG16 auspices, date and location t.b.d.; and during 11 – 18 July 2025 under ISO/IEC JTC 1/‌SC 29 auspices in Daejeon, KR.

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

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

# Administrative topics

## Organization

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

The Joint Video Experts Team (JVET) of ITU-T WP3/16 and ISO/IEC JTC 1/‌SC 29 held its twenty-ninth meeting during 11–20 January 2023 as an online-only meeting. For ISO/IEC purposes, JVET is alternatively designated ISO/IEC JTC 1/‌SC 29/‌WG 5, and this was the tenth meeting as WG 5. The JVET meeting was held under the chairmanship of Dr Jens-Rainer Ohm (RWTH Aachen/Germany).

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

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

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

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

This report contains three important annexes, as follows:

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

## Meeting logistics

Information regarding logistics arrangements for the meeting had been provided via the email reflector [jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de) and at <http://wftp3.itu.int/av-arch/jvet-site/2023_01_AC_Virtual/>.

## Primary goals

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

* JVET-AB1004 Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR
* JVET-AB1008 Additional colour type identifiers for AVC, HEVC and Video CICP (Draft 2), also issued as WG 5 preliminary WD
* JVET-AB1012 Overview of IT systems used in JVET
* JVET-AB2002 Algorithm description for Versatile Video Coding and Test Model 18 (VTM 18)
* JVET-AB2006 Additional SEI messages for VSEI (Draft 3), also issued as WG 5 DAM
* JVET-AB2010 VTM and HM common test conditions and software reference configurations for SDR 4:2:0 10 bit video
* JVET-AB2016 Common Test Conditions and evaluation procedures for neural network-based video coding technology
* JVET-AB2019 Description of algorithms and software in neural network-based video coding (NNVC)
* JVET-AB2020 Film grain synthesis technology for video applications (Draft 3), also issued as WG 5 WD
* JVET-AB2021 Draft verification test plan for VVC multilayer coding
* JVET-AB2022 Draft plan for subjective quality testing of FGC SEI message
* JVET-AB2023 Exploration Experiment on neural network-based video coding (EE1)
* JVET-AB2024 Exploration Experiment on enhanced compression beyond VVC capability (EE2)
* JVET-AB2025 Algorithm description of Enhanced Compression Model 7 (ECM 7)
* JVET-AB2027 SEI processing order SEI message in VVC (draft 2), also issued as WG 5 preliminary WD
* JVET-AB2028 Additional conformance bitstreams for VVC multilayer configurations
* JVET-AB2029 Visual quality comparison of ECM/VTM encoding

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

## Documents and document handling considerations

### General

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

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

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

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

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

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

### Late and incomplete document considerations

The formal deadline for registering and uploading non-administrative contributions had been announced as Wednesday, 4 January 2023. Any documents uploaded after 1159 hours Paris/Geneva time on Thursday 5 January 2023 were considered “officially late”, with a grace period of 12 hours (to accommodate those living in different time zones of the world). The deadline does not apply to AHG reports and other such reports which can only be produced after the availability of other input documents.

All contribution documents with registration numbers higher than JVET-AC0200 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-AC0XXX (a proposal on …), uploaded 01-XX,
* … .

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

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

* JVET-AC0XXX (a document on …), uploaded 01-XX,
* … .

Most cross-verification reports at this meeting were registered late, and/or uploaded late (except for JVET-AC0117, JVET-AC0133, and JVET-AC0168). 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-AC0102, JVET-AC0157, JVET-AC0289.

The following cross-verification reports were still missing three weeks after the end of the meeting: JVET-AC0XXX, … . They were thus marked by the JVET chair as 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-AC0048, JVET-AC0097, JVET-AC0172, JVET-AC0184, JVET-AC0187, and JVET-AC0199, which were categorized as late in the list above, based on the time of the first reasonable document upload.

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

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

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

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

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

### Outputs of the preceding meeting

All output documents of the previous meeting, particularly the meeting report JVET-AB1000, the Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR JVET-AB1004, the Additional colour type identifiers for AVC, HEVC and Video CICP (Draft 2) JVET-AB1008, the Overview of IT systems used in JVET JVET-AB1012, the Algorithm description for Versatile Video Coding and Test Model 18 (VTM 18) JVET-AB2002, the Additional SEI messages for VSEI (Draft 3) JVET-AB2006, the VTM and HM common test conditions and software reference configurations for SDR 4:2:0 10 bit video JVET-AB2010, the Common test conditions and evaluation procedures for neural network-based video coding technology JVET-AB2016, the Description of algorithms and software in neural network-based video coding (NNVC) JVET-AB2019, the Film grain synthesis technology for video applications (Draft 3) JVET-AB2020, the Draft verification test plan for VVC multilayer coding JVET-AB2021, the Draft plan for subjective quality testing of FGC SEI message JVET-AB2022, the Description of the EE on Neural Network-based Video Coding JVET-AB2023, the Description of the EE on Enhanced Compression beyond VVC capability JVET-AB2024, the Algorithm description of Enhanced Compression Model 7 (ECM 7) JVET-AB2025, the SEI processing order SEI message in VVC (Draft 1) JVET-AB2027, the Additional conformance bitstreams for VVC multilayer configurations JVET-AB2028, and the Visual quality comparison of ECM/VTM encoding JVET-AB2029, had been completed and were approved. The software implementations of HM (version 17.0), VTM (version 19.0), ECM (version 7.0), and NNVC (version 3.0) were also approved.

Only minor editorial issues were found in the meeting report JVET-AB1000; 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 optimization of encoders and receiving systems for machine analysis of coded video content
* Consideration of contributions on errata relating to standards in the domain of JVET
* Consideration of contributions on technical reports relating to standards and exploration study activities in the domain of JVET
* Consideration of contributions providing non-normative guidance relating to standards and exploration study activities in the domain of JVET
* Consideration of information contributions
* Consideration of future work items
* Coordination of visual quality testing
* Liaisons, coordination activities with other organizations
* Review of project editor and liaison assignments
* Approval of output documents and associated editing periods
* Future planning: Determination of next steps, discussion of working methods, communication practices, establishment of coordinated experiments (if any), establishment of AHGs, future meeting planning, other planning issues
* Other business as appropriate for consideration
* Closing

The agenda was approved as suggested.

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

* 1300–1500 1st “afternoon” session [break after 2 hours]
* 1520–1720 2nd “afternoon” session
* [“evening” break – nearly 4 hours]
* 2100–2300 1st “night” session [break after 2 hours]
* 2320–0120+1 2nd “night” session

It was also pointed out that the session times had been changed from meeting to meeting, such that different time zones of the world might be treated approximately equally fairly either in one meeting or another. For the current meeting, the same UTC session times were used as in the 25th JVET meeting (which had been the eighth meeting conducted as an online meeting, whereas the current one is the eleventh).

* 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. The SADL and NNVC packages for neural network-based video coding use the same licensing terms.

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

## Communication practices

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

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

It was emphasized that reflector subscriptions and email sent to the reflector must use real names when subscribing and sending messages and subscribers must respond to inquiries regarding the nature of their interest in the work. The current number of subscribers on the JVET email list was 1184 (as of 10 January 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).

It is further emphasized that the document JVET-AB1012 (to be updated during this meeting) gives valuable hints about communication practices as well as other IT resources used in JVET, such as software, conformance, and test materials.

## Terminology

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

## Opening remarks

Remarks during the opening session of the meeting Wednesday 11 January at 1300 UTC were as follows.

* Timing and organization of the meeting and online access, calendar posting of session plans
* Plans for subsequent F2F meetings in April (Antalya) and July (Geneva)
* 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 14496-10:2022 (Ed. 10), had been forwarded from DIS directly for publication 2022-01-21 (with annotated regions, shutter interval, and miscellaneous corrections) with an editing period, submitted to ITTF in 2022-05 after consultation with ISO staff on format of graphics files, upgraded to “DIS approved for registration” in ISO Project system 2022-07-04, published 2022-11-07
    - Preliminary draft text for YCgCo-Re and YCgCo-Ro issued at 26th meeting, second draft including SMPTE ST 2128 issued at 28th meeting (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, DIS ballot closed 2023-01-10
    - Preliminary draft text for YCgCo-Re and YCgCo-Ro issued at 26th meeting, second draft including SMPTE ST 2128 issued at 28th 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 2018-08-06
      * ISO/IEC 23008-8:2018/AMD 1:2019 Conformance testing for HEVC screen content coding (SCC) extensions and non-intra high throughput profiles, published 2019-10-15
    - Reference software
      * H.265.2 V4 approved 2016-12-22, published 2017-04-10
      * ISO/IEC 23008-5:2017 (Ed. 2) Reference software for high efficiency video coding, published 2017-03-01
      * ISO/IEC 23008-5:2017/AMD 1:2017 Reference software for screen content coding extensions, published 2017-11-09
  + VVC
    - 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 23090-3:2022 (Ed. 2) with operation range extensions, approval at WG level to proceed to FDIS 2022-01-21, published 2022-09-25
    - ISO/IEC 23090-3:202x (Ed. 2) / Amd.1 New level and systems-related supplemental enhancement information, CDAM 1 issued from 26th meeting, ballot closed 2022-07-14, DAM 1 issued from 27th meeting, ballot closed 2023-01-03 (ready for action at this meeting)
    - 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 23090-15:2022 V1 approval at WG level to proceed to FDIS 2022-10-15, upgraded to “DIS approved for registration” in ISO Projects system 2021-10-24, upgraded to “FDIS registered for formal approval” 2022-07-11, FDIS ballot closed 2022-11-04, published 2022-11-24
      * ISO/IEC 23090-15:202x Amd.1 Operation range extensions – DAM 1 issued from 25th meeting 2022-01-21, upgraded to “CD approved for registration as DIS” status in ISO Projects system 2022-05-31, upgraded to “DIS registered” 2022-06-22, DAM ballot closed 2022-11-15 (ready for 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 23090-16:2022 V1 approval at WG level to proceed to FDIS 2022-01-21, upgraded to “DIS approved for registration” status in ISO Projects system 2022-04-21, upgraded to “FDIS registered for formal approval” 2022-04-22, FDIS ballot initiated 2022-07-24, FDIS ballot closed 2022-09-19, published 2022-10-23
  + VSEI
    - 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, published 2022-07-25
    - ISO/IEC 23002-7:2022 (Ed. 2) approval at WG level to proceed to FDIS 2022-01-21, upgraded to “DIS approved for registration” status in ISO Projects system 2022-05-05 and “FDIS registered for formal approval” 2022-05-08, FDIS ballot closed 2022-09-27, published 2022-10-30
    - ISO/IEC 23002-7:202x (2nd Ed.) Amd.1 Request for new edition and CD for additional SEI messages issued at 27th meeting, ballot closed 2022-10-10, DAM registered 2022-11-13, pending DAM ballot
  + 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 2022-10-10, DIS registered 2022-11-13, pending DIS ballot (no action at this meeting)
    - Preliminary draft text for including SMPTE ST 2128 issued at 28th meeting (not yet formally requested as a project)
  + 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, WD 2 issued at 27th meeting, WD 3 issued at 28th meeting
  + The following freely available standards are published here in ISO/IEC:  
    <https://standards.iso.org/ittf/PubliclyAvailableStandards/index.html>
    - Various amendments of ISO/IEC 14496-4:2004
    - Various amendments of ISO/IEC 14496-5:2001
    - ISO/IEC 23008-2:2020 (Ed. 4) HEVC
  + The following standards that have been intended by JVET to be publicly available were not available at <https://standards.iso.org/ittf/PubliclyAvailableStandards/index.html> as of 2023-01-08. (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 available there)
    - ISO/IEC 23008-2:2020 (Ed. 4) Amd.1:2021: Shutter interval information SEI message, published 2021-07-12 (has not been requested)
    - ISO/IEC 23008-5:2017 (Ed. 2) Reference software for high efficiency video coding, published 2017-03-01
    - ISO/IEC 23008-5:2017/AMD 1:2017 Reference software for screen content coding extensions, published 2017-11-09
    - ISO/IEC 23008-8:2018 (Ed. 2) Conformance specification for HEVC, published 2018-08, published 2018-08-06
    - ISO/IEC 23008-8:2018/AMD 1:2019 Conformance testing for HEVC screen content coding (SCC) extensions and non-intra high throughput profiles, published 2019-10-15
    - ISO/IEC 14496-10:2022 (Ed. 10) – AVC – final text issued and public availability requested at the 25th meeting (January 2022)
    - ISO/IEC 23002-7:2022 (Ed. 2) – VSEI – FDIS issued and public availability requested at the 25th meeting (January 2022)
    - ISO/IEC 23090-3:2022 (Ed. 2) – VVC – FDIS issued and public availability requested at the 25th meeting (January 2022)
    - ISO/IEC 23090-15:2022 (Ed. 1) – VVC conformance – FDIS issued and public availability requested at the 24th meeting (October 2021)
    - ISO/IEC 23090-16:2022 (Ed. 1) – 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 was left TBD until the next meeting – perhaps it would be best to compile a list of all relevant software and conformance parts of AVC, HEVC, MPEG-2 aka H.262, CICP, and request these in bulk.
* Draft standards progression status
  + AVC and HEVC colour type indicators for YCgCo-Re, YCgCo-Ro, and SMPTE ST 2128 are drafted and pending formal action
  + HEVC new levels (from JVET-Z1005) – ISO/IEC 23008-2 DIS of new edition of HEVC was issued from 26th meeting, incorporating Amd.1 and corrigenda items (ballot closed 2023-01-10, ballot comments in [m61834](https://dms.mpeg.expert/doc_end_user/current_document.php?id=85619&id_meeting=193)); note that Amd.1 = shutter interval SEI is already included in latest ITU-T edition of H.265. It is noted that there are potential additional items (corrigenda+tickets, YCgCo-Re and YCgCo-Ro draft, SMPTE ST 2128, multiview profiles draft) which might potentially be included based on ballot comments, ITU-T consent is planned for July. It was noted that the referencing of VSEI is also somewhat different [Ed. Clarify] in the ITU-T and ISO/IEC versions of HEVC and AVC, which might be aligned at the next convenient time (basically editorial). Proceeding to FDIS should be considered at the current meeting.
  + VVC new level and systems-related supplemental enhancement information (from JVET-AA2005) – VVC DAM was issued from 27th meeting, ballot closed 2023-01-03, ballot comments in [m61833](https://dms.mpeg.expert/doc_end_user/current_document.php?id=85618&id_meeting=193). The plan is to convert this into FDIS of new edition at the current meeting. ITU-T consent is planned for July 2023 when a new edition of VSEI is also consented, to keep VVC and VSEI aligned.
  + VVC Conformance testing for operation range extensions – (from JVET-Y2026) – the DAM ballot closed 2022-11-15 (ballot comments in [m61832](https://dms.mpeg.expert/doc_end_user/current_document.php?id=85617&id_meeting=193)), and this could progress to FDAM or a new edition at the current meeting. ITU consent is planned for July 2023.
  + VSEI additional SEI messages (from JVET-AB2006) – VSEI DAM (JVET draft 3) was issued from the 28th meeting and was pending a DAM ballot to be issued. To be converted into FDIS of new edition in April (if possible) or July, ITU-T consent in July 2023. No action on this is expected at the current meeting (it is noted that there are numerous input contributions in particular on NNPF, which could be discussed as long as the ballot has not been started yet.).
  + Film grain synthesis technology for video applications (from JVET-AA2020 and JVET-AB0042) – JVET draft 3 was issued at the 28th meeting, and the ISO/IEC 23002-9 CDTR is planned to be issued at the current meeting (a request to start work on the TR had been made at the 25th meeting). The publication limit date is 2023-08-09, so an extension request may be needed, but perhaps not at the current meeting.
  + Video CICP new edition with for YCgCo-Re and YCgCo-Ro (from JVET-Z1003), an ISO/IEC 23091-2 DIS was issued from the 28th meeting and is pending its ballot, so no action on this was expected at the current meeting). ITU-T consent is planned for July 2023.
  + Video CICP colour type indicator for SMPTE ST 2128 is drafted and pending formal action. It was reported that it is expected to become finalized in the SMPTE meeting in March. It was further reported that some clarification on conversion equations is necessary (W. Husak will provide an input on that)
  + 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-AB1000 were reviewed. The following small issues in the meeting report were noted and were not considered sufficient to warrant issuing a revision. These are obviously left over from a previous report, and the correct information can be found in other places of the report:
  + In section 2.3, the twenty-seventh JVET meeting should be referenced rather the twenty-fifth.
  + In section 2.4.2, the mentioning of “placeholder” documents should have been removed, as such cases did not happen in the 28th meeting.
  + In section 9, the time when AHG setup happened should have been 1600 CEST rather than 0510 UTC
* 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
  + Conformance testing for version 2 of VVC (FDAM to be issued, DAM ballot comments in m61832)
  + New edition for HEVC (FDIS to be issued, DIS ballot comments in m61834)
  + New level and systems-related SEI for VVC (FDAM to be issued, DAM ballot comments in m61833)
  + Additional SEI in VSEI (under DAM ballot) – numerous input contributions
  + Preparation of TR for film grain (draft 4) (CDTR to be issued?)
  + Optimization of encoders and receiving systems for machine analysis of coded video content – WD of TR to be issued?
  + New edition video CICP (under DIS ballot)
  + Any action items on reference software JM/HM/VTM?
  + Additional colour type identifiers for AVC and HEVC (Draft 2 in JVET-AB1008 was issued at the last meeting) – could be included in new edition of H.265 (23008-2 DIS ballot comment on this?), no action was to be taken for AVC at this moment. It was suggested considering new editions of H.264 and ISO/IEC 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.) First version JVET-AB1012 was recently issued, update planned (also to become WG 5 output).
  + In the last meeting, public availability of experimental software packages was discussed, it was decided that the ECM branch (not the EE branches) shall be made publicly available. This has been implemented. How about NNVC? It was discussed that it would be desirable to make that available in the same way as ECM (not the training branch). It was further suggested to provide links about where to find those software packages (these are referred in document AB1012).
  + Exploration Experiments
    - Neural network-based video coding
    - Enhanced compression beyond VVC
* Liaison communication: any need for responding?
  + various from SG16 addressed to MPEG WGs in general, but nothing seems to relate to JVET/WG 5 in particular
  + JPEG has an F2F meeting next week, m62300 incoming on JPEG-AI model – revisit
* Joint meetings: with AG 5 on subjective testing, with WG 2 and VCEG on future video standardization activities, with WG 4 (and potentially WG 1) on VCM, any other (e. g. with WG 7, regarding JVET inputs on point cloud coding)?
* The number of documents was higher than for the previous meeting (150->165)
* Scheduling was discussed, and it was agreed to avoid conducting “track” sessions in parallel (some BoG parallelism could occur)
* Principles of standards development were discussed.

## Scheduling of discussions

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

* 1300–1500 1st “afternoon” session [break after 2 hours]
* 1520–1720 2nd “afternoon” session
* [“evening” break – nearly 4 hours]
* 2100–2300 1st “night” session [break after 2 hours]
* 2320–0120+1 2nd “night” session

Sessions were announced via the calendar in the JVET document site at least 22 hrs. in advance. Particular scheduling notes are shown below, although not necessarily 100% accurate or complete:

* Wed. 11 Jan., 1st day
  + Session 1:
    - 1300–1410 Opening remarks, review of practices, agenda, IPR reminder
    - 1420–1515 Reports of AHGs 1–3
  + Session 2:
    - 1535–1740 Reports of AHGs 4–15
  + Session 3:
    - 2100–2315 Review of EE1 cat. 1 and 2
  + Session 4:
    - 2335-2355 Review of EE1 cat. 3
    - 2355–0130+1 Review of EE2
* Thu. 12 Jan., 2nd day
  + Session 5:
    - 1300–1515 Review of 6.1 NNPF SEI
  + Session 6:
    - 1530–1750 Review of 4.14 VCM
  + Session 7:
    - 2100–2300 Review of EE1 and EE1 related
  + Session 8:
    - 2320–0130+1 Review of EE2 and EE2 related
* Fri. 13 Jan., 3rd day
  + BoG on NNPF (6.1, G. Sullivan) 1300-1500 and 1520-1720
  + Session 9:
    - 1300–1500 Review of EE2 (5.3.1/5.3.2)
  + Session 10:
    - 1520–1720 Review of EE2 and EE2 related (5.3.1-5.3.3)
  + Session 11:
    - 2100–2300 Review of non-EE2 (5.3.4)
  + Session 12:
    - 2320–0120+1 Review of EE1 and EE1 related (5.2.2/5.2.3)
* Mon. 16 Jan., 4th day
  + 0500–0730 MPEG information sharing session
  + Session 13:
    - 1300–1500 Further planning, Review of 4.14, 6.3, 4.2, 4.3
  + Session 14:
    - 1520–1620 Review of 4.x topics
  + 1620–1745 Joint meeting with AG 5: 4.4
  + Session 15:
    - 2100–2300 Review of non-EE2 (5.3.4)
  + Session 16:
    - 2320–0130+1 Review of non-EE2 (5.3.4)
  + NNPF BoG (G. Sullivan) 2100-2300 and 2320-0120+1
* Tue. 17 Jan., 5th day
  + 1300–1400 Joint meeting with WG 2, AG 5 and VCEG: JVET exploration status
  + Session 17:
    - 1400–1500 Review of remaining EE2 related (5.2.4), and continue non-EE2 (5.3.4)
  + Session 18:
    - 1520–1720 Review of non-EE2 (5.3.4) (chaired by Y. Ye)
  + Session 19:
    - 2100–2300 Review of EE1 related (5.2.3) and other NNVC topics (5.2.4/5.2.5)
  + Session 20:
    - 2320–0120+1 Review of non-EE2 (5.3.4) (chaired by Y. Ye)
  + NNPF BoG (G. Sullivan) 2320-0120+1
* Wed. 18 Jan., 6th day
  + 0500–0600 MPEG information sharing session
  + Session 21:
    - 1300–1400 tbd, NNPF BoG report, 4.2, 4.3
  + 1400–1525 Joint meeting with WG 4 on VCM and lenslet video coding
  + 1530–1720 BoG on VCM (C. Hollmann, S.Liu)
  + 1530–1620 Joint meeting with AG 5: 4.4
  + Session 22:
    - 1620–1730 4.2, 4.3
  + Session 23:
    - 2100–2300 4.3, 6.2, 4.2, other 4.x
  + Session 24:
    - 2320–0120+1 remaining doc 5.3.4 review
  + NNPF BoG (G. Sullivan) 2320-0120+1
* Thu. 19 Jan., 7th day
  + Session 25:
    - 1300–1500 BoG reports, review remaining docs 6.2, 4.12, 4.15, output planning, revisits
  + Session 26 (chaired by Y. Ye):
    - 1520–1720 Review docs 4.8, 4.9, 4.10, 4.11
  + Session 27:
    - 2100–2310 Remaining doc reviews 4.11, 4.1; revisits
  + Session 28:
    - 2330–0120+1 Revisits, review DoCR & liaison, output review, AHG/EE planning, a.o.b.
* Fri. 20 Jan., 8th day
  + Plenary:
    - 1300–1600 Remaining doc review, review and approval of output docs, AHGs
    - 1605-1705 Recommendations, future planning, a.o.b.
  + 2100–2300 MPEG information sharing session
  + XXXX+1–XXXX+1 WG 5 Closing plenary: Approval of meeting recommendations

## Contribution topic overview

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

* AHG reports (15) (section 3)
* Project development (section 4)
  + Deployment and advertisement of standards (2)
  + Text development and errata reporting (3)
  + Test conditions (4)
  + Subjective quality testing and verification testing (5)
  + Test Material (0)
  + Quality assessment (1)
  + Conformance test development (0)
  + Software development (3)
  + Implementation studies and complexity analysis (1)
  + AHG7: Low latency and constrained complexity (1)
  + Encoding algorithm optimization (3)
  + Profile/tier/level specification (1)
  + Proposed modification of system interface (0)
  + AHG15: Optimization of encoders and receiving systems for machine analysis of coded video content (8)
  + General aspects of standards development and applications of standards (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 (29) (section 5.2)
  + AHG12 and EE2: Enhanced compression beyond VVC capability (92) (section 5.3)
* AHG9: High-level syntax (HLS) proposals (section 6) with subtopics
  + SEI messages on neural-network post filter (22) (section 6.1)
  + SEI messages on topics other than NNPF (4) (section 6.2)
  + Film grain synthesis (1) (section 6.3)
  + Non-SEI HLS aspects (0) (section 6.4)
* Joint meetings, plenary discussions, BoG reports (1), liaison (0), 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 (15)

These reports were discussed Wednesday 11 Jan. 2023 during 1420–1515 and 1535–1740 UTC (chaired by JRO).

[JVET-AC0001](https://jvet-experts.org/doc_end_user/current_document.php?id=12242) JVET AHG report: Project management (AHG1) [J.-R. Ohm (chair), G. J. Sullivan (vice-chair)]

The work of the JVET overall had proceeded well in the interim period with approximately 10% 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_10_AB_Mainz/>). 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-AB1000) [Posted 2022-11-25, also submitted as WG 5 N 156]
* Errata report items for VVC, HEVC, AVC, Video CICP, and CP usage TR (JVET-AB1004) [Posted 2022-12-16]
* Additional colour type identifiers for AVC, HEVC and Video CICP (Draft 2) (JVET-AB1008) [Posted 2023-01-09, also submitted as WG 5 WD N 163]
* Overview of IT systems used in JVET (JVET-AB1012) [Posted 2023-01-10]
* Algorithm description for Versatile Video Coding and Test Model 18 (VTM 18) (JVET-AB2002) [Posted 2022-12-19, last update 2023-01-06, also submitted as WG 5 WD N 161]
* Additional SEI messages for VSEI (Draft 3) (JVET-AB2006) [Posted 2023-01-05, also submitted as WG 5 DAM N 158]
* VTM and HM common test conditions and software reference configurations for SDR 4:2:0 10 bit video (JVET-AB2010) [Posted 2022-12-19]
* Common Test Conditions and evaluation procedures for neural network-based video coding technology (JVET-AB2016) [Posted 2022-12-24]
* Description of algorithms and software in neural network-based video coding (NNVC) (JVET-AB2019) [Posted 2022-12-02, also submitted as WG 5 N 165]
* Film grain synthesis technology for video applications (Draft 3) (JVET-AB2020) [Posted 2022-12-29, also submitted as WG 5 WD N 159]
* Draft verification test plan for VVC multilayer coding (JVET-AB2021) [Posted 2022-11-22, last update 2022-12-16]
* Draft plan for subjective quality testing of FGC SEI message (JVET-AB2022) [Posted 2022-11-22, last update 2022-12-16]
* Exploration experiment on Neural Network-based Video Coding (EE1) (JVET-AB2023) [Posted 2022-10-27, last update 2022-11-29, also submitted as WG 5 N 164]
* Exploration experiment on Enhanced Compression beyond VVC capability (EE2) (JVET-AB2024) [Posted 2022-10-28, last update 2022-11-14, also submitted as WG 5 N 166]
* Algorithm description of Enhanced Compression Model 7 (ECM 7) (JVET-AB2025) [Posted 2022-12-22, also submitted as WG 5 N 167]
* SEI processing order SEI message in VVC (Draft 1) (JVET-AB2027) [Posted 2022-11-15, also submitted as WG 5 WD N 160]
* Additional conformance bitstreams for VVC multilayer configurations (JVET-AB2028) [Posted 2022-12-30]
* Visual quality comparison of ECM/VTM encoding (JVET-AB2029) [Posted 2022-12-23]

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

* Recommendations of the 9th WG 5 meeting (WG 5 N 155)
* Disposition of comments received on ISO/IEC 23002-7:202x (2nd Ed.) CDAM1 (WG 5 N 157)
* Text of ISO/IEC DIS 23091-2:202x Coding-independent code points - Part 2: Video (3rd edition) (WG 5 N 162)
* Request for ISO/IEC 23002-7:202x (2nd Ed.) Amd.1 Additional SEI messages (WG 5 N 125)
* Liaison statement to ISO/IEC JTC 1/SC 29/WG 1 (JPEG) on JPEG AI and NNVC (WG 5 N 168)
* Liaison statement to SMPTE on ITP-PQ-C2 colour space (WG 5 N 169)
* List of AHGs established at the 9th WG 5 meeting (WG 5 N 151)

The fifteen *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 29th meeting had been announced in the JVET reflector, in the JVET logistics document (<https://www.itu.int/wftp3/av-arch/jvet-site/2023_01_AC_Virtual/JVET-AC_Logistics.docx>), and in the WG 5 calling notice (N 171) and agenda (N 172) for the 10th WG 5 meeting.

Software integration was finalized approximately according to the plan.

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

Roughly 165 input contributions (not counting the AHG reports, summary reports and crosschecks) had been registered for consideration at the current meeting.

It is further noted that, starting from the twentieth JVET meeting, work items which had originally been conducted by the Joint Collaborative Team on Video Coding (JCT-VC) were continued to be conducted in JVET as a single joint team, as negotiated by the parent bodies. This particularly consists of work on

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

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

A preliminary basis for the document subject allocation and meeting notes for the 29th meeting had been made publicly available on the ITU-hosted ftp site <http://wftp3.itu.int/av-arch/jvet-site/2023_01_AC_Virtual/>.

[JVET-AC0002](https://jvet-experts.org/doc_end_user/current_document.php?id=12243) 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)]

**Output documents produced**

**JVET-AB1004 - Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP Usage TR**

Incorporated items at the JVET-AB meeting:

* For VVC, incorporated the following items:
  + JVET-AB0223: Text improvement for Timing / DU information SEI message in HEVC and VVC
  + Added a NOTE in the HRD text to mention that "leading pictures" associated with a DRAP picture may not be correctly decodable, per the discussion of JVET-AB0055 (On leading pictures design in DRAP SEI Message).
  + Added a NOTE to mention that the definition of IRAP in VSEI covers the case of GDR with ph\_recovery\_poc\_cnt equal to 0, per the discussion of JVET-AB0057 (On the associated IRAP for DRAP and EDRAP pictures).
  + Tickets [#1564](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1564), [#1568](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1568), [#1569](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1569), and [#1572](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1572) (some of them need more confirmation)
  + JVET-AB0120: Addition of interpolation for "initial CPB removal delay offset"
* For HEVC, incorporated the following items:
  + JVET-AB0223: Text improvement for Timing / DU information SEI message in HEVC and VVC
  + On conformance indication of the Multiview Main, Scalable Main, Scalable Main 10, and 3D Main profiles (resulted from a discussion between Alexis Tourapis and Gary Sullivan – thanks!)
  + On inference of high\_precision\_offsets\_enabled\_flag (resulted from a discussion between [Andy Fu](mailto:andy.fu@allegrodvt.com) and Gary Sullivan – thanks!)
  + On a redundant sentence in the definitionn of next\_bits( n ) from Gary Sullivan
  + On addition the parenthetic abbreviation TB to the text in clause 9.3.3.1 (suggested by [Cliff Reader](mailto:cliff@reader.com)– thanks!)
  + On indention of item 3 in the paragraph just below Equation 8-14 in clause 8.4.1 (reported and suggested by [Cliff Reader](mailto:cliff@reader.com) – thanks!)
  + On redudnant call to clause 8.6.1 (reported by [Cliff Reader](mailto:cliff@reader.com) – thanks!)
  + JVET-AB0120: Addition of interpolation for "initial CPB removal delay offset"
* For AVC, incorporated the following items:
  + JVET-AB0120: Addition of interpolation for "initial CPB removal delay offset"

**JVET-AB1008 - Additional colour type identifiers for AVC and HEVC**

This document contains the draft text for the specification of additional colour type identifiers for AVC (Rec. ITU-T H.264 | ISO/IEC 14496-10), HEVC (Rec. ITU-T H.265 | ISO/IEC 23008-2), and CICP (Rec. ITU-T H.273 | ISO/IEC 23091-2). Text modifications are provided for specification of code point identifiers for YCgCo-R colour representation with equal luma and chroma bit depths and for Draft SMPTE ST 2128. The new code points for for YCgCo-R are referred to as YCgCo-Re and YCgCo-Ro, where the number of bits added to a source RGB bit depth is 2 (i.e., even) and 1 (odd), respectively. Draft SMPTE ST 2128 specifies a colour representation referred to as ITP. Revision marking is provided to show modifications relative to the basis texts (based on the 2021-08 edition of Rec. ITU-T H.264 (the most recent edition), the 2020 edition of ISO/IEC 23008-2 (not the 5th edition currently under preparation), and ISO/IEC DIS 23091-2:202x (3rd Ed.), respectively). Equation numbers and their cross-references that are maintained automatically have been updated without revision marking.

**JVET-AB2002 - Algorithm description for Versatile Video Coding and Test Model 19 (VTM19)**

The JVET established the VVC Test Model 19 (VTM19) algorithm description and encoding method at its 28th meeting (20 – 28 October 2022, Mainz). This document serves as a source of general tutorial information on the VVC design and also provides an encoder-side description of VTM19. The VVC has been developed by a joint collaborative team of ITU-T and ISO/IEC experts known as the Joint Video Experts Team (JVET), which is a partnership of ITU-T Study Group 16 Question 6 (known as VCEG) and ISO/IEC JTC 1/SC 29/WG 11 (known as MPEG). This draft new standard has been designed with two primary goals. The first of these is to specify a video coding technology with a compression capability that is substantially beyond that of the prior generations of such standards, and the second is for this technology to be highly versatile for effective use in a broadened range of applications. Some key application areas for the use of this standard particularly include ultra-high-definition video (e.g., with 3840×2160 or 7620×4320 picture resolution and bit depth of 10 or 12 bits as specified in Rec. ITU-R BT.2100), video with a high dynamic range and wide colour gamut (e.g., with the perceptual quantization or hybrid log-gamma transfer characteristics specified in Rec. ITU-R BT.2100), and video for immersive media applications such as 360° omnidirectional video projected using a common projection format such as the equirectangular or cubemap projection format, in addition to the applications that have commonly been addressed by prior video coding standards.

Ed. Notes:

VVC Test Model 19 (VTM19) algorithm description and encoding method v2

* Incorporated JVET-AB0072: VTM Encoder Implementation for Green-MPEG SEI Messaging

VVC Test Model 19 (VTM19) algorithm description and encoding method v1

* Incorporated JVET-AA0110 and JVET-AB0267: phase indication SEI
* Incorporated JVET-AB0070: post-filter hint SEI
* Incorporated JVET-AA0102, JVET-AA0101, JVET-AB0051, and JVET-AB0069: processing order SEI
* MTS signaling bug fix in section 3.5.2
* Incorporated JVET-AB0080 and JVET-AB0081; RPR controls and filters
* Incorporated JVET-V0078; QP control for very smooth blocks

**JVET-AB2006 - Additional SEI messages for VSEI (Draft 3)**

This document contains the draft text for changes to the versatile supplemental enhancement information messages for coded video bitstreams (VSEI) standard (Rec. ITU-T H.274 | ISO/IEC 23002-7), to specify additional SEI messages, including the shutter interval information SEI message, neural-network post-filter characteristics SEI message, neural-network post-filter activation SEI message, phase indication SEI message, and post-filter hint SEI message.

**JVET-AB2027 - SEI processing order SEI message in VVC (Draft 2)**

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.

1. **Related input contributions**

It is noted that document JVET-AC0275 contains comments of the Finnish NB on the 23002-7 DAM1 that are missing from the official ballot response.

1. **Remaining bug tickets**

* [#1564](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1564) Intra prediction ref pixel array bounds too small for wide angle
* [#1572](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1572) Sub clause C.1 -- Regarding number of bitstream conformance tests to be performed

1. **Recommendations**

The AHG recommends to:

* Approve JVET-AB1004, JVET-AB1008, JVET-AB2002, JVET-AB2006, and JVET-AB2027. documents as JVET outputs,
* Compare the VVC documents with the VVC software and resolve any discrepancies that may exist, in collaboration with the software AHG,
* Encourage the use of the issue tracker to report issues with the text of both the VVC specification text and the algorithm and encoder description,
* Continue to improve the editorial consistency of VVC text specification and Test Model documents,
* Ensure that, when considering changes to VVC, properly drafted text for addition to the VVC Test Model and/or the VVC specification text is made available in a timely manner,
* Review bug tickets, and other AHG2 related inputs and act on them if found to be necessary.

(It is noted that some of the recommendations should be formulated more generally, also covering other standards than VVC).

[JVET-AC0003](https://jvet-experts.org/doc_end_user/current_document.php?id=12244) 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 19.0](https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/-/releases/VTM-19.0) (Nov. 2022)
* [HM-17.0](https://vcgit.hhi.fraunhofer.de/jvet/HM/-/releases/HM-17.0) (Jan. 2023)
* [HM-16.21+SCM-8.8](https://vcgit.hhi.fraunhofer.de/jvet/HM/-/tags/HM-16.21+SCM-8.8) (Mar. 2020)
* [SHM 12.4](https://vcgit.hhi.fraunhofer.de/jvet/SHM/-/tags/SHM-12.4) (Jan. 2018)
* [HTM 16.3](https://vcgit.hhi.fraunhofer.de/jvet/HTM/-/tags/HTM-16.3) (Jul. 2018)
* [JM 19.0](https://vcgit.hhi.fraunhofer.de/jvet/JM/-/tags/JM-19.0)
* [JSVM 9.19.15](https://vcgit.hhi.fraunhofer.de/jvet/jsvm/-/tags/JSVM_9_19_15)
* [JMVC 8.5](https://vcgit.hhi.fraunhofer.de/jvet/jmvc/-/tags/JMVC_8_5)
* [3DV ATM 15.0](https://vcgit.hhi.fraunhofer.de/jvet/3dv-atm/-/tags/3DV-ATM_v15.0) (no version history)
* [HDRTools 0.23](https://gitlab.com/standards/HDRTools/-/tags/v0.23) (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.

1. **Software development**

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

The server is located at:

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

The registration and development workflow are documented at:

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

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

1. **VTM related activities**

The VTM software can be found at

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

The software development continued on the GitLab server. VTM versions 18.1 and 18.2 were tagged on Oct. 24 and Oct. 25, and VTM version 19.0 was tagged on Nov. 30. VTM 19.1 is expected during the 29th JVET meeting.

VTM 18.1 was tagged on Oct. 24, 2022. Changes include:

* 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
* Add QPIncrementFrame description in SW manual

VTM 18.2 was tagged on Oct. 25, 2022. Changes include:

* Remove macros from previous cycle

VTM 19.0 was tagged Nov. 30, 2022. Changes include:

* JVET-AB0085: Support multiple PTL signalling (encoder only)
* JVET-AB0047: Moving the gated syntax that contains nnpfc\_uri\_tag and nnpfc\_uri...
* JVET-AB0080: GOP-based RPR encoder control also including a fix for chroma QP
* JVET-AB0081: Increased length of filters used for upscaling reconstructed pictures
* JVET-AB0072: VTM Encoder Implementation for Green-MPEG SEI Messaging
* Fix which conformance window is used in case of field coding
* Clean up code related to sign derivation
* Fix variable names
* Fix which PTL is used for checks
* Fix trailing white space, variable names, braces
* Fix variable names and spacing
* Fix variable name
* Fix inference of bit-depth for YUV files
* Fix crash during checks for layer ID of ALF APS
* Use static\_vector instead of std::vector for list of subprofiles
* Remove default parameter values and fix braces/indentation
* Use vector object to hold frame-level delta qps
* Avoid code duplication and repeated memory allocation
* Combine redundant ALF constants m\_NUM\_BITS and m\_scaleBits
* Reduce number of calls to Rice parameter derivation
* Use strong-typed enum for BPMType
* Clean up Analyze class and related code
* Updated formatRGB.cfg to include RGBtoGBR conversion
* Address compiler warnings/errors
* Clean up adaptive BT size related code
* Fix variable names
* Clean up GPM encoder code
* Use named constant for maximum CU size for IBC
* Fix !2364: distinguish number of SAD and SATD candidates
* Clean up BDOF condition checking
* Fix typo in function name
* Use bool type and constants for boolean fields
* Fix variable names and indentations
* Remove default parameter values for printOut function
* Use static\_vector instead of custom structures with separate array and item count
* Use strong-typed enum for TransType
* Use strong-typed enum for HashType
* Clean up handling of ALF data in Picture object
* Use named constants for CBF masks
* Clean up SAO code and use strong typed enums
* Remove most default parameters from motion compensation functions
* Clean up matrix multiplication functions
* Clean up HMVP related functions
* Setup general HRD parameters
* Use strong-typed enum for CoeffScanType
* Clean up ComprCUCtx data structure
* Rename JVET\_C0024\_ZERO\_OUT\_TH and clean up related functions
* Clean up deblocking filter code
* Remove -Wno-class-memaccess compiler option
* Don't optimize for deblocking when deblocking is disabled
* Clean up applyDeblockingFilterParameterSelection()
* Clean up checking slice/tile/subpic conditions in deblocking filter
* Clean up filter length handling in deblocking filter
* Clean up edge/bs handling in deblocking filter
* Avoid double braces for init
* Use enumeration for BDPCM mode
* Clean up motion discontinuity processing in deblocking filter
* Define type to hold list of ALF APS Ids and clean up related code
* Allow GREEN\_METADATA\_SEI\_ENABLED to be set externally
* Fix variable names
* Use proper preprocessor guards for GREEN\_METADATA\_SEI\_ENABLED
* Avoid using UNDEFINED which is defined as a macro by HDRLib
* Fix: add check that was removed during integration
* Fix SEIBufferingPeriod::copyTo()
* Fix for zero-sized SEI NAL units
* Fix parsing of pt\_sublayer\_delays\_present\_flag
* Fix activateAPS() to update pointers to APSs
* Store latest DRAP POC in DecLib and set it to Slice
* Fix checkCRA() for field pictures

VTM 19.1 is expected to be tagged during the 29th JVET meeting. Changes include so far:

* JVET-T0056: SEI manifest and SEI prefix indication
* JVET-AB0069 change the length of the po\_sei\_processing\_order[i] syntax element
* JVET-AB0135 and BC#017 Add nnpfc\_total\_kilobyte\_size and replace Remove unused functions
* M60678: ballot comments of FI\_03
* JVET-AB0058: signal frame upsampling in neural network post-filter characteristics
* JVET-AB0051: Modification of SEI processing order SEI message
* JVET-AB0049: Modifications of NNPFC and NNPFA
* JVET-AB0050: Add two flags for NNPFA
* JVET-AB0070: post-filter hint SEI
* Use constant for subblock size in deblocking filter
* Fix variable names
* Define constant for VDPU size and fix function name
* Clean up intra and CIIP code
* Fix variable names
* Fixes in SubpicMergeApp for NumSubPics and SubPicId setters
* Use strong-typed enum for MtsType
* Clean deblocking filter code
* SubPicMergeApp: remove duplicate ceilLog2 function
* Clean up DMVR code
* Fix compilation when disabling REUSE\_CU\_RESULTS
* Use named constants
* Use named variable for unset cost
* Clean up reading/parsing of number of active refs
* Define named reference index for IBC
* Clean up ISP and use strong typed enum
* Use strong typed enum for SEI payload type
* Rename dynamic\_cache to Pool
* nnpfc\_complexity\_idc with a flag
* Rename LFCUParam data structure
* Clean up MMVD related code
* Define RefSetArray template to define arrays indexed by listIdx and refIdx
* Move new macros into section where they are indicated to be removed in next cycle
* Fix compilation when JVET\_R0351\_HIGH\_BIT\_DEPTH\_ENABLED is enabled
* Use python3 as interpreter for cmake/CMakeBuild/bin/cmake.py
* Use single variable to control ENABLE\_TRACING
* Fix a gcc12 compiling warning/error
* Add option to enable RExt\_\_HIGH\_BIT\_DEPTH\_SUPPORT from command line
* Fix variable names
* Clean up hash related functions
* Remove Ubuntu 18.04 from auto-build pipeline
* Use ptrdiff\_t for strides
* Use strong enum type for AffineModel
* Clean buffer usage in merge RD checking functions
* Clean up BCW and AMVR related code
* Clean up merge context data structure
* Use strong typed enum for MESearchMethod
* Remove unused field
* Use strong typed ChannelType
* Use strong typed enum for MergeType
* Fix gcc 12 compiling errors
* Replace m\_subPuMC by a local variable
* Clean up MV precision handling functions and use strong-typed enum
* Define MergeIdxPair type to hold pair of merge indices
* Clean up inter merge related functions
* Clean up width/height in InterPrediction::xPredInterBlk
* Remove m\_iRefListIdx field from InterPrediction class
* Use strong typed enum for APS type
* Remove unused variable
* Use proper data type for RPR scaling
* Fix variable names and use named constants
* Use strong-type enum for DFunc
* Change copyright year to 2023
* Remove unused field
* Fix parameter type in function definition: use ptrdiff\_t instead of uint32\_t
* Clean up MVD encoding
* Group partition contexts for faster copy
* Define enum for MIP size and use Pel elements for buffers
* Improve adherence to coding guidelines in EncGOP.cpp
* Improve adherence to coding guidelines
* Remove mvRefine field from InterPredictionData
* Fix macOS build with trace file enabled
* Use Size type instead of array of two values
* Remove g\_tbMax and use floorLog2 instead
* Move sei\_processing\_order.cfg into sei\_vui subdirectory
* Remove duplicate function and fix data types
* Add Ubuntu 22.04 auto-build
* Remove macros for HLS syntax element read/write

***CTC Performance***

The following tables shows **VTM 19.0** performance over **HM 17.0**.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **All Intra Main10** |  |  |
|  |  |  | **Over HM-17.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | -29.50% | -32.83% | -33.81% | 1289% | 162% |
| Class A2 | -29.73% | -24.40% | -21.58% | 2117% | 171% |
| Class B | -22.33% | -27.21% | -31.01% | 2370% | 170% |
| Class C | -22.90% | -19.56% | -23.23% | 3432% | 173% |
| Class E | -26.05% | -25.89% | -24.34% | 1883% | 155% |
| **Overall** | -25.50% | -25.76% | -27.06% | 2196% | 167% |
| Class D | -18.79% | -13.79% | -13.86% | 3948% | 163% |
| Class F | -39.48% | -40.18% | -42.90% | 4700% | 162% |
|  |  |  |  |  |  |
|  |  |  | **Random access Main10** |  |  |
|  |  |  | **Over HM-17.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | -40.60% | -40.73% | -47.09% | 548% | 157% |
| Class A2 | -44.00% | -41.71% | -40.78% | 646% | 174% |
| Class B | -37.44% | -50.14% | -48.50% | 633% | 153% |
| Class C | -33.90% | -36.28% | -38.27% | 869% | 162% |
| Class E |  |  |  |  |  |
| **Overall** | -38.44% | -42.88% | -43.95% | 672% | 160% |
| Class D | -31.72% | -32.70% | -31.98% | 998% | 161% |
| Class F | -46.08% | -49.81% | -50.71% | 485% | 132% |
|  |  |  |  |  |  |
|  |  |  | **Low delay B Main10** |  |  |
|  |  |  | **Over HM-17.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -30.44% | -37.38% | -34.86% | 609% | 141% |
| Class C | -28.20% | -21.26% | -21.05% | 769% | 141% |
| Class E | -31.11% | -35.63% | -29.37% | 293% | 117% |
| **Overall** | -29.86% | -31.57% | -28.89% | 548% | 135% |
| Class D | -26.80% | -16.18% | -15.51% | 825% | 150% |
| Class F | -42.24% | -44.07% | -44.19% | 418% | 124% |
|  |  |  |  |  |  |
|  |  |  | **Low delay P Main10** |  |  |
|  |  |  | **Over HM-17.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -34.54% | -38.33% | -36.21% | 564% | 154% |
| Class C | -28.15% | -17.66% | -18.25% | 724% | 162% |
| Class E | -33.00% | -38.21% | -31.27% | 286% | 130% |
| **Overall** | -32.03% | -31.41% | -28.99% | 517% | 150% |
| Class D | -26.70% | -12.74% | -11.61% | 773% | 161% |
| Class F | -40.18% | -41.55% | -41.78% | 446% | 131% |

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 19.0** performance compared to **VTM 18.0** using SDR CTC:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **All Intra Main10** |  |  |
|  |  |  | **Over VTM-18.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | 0.00% | 0.00% | 0.00% | 100% | 98% |
| Class A2 | 0.00% | 0.00% | 0.00% | 100% | 98% |
| Class B | 0.00% | 0.00% | 0.00% | 100% | 94% |
| Class C | 0.00% | 0.00% | 0.00% | 100% | 93% |
| Class E | 0.00% | 0.00% | 0.00% | 100% | 95% |
| **Overall** | 0.00% | 0.00% | 0.00% | 100% | 95% |
| Class D | 0.00% | 0.00% | 0.00% | 100% | 91% |
| Class F | 0.00% | 0.00% | 0.00% | 99% | 94% |
|  |  |  |  |  |  |
|  |  |  | **Random access Main10** |  |  |
|  |  |  | **Over VTM-18.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | 0.00% | 0.00% | -0.01% | 99% | 99% |
| Class A2 | 0.00% | 0.00% | 0.01% | 98% | 100% |
| Class B | 0.00% | 0.00% | -0.01% | 100% | 98% |
| Class C | 0.00% | -0.01% | 0.02% | 99% | 95% |
| Class E |  |  |  |  |  |
| **Overall** | 0.00% | 0.00% | 0.00% | 99% | 98% |
| Class D | -0.01% | 0.01% | -0.01% | 99% | 95% |
| Class F | 0.00% | 0.00% | 0.00% | 99% | 94% |
|  |  |  |  |  |  |
|  |  |  | **Low delay B Main10** |  |  |
|  |  |  | **Over VTM-18.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -0.01% | 0.07% | -0.08% | 98% | 94% |
| Class C | 0.01% | 0.04% | 0.01% | 100% | 97% |
| Class E | 0.01% | -0.02% | -0.10% | 102% | 101% |
| **Overall** | 0.00% | 0.04% | -0.06% | 99% | 96% |
| Class D | 0.00% | -0.08% | 0.09% | 101% | 96% |
| Class F | -0.03% | 0.04% | -0.05% | 100% | 96% |
|  |  |  |  |  |  |
|  |  |  | **Low delay P Main10** |  |  |
|  |  |  | **Over VTM-18.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | 0.00% | 0.00% | 0.00% | 98% | 95% |
| Class C | 0.00% | 0.00% | 0.00% | 101% | 98% |
| Class E | 0.00% | 0.00% | 0.00% | 103% | 103% |
| **Overall** | 0.00% | 0.00% | 0.00% | 100% | 98% |
| Class D | 0.00% | 0.00% | 0.00% | 100% | 92% |
| Class F | 0.00% | 0.00% | 0.00% | 100% | 95% |

The following tables shows **VTM 19.0** performance over **HM 17.0** using HDR CTC:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access** | | | | | | | | | |
|  | **Over HM17.0** | | | | | | | | | |
|  |  |  | **wPSNR** |  |  | **PSNR** |  |  |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | -39.54% | -37.59% | -36.68% | -54.95% | -48.45% | -33.68% | -49.45% | -40.85% | 287% | 90% |
| Class H2 |  |  |  |  |  | -32.01% | -57.74% | -63.32% | 270% | 89% |
| **Overall** | -39.54% | -37.59% | -36.68% | -54.95% | -48.45% | -33.07% | -52.47% | -49.02% | 280% | 89% |
|  |  |  |  |  |  |  |  |  |  |  |
|  | **All Intra** | | | | | | | | | |
|  | **Over HM17.0** | | | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | -41.42% | -27.31% | -26.80% | -57.74% | -52.84% | -24.00% | -52.95% | -45.18% | 1474% | 101% |
| Class H2 |  |  |  |  |  | -21.76% | -47.21% | -50.55% | 1179% | 94% |
| **Overall** | -41.42% | -27.31% | -26.80% | -57.74% | -52.84% | -23.18% | -50.86% | -47.13% | 1359% | 98% |

The following tables show **VTM 19.0** performance compared to **VTM 18.0** using HDR CTC:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access** | | | | | | | | | |
|  | **Over VTM18.0** | | | | | | | | | |
|  |  |  | **wPSNR** |  |  | **PSNR** |  |  |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | 0.04% | 0.01% | 0.00% | -0.01% | 0.01% | 0.00% | 0.00% | 0.01% | 95% | 92% |
| Class H2 |  |  |  |  |  | 0.00% | 0.00% | 0.00% | 97% | 91% |
| **Overall** | 0.04% | 0.01% | 0.00% | -0.01% | 0.01% | 0.00% | 0.00% | 0.01% | 96% | 92% |
|  |  |  |  |  |  |  |  |  |  |  |
|  | **All Intra** | | | | | | | | | |
|  | **Over VTM18.0** | | | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.01% | 0.00% | 96% | 91% |
| Class H2 |  |  |  |  |  | 0.00% | 0.00% | 0.00% | 90% | 88% |
| **Overall** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 94% | 90% |

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

***Issues in VTM affecting conformance***

The following issues in VTM master branch may affect conformance:

* Missing HLS features (see sections below)

Merge request 2343 (“Fix which PTL is used for checks”) fixes which Profile/Tier/Level data structure is used (SPS or VPS) for some conformance checks. After this correction (i.e., VTM versions 19.0 and later), the following conformance streams fail to decode:

* ILRPL\_A\_Huawei\_2
* OLS\_A\_Tencent\_5
* OLS\_B\_Tencent\_5
* OLS\_C\_Tencent\_5
* OPI\_B\_Nokia\_3
* SPATSCAL444\_A\_Qualcomm\_2
* SPATSCAL\_A\_Qualcomm\_3
* VPS\_A\_INTEL\_3
* VPS\_B\_ERICSSON\_1
* VPS\_C\_ERICSSON\_2

***Status of implementation of proposals of previous JVET meetings***

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

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

1. **HM related activities**

HM 17.0 was tagged on Jan. 5, 2023. Changes include:

* JVET-AA0130: VTM and HM common test conditions for high bit depth
* JVET-AB0271: per-sequence config files for HDR content
* JCTVC-AD0021(JVET-T0056) SEI manifest & SEI prefix indication
* Build fixes for Xcode 14.2
* Remove Ubuntu 18.04 from auto-build pipeline
* Change Python executable to python3 for build environment scripts

The following tables show **HM 17.0** performance compared to **HM 16.26** under SRD CTC:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **All Intra Main10** |  |  |
|  |  |  | **Over HM-16.26** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | 0.00% | 0.00% | 0.00% | 101% | 98% |
| Class A2 | 0.00% | 0.00% | 0.00% | 102% | 97% |
| Class B | 0.00% | 0.00% | 0.00% | 101% | 102% |
| Class C | 0.00% | 0.00% | 0.00% | 100% | 102% |
| Class E | 0.00% | 0.00% | 0.00% | 101% | 101% |
| **Overall** | 0.00% | 0.00% | 0.00% | 101% | 100% |
| Class D | 0.00% | 0.00% | 0.00% | 101% | 100% |
| Class F | 0.00% | 0.00% | 0.00% | 101% | 100% |
|  |  |  |  |  |  |
|  |  |  | **Random access Main10** |  |  |
|  |  |  | **Over HM-16.26** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | 0.00% | 0.00% | 0.00% | 101% | 102% |
| Class A2 | 0.00% | 0.00% | 0.00% | 101% | 103% |
| Class B | 0.00% | 0.00% | 0.00% | 101% | 105% |
| Class C | 0.00% | 0.00% | 0.00% | 100% | 100% |
| Class E |  |  |  |  |  |
| **Overall** | 0.00% | 0.00% | 0.00% | 100% | 103% |
| Class D | 0.00% | 0.00% | 0.00% | 99% | 99% |
| Class F | 0.00% | 0.00% | 0.00% | 101% | 101% |
|  |  |  |  |  |  |
|  |  |  | **Low delay B Main10** |  |  |
|  |  |  | **Over HM-16.26** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | 0.00% | 0.00% | 0.00% | 100% | 104% |
| Class C | 0.00% | 0.00% | 0.00% | 101% | 115% |
| Class E | 0.00% | 0.00% | 0.00% | 100% | 110% |
| **Overall** | 0.00% | 0.00% | 0.00% | 101% | 109% |
| Class D | 0.00% | 0.00% | 0.00% | 100% | 102% |
| Class F | 0.00% | 0.00% | 0.00% | 100% | 106% |
|  |  |  |  |  |  |
|  |  |  | **Low delay P Main10** |  |  |
|  |  |  | **Over HM-16.26** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | 0.00% | 0.00% | 0.00% | 101% | 104% |
| Class C | 0.00% | 0.00% | 0.00% | 101% | 108% |
| Class E | 0.00% | 0.00% | 0.00% | 101% | 106% |
| **Overall** | 0.00% | 0.00% | 0.00% | 101% | 106% |
| Class D | 0.00% | 0.00% | 0.00% | 100% | 104% |
| Class F | 0.00% | 0.00% | 0.00% | 101% | 106% |

The following table shows **HM 17.0** performance compared to **HM 16.26** under HDR CTC:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access** | | | | | | | | | |
|  | **Over HM16.26** | | | | | | | | | |
|  |  |  | **wPSNR** |  |  | **PSNR** |  |  |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 101% | 99% |
| Class H2 |  |  |  |  |  | 0.00% | 0.00% | 0.00% | 98% | 94% |
| **Overall** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 97% |
|  |  |  |  |  |  |  |  |  |  |  |
|  | **All Intra** | | | | | | | | | |
|  | **Over HM16.26** | | | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 102% | 108% |
| Class H2 |  |  |  |  |  | 0.00% | 0.00% | 0.00% | 102% | 105% |
| **Overall** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 102% | 107% |

The following MRs are pending [with status indicated]:

* JVET-X0048: implementation of film grain characteristics (FGC) SEI message [waiting for proponent feedback]
* Port the Y4M support [one issue remains]
* Mark the current picture as short-term ref (for SCM) [need SCC expert reviewer]

The HM SCC (SCM) branch (HM-16.21+SCM-8.8) has not been updated for the recent HM versions. Updating SCM to, for example, HM-17.0+SCM-8.8 should be considered. It may though be helpful to move SCC related functionality into separate source files. Volunteer work towards merging the branches would be appreciated.

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

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.

1. **360Lib related activities**

The latest 360Lib software (360Lib-13.4) can be found at <https://vcgit.hhi.fraunhofer.de/jvet/360lib/-/tags/360Lib-13.4>

The following table is for the projection formats comparison using VTM-19.0 according to 360-degree video CTC (JVET-U2012) compared to that using VTM-18.0 (VTM-18.0 as anchor).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **PERP: VTM-19.0 over VTM-18.0** | | | | | |
|  | **End-to-end**  **WS-PSNR** | | | **End-to-end**  **S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | 0.00% | 0.00% | 0.01% | 0.00% | -0.01% | 0.01% |
| Class S2 | 0.00% | -0.02% | 0.01% | 0.00% | -0.01% | 0.01% |
| **Overall** | 0.00% | -0.01% | 0.01% | 0.00% | -0.01% | 0.01% |

The following table compares generalized cubemap (GCMP) coding and padded equi-rectangular projection (PERP) coding using VTM-19.0 (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% | -5.53% | -6.05% | -11.50% | -5.47% | -6.00% |
| Class S2 | -3.72% | 0.80% | 1.21% | -3.69% | 0.89% | 1.29% |
| **Overall** | -8.40% | -3.00% | -3.14% | -8.38% | -2.93% | -3.08% |

The following tables are for PERP and GCMP coding comparison between VTM-19.0 and HM-16.22 (HM as anchor), respectively.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **VTM-19.0 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% | -41.28% |
| Class S2 | -36.88% | -37.36% | -39.69% | -36.87% | -37.40% | -39.75% |
| **Overall** | -33.30% | -38.23% | -40.65% | -33.29% | -38.27% | -40.67% |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **VTM-19.0 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% | -40.81% | -42.74% |
| Class S2 | -39.84% | -39.57% | -41.53% | -39.85% | -39.58% | -41.57% |
| **Overall** | -37.30% | -40.35% | -42.29% | -37.28% | -40.32% | -42.27% |

1. **SCM related activities**

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

1. **SHM related activities**

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

1. **HTM related activities**

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

The next release will include the following changes:

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

1. **HDRTools related activities**

There had not been any updates of HDRTools.

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

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

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

1. **Bug tracking**

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

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

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

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

Bug tracking for HDRTools is located at:

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

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

1. **Software repositories**

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

1. **CTC alignment and merging**

Following the merger for high-bit depth CTCs (JVET-AA0130), there are currently 7 related JVET 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

Older CTC documents are:

JCTVC-X1007 SHVC test conditions

JCT3V-G1100 3DV test conditions

JCTVC-Z1015 SCM test conditions

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

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

1. **Guidelines for reference software development**

Guidelines for VVC and HEVC reference software development have been updated and submitted as JVET-AC0166 and JVET-AC0204.

Updates include clarification of existing rules, removal of outdated rules, and updates of supported compiler versions. It is also proposed to bump the C++ version to C++17 such as to have better support for constexpr constructs and additional features such as [[maybe\_unused]] attributes and overaligned memory allocation.

1. **Recommendations**

The AHG recommends to:

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

It was discussed that all conformance streams failing to decode as reported above (due to a software bug fix which caused to generate erroneous streams) are relating to multi-layer configurations, and this could be either resolved by a new edition, or in the context of the ongoing multi-layer conformance activity.

Target a new edition of software in the July meeting, to reflect the bug fixing done more recently. This would also avoid misalignment between published software and conformance versions.

[JVET-AC0004](https://jvet-experts.org/doc_end_user/current_document.php?id=12245) JVET AHG report: Test material and visual assessment (AHG4) [V. Baroncini, T. Suzuki, M. Wien (co-chairs), S. Liu, S. Puri, A. Segall, P. Topiwala, S. Wenger, J. Xu, Y. Ye (vice-chairs)]

**Activities**

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

A teleconference was organized by AG5 and JVET AhG4 2022-12-16 08:30h UTC. The topic was the Verification Test preparation for VVC Multilayer Coding. The report of the AhG meeting is available in JVET-AC0042.

***Subjective test preparation for film grain synthesis***

The teleconference was organized by AG5 and JVET AhG4 2022-12-16 17:00h UTC. The topic was the preparation for subjective testing of film grain synthesis. The report of the AhG meeting is available in JVET-AC0043.

***Test sequences***

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

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

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

**Related contributions**

|  |  |  |
| --- | --- | --- |
| [JVET-AC0041](https://jvet-experts.org/doc_end_user/current_document.php?id=12226) | [AHG4] On visual volumetric video-based coding (V3C) testing conditions | [S. Schwarz](mailto:sebastian.schwarz@nokia.com), M. M. Hannuksela, P. Rondao Alface, L. Kondrad, L. Ilola (Nokia) |
| [JVET-AC0042](https://jvet-experts.org/doc_end_user/current_document.php?id=12227) | AHG4: Report on AHG meeting on verification test preparation for VVC multilayer coding | [M. Wien (RWTH)](mailto:wien@lfb.rwth-aachen.de) |
| [JVET-AC0043](https://jvet-experts.org/doc_end_user/current_document.php?id=12228) | AHG4: Report on AHG meeting on subjective test preparation for film grain synthesis | [M. Wien (RWTH)](mailto:wien@lfb.rwth-aachen.de) |
| [JVET-AC0044](https://jvet-experts.org/doc_end_user/current_document.php?id=12229) | [AHG4] Occupancy-only PSNR calculations for V3C V-PCC coding evaluation | [S. Schwarz](mailto:sebastian.schwarz@nokia.com), M. M. Hannuksela (Nokia) |
| [JVET-AC0145](https://jvet-experts.org/doc_end_user/current_document.php?id=12347) | AHG4: experiments in preparation of dual-layer VVC visual tests | [P. de Lagrange (InterDigital)](mailto:philippe.delagrange@interdigital.com) |
| [JVET-AC0181](https://jvet-experts.org/doc_end_user/current_document.php?id=12385) | AHG4: experiments in preparation of film grain visual tests | [P. de Lagrange (InterDigital)](mailto:philippe.delagrange@interdigital.com) |
| [JVET-AC0267](https://jvet-experts.org/doc_end_user/current_document.php?id=12472) | Training Methods in Visual Assessment: Potential Improvements for Expert Viewing Tests | [M. Wien (RWTH)](mailto:wien@lfb.rwth-aachen.de), [V. Baroncini (VABTech)](mailto:baroncini@gmx.com) |

**Recommendations**

The AHG recommends:

* To review and discuss the input related to the verification test plan for VVC multilayer coding proposed in JVET-AC0145.
* To review and discuss the input related to the subjective tests for the FGC SEI message proposed in JVET-AC0181.
* Review JVET-AC0041 and JVET-AC0044 in a joint meeting with AG5 and WG7
* Review JVET-AC0267 in a joint session with AG 5
* 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.

It was suggested that the aspects targeted in JVET-AC0041 should be discussed at a higher level with other WGs (2, 4, 7) and VCEG first, e.g. whether it is desirable to develop dedicated tools or profiles for PCC in JVET.

[JVET-AC0005](https://jvet-experts.org/doc_end_user/current_document.php?id=12246) 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)]

1. **Activities**

The AHG communication is conducted through the main JVET reflector, jvet@lists.rwth-aachen.de, with [AHG5] in message headers. However, no correspondence marked as AHG5 was sent between the 28th and 29th meetings.

1. **Timeline**

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

* **VVCv1 conformance:**
  + ISO/IEC FDIS 23090-15 issued from 2021-10 meeting, FDIS registered for formal approval 2022-07-11, FDIS ballot closed 2022-11-04, standard published 2022-11-24
  + H.266.1 V1 Consent 2022-01-28, Last Call began 2022-04-01, Approved 2022-04-29, pre-published 2022-05-17, standard published 2022-07-12.
* **VVCv2 conformance:**
  + ISO/IEC 23090-15/Amd.1 CDAM: 2021-10
  + ISO/IEC 23090-15/Amd.1 DAM: 2022-01
  + DAM ballot closed 2022-11-15, ready to proceed to FDAM at current meeting 2023-01
  + H.266.1 V2 Consent planned for 2023-07

1. **Status on bitstream submission**

There were no changes to the submitted bitstreams. The status at the time of preparation of this report is as follows:

* conformance bitstreams for VVC:
  + 104 bitstream categories have been identified
  + At least one bitstream has been submitted in each identified category
  + 283 total bitstreams have been provided, checked, and made available
  + No changes between 27th and 28th meeting.
* conformance bitstreams for VVC operation range extensions:
  + 57 bitstream categories have been identified
  + 1 bitstream of 1 identified category has been re-generated
  + 128 (was 127) bitstreams of 57 (was 56) identified categories have been cross-checked and uploaded.

1. **Activities and Discussion**

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

VVC activities:

Ticket #1581 was filed. According to it the following streams:

* ILRPL\_A\_Huawei\_2
* OLS\_A\_Tencent\_5
* OLS\_B\_Tencent\_5
* OLS\_C\_Tencent\_5
* OPI\_B\_Nokia\_3
* SPATSCAL444\_A\_Qualcomm\_2
* SPATSCAL\_A\_Qualcomm\_3
* VPS\_A\_INTEL\_3
* VPS\_B\_ERICSSON\_1
* VPS\_C\_ERICSSON\_2

need to be re-generated. The issue may be resolved after applying additional encoder fix related to JVET-AB0085.

VVC operation range extensions activities:

None.

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

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

1. **Contributions**

No AhG5 related contributions to this meeting.

1. **Ftp site information**

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

* VVC:

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

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

* VVC operation range extensions:

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

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

The ftp site for uploading bitstream file is as follows.

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

(user id: avguest, passwd: Avguest201007)

If using FileZilla, the following configuration is suggested:

Graphical user interface, text, application, email

Description automatically generated

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

1. **Recommendations**

The AHG recommends the following:

* Proceed with the generation, cross-checking, and documentation of the additional conformance bitstreams for VVC, in particular conformance bitstreams for VVC multilayer configurations (JVET-AB2028).
* Maintain and update the conformance bitstream database and contribute to report problems in JVET document 1004.
* Start the generation, cross-checking, and documentation of the conformance streams for the HEVC multiview profiles supporting extended bit depth (JVET-AA1011).

Based on ballot comment JP-006, it would be possible to include the corrected v1 bitstreams into the FDAM/FDIS.

[JVET-AC0006](https://jvet-experts.org/doc_end_user/current_document.php?id=12240) JVET AHG report: ECM software development (AHG6) [V. Seregin (chair), J. Chen, F. Le Léannec, K. Zhang (vice-chairs)]

1. **Software development**

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

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

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

The following adopted aspects were integrated into ECM-7.0:

JVET-AB0155: Spatial GPM (Test 1.6b) (MR 271)

JVET-AB0092: GLM with luma value (Test 1.8b) (MR 267)

JVET-AB0157: Template based MRL and intra fusion (Test 1.12a) (MR 275)

JVET-AB0143: CCCM with template selection (Test 1.13a) (MR 260)

JVET-AB0130: IntraTMP search speedup (Test 1.14). Enable Intra TMP in classes A-E (CTC) (MR 274)

JVET-AB0078: AMVPmerge mode for LB (Test 2.1a) (MR 253)

JVET-AB0079: TM-based BCW index derivation (Test 2.2) (MR 262)

JVET-AB0112: Affine DMVR (Test 2.6) (MR 263)

JVET-AB0061: Using block vector derived from IntraTMP for IBC (Test 3.2) (MR 257)

JVET-AB0067: Modification of LFNST for MIP coded block (Test 4.1) (MR 272)

JVET-AB0184: Extended fixed filter output for ALF (Test 5.1) (MR 261)

JVET-AB0082: Fix for ARMC and use 12-tap filter for picture upscaling (MR 276)

JVET-AB0174: Division free CCCM (MR 257)

JVET-AB0189: RMVF bit length control (MR 265)

Included fixes:

Fix to use unfiltered source for HDRmetrics when MCTF is enabled (MR 273)

Fix ISP CHECKD condition (MR 277)

Fix out of bound access to array EXT\_REF\_LINE\_IDX (MR 281)

Disable EncDbOpt when deblocking filter is disabled (MR 288)

Picture padding to include template (290)

Check if IBC is used before resetting IBC buffer (MR 297)

Memory deallocation for CCSAO (MR 298)

ECM-7.0 was tagged on December 7, 2022.

There are no changes to VTM-11 anchor, the latest version is VTM-11ecm6 which was tagged on September 4, 2022.

***CTC Performance***

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

Next tables show ECM-7.0 performance over ECM-6.0 anchor.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | |
|  | **Over ECM-6.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -0.57% | -1.41% | -1.14% | 121% | 104% |
| Class A2 | -1.03% | -2.81% | -3.34% | 119% | 107% |
| Class B | -1.00% | -2.22% | -2.00% | 120% | 110% |
| Class C | -0.83% | -1.17% | -1.16% | 123% | 121% |
| Class E | -1.52% | -1.92% | -1.78% | 125% | 119% |
| **Overall** | -0.98% | -1.90% | -1.86% | 122% | 112% |
| Class D | -0.65% | -0.71% | -0.94% | 115% | 112% |
| Class F | -0.78% | -2.74% | -2.70% | 105% | 105% |
| Class TGM | -1.81% | -5.38% | -4.66% | 102% | 92% |
|  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | |
|  | **Over ECM-6.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -0.86% | -0.75% | -1.37% | 104% | 93% |
| Class A2 | -1.08% | -1.93% | -2.16% | 103% | 98% |
| Class B | -0.93% | -1.55% | -1.60% | 110% | 96% |
| Class C | -0.62% | -0.84% | -0.71% | 115% | 107% |
| Class E |  |  |  |  |  |
| **Overall** | -0.86% | -1.28% | -1.43% | 109% | 99% |
| Class D | -0.58% | -0.79% | -0.92% | 115% | 105% |
| Class F | -1.12% | -2.69% | -2.59% | 109% | 102% |
| Class TGM | -1.05% | -3.12% | -2.73% | 108% | 99% |
|  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | |
|  | **Over ECM-6.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -0.81% | -1.44% | -1.25% | 105% | 93% |
| Class C | -0.67% | -0.93% | -0.90% | 108% | 93% |
| Class E | -0.64% | 0.17% | -0.71% | 115% | 95% |
| **Overall** | -0.72% | -0.87% | -1.00% | 108% | 93% |
| Class D | -0.77% | -0.48% | -0.14% | 117% | 101% |
| Class F | -0.74% | -2.06% | -1.44% | 107% | 100% |
| Class TGM | -0.37% | -1.36% | -1.10% | 102% | 94% |
|  |  |  |  |  |  |
|  | **Low delay P Main 10** | | | | |
|  | **Over ECM-6.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -0.46% | -0.79% | -0.49% | 103% | 93% |
| Class C | -0.40% | -0.21% | -0.68% | 105% | 95% |
| Class E | -0.27% | -0.04% | -0.03% | 114% | 99% |
| **Overall** | -0.39% | -0.41% | -0.44% | 106% | 95% |
| Class D | -0.50% | -0.17% | 0.24% | 111% | 98% |
| Class F | -0.42% | -1.65% | -1.25% | 103% | 97% |
| Class TGM | -0.43% | -1.44% | -1.20% | 100% | 92% |

The below tables show ECM-7.0 performance comparing to VTM-11.0ecm6 anchor.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | |
|  | **Over VTM-11.0ecm6** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -8.67% | -17.52% | -24.11% | 617% | 348% |
| Class A2 | -12.88% | -21.39% | -24.64% | 576% | 348% |
| Class B | -7.62% | -20.59% | -18.79% | 630% | 392% |
| Class C | -8.12% | -12.50% | -12.91% | 635% | 411% |
| Class E | -9.72% | -17.97% | -16.91% | 619% | 435% |
| **Overall** | -9.13% | -17.98% | -19.03% | 618% | 387% |
| Class D | -6.68% | -10.00% | -10.01% | 624% | 416% |
| Class F | -18.45% | -26.68% | -26.24% | 444% | 397% |
| Class TGM | -31.47% | -38.80% | -37.65% | 386% | 367% |
|  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | |
|  | **Over VTM-11.0ecm6** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -19.46% | -21.91% | -29.05% | 583% | 646% |
| Class A2 | -21.45% | -28.23% | -31.25% | 586% | 894% |
| Class B | -17.65% | -28.04% | -25.93% | 614% | 830% |
| Class C | -19.26% | -22.56% | -22.39% | 605% | 892% |
| Class E |  |  |  |  |  |
| **Overall** | -19.20% | -25.39% | -26.67% | 600% | 817% |
| Class D | -20.25% | -23.80% | -24.06% | 605% | 959% |
| Class F | -22.05% | -29.05% | -29.34% | 520% | 556% |
| Class TGM | -30.00% | -35.63% | -35.75% | 556% | 438% |
|  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | |
|  | **Over VTM-11.0ecm6** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -16.16% | -30.50% | -28.78% | 529% | 695% |
| Class C | -17.44% | -25.63% | -25.66% | 552% | 714% |
| Class E | -14.58% | -22.41% | -23.17% | 552% | 483% |
| **Overall** | -16.19% | -26.85% | -26.34% | 542% | 640% |
| Class D | -19.65% | -26.43% | -26.39% | 554% | 850% |
| Class F | -20.49% | -29.26% | -27.88% | 514% | 496% |
| Class TGM | -28.89% | -35.39% | -35.06% | 524% | 401% |
|  |  |  |  |  |  |
|  | **Low delay P Main 10** | | | | |
|  | **Over VTM-11.0ecm6** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -15.41% | -30.16% | -27.87% | 486% | 694% |
| Class C | -16.91% | -24.79% | -24.39% | 460% | 701% |
| Class E | -14.28% | -22.01% | -22.86% | 519% | 501% |
| **Overall** | -15.63% | -26.33% | -25.46% | 485% | 642% |
| Class D | -19.78% | -26.11% | -26.07% | 456% | 758% |
| Class F | -19.49% | -28.95% | -27.62% | 500% | 481% |
| Class TGM | -25.88% | -33.29% | -33.07% | 543% | 399% |

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

1. **Recommendations**

The AHG recommends to:

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

[JVET-AC0007](https://jvet-experts.org/doc_end_user/current_document.php?id=12247) JVET AHG report: Low latency and constrained complexity (AHG7) [A. Duenas, T. Poirier, S. Liu (co-chairs), L. Wang, J. Xu (vice-chairs)]

**Related Contributions**

[JVET-AC0180](https://jvet-experts.org/doc_end_user/current_document.php?id=12384) **AHG 7: Reference frame padding for GDR [T.Poirier, F. Aumont (Interdigital)]**

This contribution proposes to pad the dirty area using samples from the clean area instead of setting unavailable samples to 2*BD-1* for reference pictures used by CUs in the clean (refreshed) area.

**Recommendations**

The AHG recommends reviewing input contributions and:

* to study the impact of the padding of reference frames for GDR and the option of enabling GDR for low latency and controlled complexity.

Consider continuing this as part of the encoder optimization activity (with additional mandates), normative aspects could be studied in contxt of ECM.

[JVET-AC0008](https://jvet-experts.org/doc_end_user/current_document.php?id=12248) 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)]

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

1. Contributions

No contributions relevant to the study of high bit depth, high bit rate or high frame rate coding have been registered for the 29th meeting.

1. Benchmarks

The following provide a comparison of VTM19.0 and HM17.0 for 12 and 16 bit content. The tables below report BD-Rate results computed using PSNR only.

*Low QP Range, BD-rate gain on PSNR.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **HDR PQ** |  |  | **AI** |  |  |
|  | **VTM19.0 over HM17.0** | | | | |
|  | psnrY | psnrU | psnrV | EncT | DecT |
| PQ444 | -5.73% | -6.65% | -7.35% | 2932% | 164% |
| PQ422 | -8.96% | -12.80% | -13.44% | 2473% | 158% |
| **Overall** | -7.34% | -9.73% | -10.39% | 2702% | 161% |
|  |  |  |  |  |  |
|  |  |  | **LDB** |  |  |
|  | **VTM19.0 over HM17.0** | | | | |
|  | psnrY | psnrU | psnrV | EncT | DecT |
| PQ444 | -6.65% | -4.92% | -6.94% | 288% | 160% |
| PQ422 | -7.86% | -12.08% | -13.09% | 366% | 157% |
| **Overall** | -7.26% | -8.50% | -10.02% | 327% | 158% |
|  |  |  |  |  |  |
|  |  |  | **RA** |  |  |
|  | **VTM19.0 over HM17.0** | | | | |
|  | psnrY | psnrU | psnrV | EncT | DecT |
| PQ444 | -6.28% | -4.63% | -5.92% | 322% | 160% |
| PQ422 | -7.83% | -11.79% | -12.02% | 450% | 158% |
| **Overall** | -7.05% | -8.21% | -8.97% | 386% | 159% |
|  |  |  |  |  |  |
| **Overall PQ** | -7.22% | -8.81% | -9.79% | 1139% | 159% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **HDR HLG** |  |  | **AI** |  |  |
|  | **VTM19.0 over HM17.0** | | | | |
|  | psnrY | psnrU | psnrV | EncT | DecT |
| HLG444 | -4.58% | -6.22% | -6.60% | 3569% | 158% |
| HLG422 | -5.32% | -7.62% | -7.71% | 3021% | 152% |
| **Overall** | -4.95% | -6.92% | -7.16% | 3295% | 155% |
|  |  |  |  |  |  |
|  |  |  | **LDB** |  |  |
|  | **VTM19.0 over HM17.0** | | | | |
|  | psnrY | psnrU | psnrV | EncT | DecT |
| HLG444 | -6.30% | -8.70% | -8.34% | 393% | 148% |
| HLG422 | -6.04% | -10.29% | -9.16% | 476% | 148% |
| **Overall** | -6.17% | -9.49% | -8.75% | 435% | 148% |
|  |  |  |  |  |  |
|  |  |  | **RA** |  |  |
|  | **VTM19.0 over HM17.0** | | | | |
|  | psnrY | psnrU | psnrV | EncT | DecT |
| HLG444 | -6.37% | -8.59% | -7.92% | 440% | 150% |
| HLG422 | -6.05% | -9.75% | -8.20% | 594% | 153% |
| **Overall** | -6.21% | -9.17% | -8.06% | 517% | 152% |
|  |  |  |  |  |  |
| **Overall HLG** | -5.78% | -8.53% | -7.99% | 1416% | 152% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **SVT RGB** |  |  | **AI** |  |  |
|  | **VTM19.0 over HM17.0** | | | | |
|  | psnrG | psnrB | psnrR | EncT | DecT |
| SVT16 | -2.64% | -1.83% | -1.85% | 5548% | 145% |
| SVT12 | -4.29% | -3.19% | -3.28% | 5389% | 145% |
| **Overall** | -3.47% | -2.51% | -2.56% | 5468% | 145% |
|  |  |  |  |  |  |
|  |  |  | **LDB** |  |  |
|  | **VTM19.0 over HM17.0** | | | | |
|  | psnrG | psnrB | psnrR | EncT | DecT |
| SVT16 | -2.52% | -1.14% | -1.13% | 711% | 148% |
| SVT12 | -4.61% | -2.13% | -2.14% | 653% | 145% |
| **Overall** | -3.57% | -1.64% | -1.63% | 682% | 146% |
|  |  |  |  |  |  |
|  |  |  | **RA** |  |  |
|  | **VTM19.0 over HM17.0** | | | | |
|  | psnrG | psnrB | psnrR | EncT | DecT |
| SVT16 | -2.66% | -1.20% | -1.19% | 852% | 149% |
| SVT12 | -4.71% | -2.21% | -2.22% | 793% | 145% |
| **Overall** | -3.69% | -1.71% | -1.71% | 822% | 147% |
|  |  |  |  |  |  |
| **Overall RGB** | -3.57% | -1.95% | -1.97% | 2324% | 146% |

*Standard QP Range, BD-rate gain on PSNR.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **RA** | | | | |
|  | **VTM19.0 over HM17.0** | | | | |
|  | psnrY | psnrU | psnrV | EncT | DecT |
| Class H1 | -26.50% | -69.30% | -56.97% | 290% | 177% |
| Class H2 | -33.05% | -57.93% | -57.02% | 293% | 161% |
| **Overall** | -26.50% | -69.30% | -56.97% | 291% | 169% |
|  |  |  |  |  |  |
|  | **AI**  **VTM19.0 over HM17.0** | | | | |
|  | psnrY | psnrU | psnrV | EncT | DecT |
| Class H1 | -12.58% | -68.09% | -54.83% | 2104% | 185% |
| Class H2 | -23.06% | -49.57% | -50.26% | 2530% | 174% |
| **Overall** | -12.58% | -68.09% | -54.83% | 2307% | 179% |

1. Recommendations

The AHG recommends the following:

* To close the group.

The overall results above are probably only considering class H2. Correct from V2.

It is commented that it would be relevant to continue running the CTC in JVET-AA2018 also for future meeting cycles, but not necessary to retain a separate AHG for that. It is noted that K. Sharman is already one of the AHG3 chairs.

[JVET-AC0009](https://jvet-experts.org/doc_end_user/current_document.php?id=12249) 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)]

1. **Related contributions**

A total of 27 contributions are identified relating to the mandates of AHG9. The number of contributions relating to each mandate is as follows

* 23 contributions relate to the mandate to study the SEI messages in VSEI, VVC, HEVC, and AVC;
  + 21 contributions relate to the neural-network post-filter SEI messages, including 1 summary of proposals on NNPF SEI messages.
  + 1 contribution relates to the SEI processing order SEI message
  + 1 contribution relates to preparing the FDAM/FDIS text of VVC 2nd Ed.
* 0 contributions relate to the mandate to collect software and showcase information for SEI messages;
* 4 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.

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

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

**Summary of proposals related to NNPF SEI messages**

[JVET-AC0208](https://jvet-experts.org/doc_end_user/current_document.php?id=12412) AHG9: Summary of Proposals Related to Neural-Network Post-Filter SEI Messages [S. Deshpande (Sharp)]

**Proposals on NNPF SEI messages**

[JVET-AC0047](https://jvet-experts.org/doc_end_user/current_document.php?id=12232) AHG9: On bugfix for NNPFC and NNPFA SEI messages [T. Chujoh, Y. Yasugi, T. Ikai (Sharp)]

[JVET-AC0061](https://jvet-experts.org/doc_end_user/current_document.php?id=12263) AHG9: Comments on Neural-network Post-filter Characteristics SEI Message [S. Deshpande (Sharp)]

[JVET-AC0062](https://jvet-experts.org/doc_end_user/current_document.php?id=12264) AHG9: On NNPFC SEI Message [S. Deshpande, A. Sidiya (Sharp)]

[JVET-AC0074](https://jvet-experts.org/doc_end_user/current_document.php?id=12276) AHG9: On the VVC use of the NNPFC SEI message for picture rate upsampling [M. M. Hannuksela (Nokia)]

[JVET-AC0075](https://jvet-experts.org/doc_end_user/current_document.php?id=12277) AHG9: On NNPFC and NNPFA SEI messages for picture rate upsampling post-filter [M. M. Hannuksela (Nokia)]

[JVET-AC0076](https://jvet-experts.org/doc_end_user/current_document.php?id=12278) AHG9: AHG15: On the NNPFC SEI message for machine analysis [M. M. Hannuksela, F. Cricri, J. I. Ahonen, H. Zhang (Nokia)]

[JVET-AC0085](https://jvet-experts.org/doc_end_user/current_document.php?id=12287) AHG9: On neural-network post-filter characteristics (NNPFC) SEI message for temporal upsampling towards machine vision [S. Wang, J. Chen, Y. Ye (Alibaba), S. Wang (CityU)]

[JVET-AC0127](https://jvet-experts.org/doc_end_user/current_document.php?id=12329) AHG9: Combination of picture rate upsampling with other NNPF purposes [Y.-K. Wang, J. Xu, Y. Li, L. Zhang, K. Zhang, J. Li, C. Lin (Bytedance)]

[JVET-AC0128](https://jvet-experts.org/doc_end_user/current_document.php?id=12330) AHG9: On the signalling of complexity information in NNPFC SEI message [Hendry, S. Kim (LGE)]

[JVET-AC0129](https://jvet-experts.org/doc_end_user/current_document.php?id=12331) AHG9: On the NNPFC SEI message update and activation [Hendry, J. Nam, H. Jang, S. Kim, J. Lim (LGE)]

[JVET-AC0131](https://jvet-experts.org/doc_end_user/current_document.php?id=12333) AHG9: On NNPFC SEI message repetition [Hendry, J. Nam, H. Jang, S. Kim, J. Lim (LGE)]

[JVET-AC0132](https://jvet-experts.org/doc_end_user/current_document.php?id=12334) AHG9: On design for region-based neural-network post-filter SEI message [Hendry, J. Nam, H. Jang, S. Kim, J. Lim (LGE)]

[JVET-AC0134](https://jvet-experts.org/doc_end_user/current_document.php?id=12336) AHG9: Signalling of NNPF quality improvement [C. Lin, Y.-K. Wang, K. Zhang, J. Li, Y. Li, L. Zhang (Bytedance)]

[JVET-AC0135](https://jvet-experts.org/doc_end_user/current_document.php?id=12337) AHG9: Separate activation of color components for neural-network post-filter [C. Lin, Y. Li, Y.-K. Wang, K. Zhang, J. Li, L. Zhang (Bytedance)]

[JVET-AC0136](https://jvet-experts.org/doc_end_user/current_document.php?id=12338) AHG9: On additional NNPF purposes [C. Lin, Y.-K. Wang, K. Zhang, Y. Li, J. Li, L. Zhang (Bytedance)]

[JVET-AC0152](https://jvet-experts.org/doc_end_user/current_document.php?id=12356) AHG9: Bit-masking based representation of nnpfc\_purpose [J. Xu, Y.-K. Wang, L. Zhang, Y. Li (Bytedance)]

[JVET-AC0153](https://jvet-experts.org/doc_end_user/current_document.php?id=12357) AHG9: Bitdepth increase indication in the NNPFC SEI message [J. Xu, Y.-K. Wang, L. Zhang (Bytedance)]

[JVET-AC0154](https://jvet-experts.org/doc_end_user/current_document.php?id=12358) AHG9: Miscellaneous cleanups of the neural-network post-filter SEI messages [J. Xu, Y.-K. Wang, L. Zhang, C. Lin (Bytedance)]

[JVET-AC0174](https://jvet-experts.org/doc_end_user/current_document.php?id=12378) AHG9: Multiple input pictures for neural-network post-processing filter [Y. Li, Y.-K. Wang, C. Lin, J. Li, J. Xu, K. Zhang, L. Zhang (Bytedance)]

[JVET-AC0299](https://jvet-experts.org/doc_end_user/current_document.php?id=12504) AHG9: On indicating processing order in the NNPFC SEI messag [T. Shao, A. Arora, P. Yin, S. McCarthy, T. Lu, F. Pu, W. Husak (Dolby), Hendry, S. Kim (LGE)]

**SEI processing order SEI message**

[JVET-AC0058](https://jvet-experts.org/doc_end_user/current_document.php?id=12260) AHG9: On the SEI processing order SEI message [Y. He, M. Karczewicz (Qualcomm)]

**Preparation of FDAM/FDIS text of VVC 2nd Ed.**

[JVET-AC0275](https://jvet-experts.org/doc_end_user/current_document.php?id=12480) AHG9: Comments on VVC 2nd Ed. DAM1 [M. M. Hannuksela, M. Santamaria (Nokia)]

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

[JVET-AC0077](https://jvet-experts.org/doc_end_user/current_document.php?id=12279) AHG9/AHG15: On bitstreams that are potentially suboptimal for user viewing [M. M. Hannuksela, F. Cricri, H. Zhang (Nokia)]

[JVET-AC0088](https://jvet-experts.org/doc_end_user/current_document.php?id=12290) AHG9: Generative Face Video SEI Message [B. Chen, J. Chen, S. Wang, Y. Ye (Alibaba), S. Wang (CityU)]

[JVET-AC0122](https://jvet-experts.org/doc_end_user/current_document.php?id=12324) AHG9: Attenuation Map Information SEI for reducing energy consumption of displays [C.-H. Demarty, F. Aumont, E. Reinhard, L. Blondé, O. Le Meur, Z. Ameur, E. Francois (InterDigital)]

[JVET-AC0141](https://jvet-experts.org/doc_end_user/current_document.php?id=12343) AHG9: Alternative Picture Timing SEI [H.-B. Teo, J. Gao, C.-S. Lim, K. Abe (Panasonic)]

**Activities**

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

**Recommendations**

The AHG recommends to:

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

[JVET-AC0010](https://jvet-experts.org/doc_end_user/current_document.php?id=12250) JVET AHG report: Encoding algorithm optimization (AHG10) [P. de Lagrange, A. Duenas, R. Sjöberg, A. Tourapis (AHG chairs)]

1. **Related contributions**

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

It is noted that some of the mandates of AHG15 are related to test conditions and encoder optimization, and that some level of coordination would probably be beneficial.

***Test conditions***

**JVET-AC0078 – EE1-2.1: Updates on RPR encoder and post-filter**

The JVET-AB0102 contribution listed in the AHG10 report for previous JVET meeting proposed using a secondary anchor for EE1 super-resolution tests, that includes more intermediate resolutions (adding 4/5 ratio), GOP-based decision of coding resolution based on QP and high-frequency content in video source, and longer-tap filters for upscale.

This contribution reports test results in the EE1 context for these three aspects. GOP-based decision brings 0.7% luma gain in RA using only 1/1 and 1/2 resolutions, 1.23% when adding 2/3 and 4/5 resolutions, and longer tap filters bring around 0.3% chroma gain.

**JVET-AC0096 – AHG10/12: Suggestion for new CTC for RPR in VTM and ECM**

This contribution proposes four changes to RPR CTCs for both VTM and ECM, to exercise the tool more and adjust viewing conditions:

* Introduce more resolution changes: more often, and more resolution variants (switching every 32 frames and alternating between full resolution and one of the reduced resolutions: 4/5, 2/3, 1/2)
* Harmonize bitrates between the different resolutions by adjusting QPs depending on resolution (-6 for 1/2, -4 for 2/3, -2 for 4/5)
* Evaluate performance by using PSNR of upscaled reconstructed pictures only
* Test in RA configuration in addition to LDB

Performance numbers are reported using the proposed test conditions, for both VTM and ECM, and also for GOP-based RPR using VTM.

**JVET-AC0138 – Adjusting luma/chroma BD-rate balance in ECM**

This contribution proposes updates of ECM CTCs to balance luma/chroma gains compared to VTM. More specifically, adjustments to “CbQpOffset”, “CrQpOffset” and “LMCSOffset” parameters are proposed, depending on SDR/HDR and AI/others.

For AI / RA / LDB where imbalance was around +9% / +6% / +10% gain for chroma, imbalance is reduced to +0% / +4% / +5% respectively, with luma gain improvement of +0.9% / +0.2% / +0.16%.

**JVET-AC0149 – AHG10/12: Reduced I-frame QP for RA**

This contribution proposes to change the QP offset for intra frames in RA CTCs for both VTM and ECM: changing the “IntraQPOffset” config parameter from -3 to -4 is reported to bring 0.28%/2.62%/2.76% Y/Cb/Cr BDRate gains in the VTM, and similar gains for ECM. Subjective benefits are reported for static content, with no subjective problems for other content.

***Encoder optimizations***

**JVET-AC0139 – AHG10: Encoder Optimization of VTM Merge Functions**

In this contribution, an encoder only optimization on merge is proposed for VTM software. 97%, 99% and 92% encoding time is reported with 0.02%, 0.01% and 0.03% luma BD-rate loss for RA, LDB and LDP configurations, respectively.

Three merge functions on regular merge, sub-block merge and geometric merge are replaced with one unified merge function, which is said to remove redundant code. The code changes also include a new data structure to hold candidate information and prediction buffers, and memory optimization of prediction buffers. The maximum number of RDO checks after SATD-based candidate pruning is capped by a new configuration parameter (NumMaxRdCfg). The “Best Skip based Fast Merge” method is removed based on infrequent use in CTCs, marginal impact on encoding time, and losses in specific conditions.

1. **Recommendation**

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

[JVET-AC0011](https://jvet-experts.org/doc_end_user/current_document.php?id=12251) JVET AHG report: Neural network-based video coding (AHG11) [E. Alshina, S. Liu, A. Segall (co-chairs), F. Galpin, J. Li, T. Shao, H. Wang, Z. Wang, M. Wien, P. Wu (vice-chairs)]

***Common Test Conditions***

**Document**

The AHG released revised common test conditions as decided at the 28th meeting. The final version was uploaded as document JVET-AB2016 on December 24, 2022.

**Anchor Encoding**

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

***EE Coordination***

The AHG finalized, conducted and discussed the EE on NN based video coding. The final version of the EE description was uploaded to the document repository on November 14, 2022.

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

|  |  |  |
| --- | --- | --- |
| JVET-AC0023 | 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 |

**Anchor Encoding**

The performance of the VTM-11.0\_nnvc-3.0 anchor compared to VTM-11.0\_nnvc-2.0 is reported below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Random access Main 10** | | |  |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | |  |
|  | Y-PSNR | U-PSNR | V-PSNR | PSNR Overlap |
| Class A1 | 0.00% | 0.00% | 0.00% | 100% |
| Class A2 | 0.00% | 0.00% | 0.00% | 100% |
| Class B | -0.01% | -0.03% | -0.05% | 100% |
| Class C | 0.00% | 0.00% | 0.01% | 100% |
| Class E | - | - | - | - |
| **Overall** | 0.00% | -0.01% | -0.01% | 100% |
| Class D | 0.00% | 0.00% | 0.00% | 100% |
| Class F | 0.00% | -0.01% | 0.00% | 100% |
| Class H | - | - | - | - |
|  |  |  |  |  |
|  |  | **Low delay B Main 10** |  |  |
|  |  | **BD-rate Over VTM-11.0\_nnvc-2.0** |  |  |
|  | Y-PSNR | U-PSNR | V-PSNR | PSNR Overlap |
| Class A1 | - | - | - | - |
| Class A2 | - | - | - | - |
| Class B | 0.00% | -0.05% | 0.03% | 100% |
| Class C | -0.03% | 0.03% | 0.01% | 100% |
| Class E | 0.00% | 0.00% | 0.03% | 100% |
| **Overall** | -0.01% | -0.01% | 0.02% | 100% |
| Class D | 0.00% | 0.06% | 0.00% | 100% |
| Class F | -0.02% | 0.18% | 0.08% | 100% |
| Class H | - | - | - | - |
|  |  |  |  |  |
|  | **All Intra Main 10** | | |  |
|  | **BD-rate Over VTM-11.0\_nnvc-2.0** | | |  |
|  | Y-PSNR | U-PSNR | V-PSNR | PSNR Overlap |
| Class A1 | 0.00% | 0.00% | 0.00% | 100% |
| Class A2 | 0.00% | 0.00% | 0.00% | 100% |
| Class B | 0.00% | -0.01% | -0.02% | 100% |
| Class C | 0.00% | 0.00% | 0.00% | 100% |
| Class E | 0.00% | 0.00% | 0.00% | 100% |
| **Overall** | 0.00% | 0.00% | -0.01% | 100% |
| Class D | 0.00% | 0.00% | 0.00% | 100% |
| Class F | 0.00% | 0.00% | 0.00% | 100% |
| Class H | - | - | - | - |
|  |  |  |  |  |

*Note that the encoder and decoder run-time data was not accurate and has been removed.*

***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 29th meeting is provided below:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Title** | **Common Test Conditions** | **Results** | | | **Training Data** | |
| **RA** | **LDB** | **AI** | **CTC** | **Additional** |
| **Loop Filter** | | | | | | | |
| [JVET-AC0065](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12267) | Non-EE1: On flipping of input and output of model in NNVC filter set 0 | Yes | Yes | Yes | No | - | - |
| [JVET-AC0179](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12383) | AHG11: Swin-Transformer based In-Loop Filter for Natural and Screen Contents | No | No | No | No | - | DIV2K |
| **Post Filtering** | | | | | | | |
| [*JVET-AC0047*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12232) | *AHG9: On bugfix for NNPFC and NNPFA SEI messages* | No | No | No | No | - | - |
| [*JVET-AC0061*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12263) | *AHG9: Comments on Neural-network Post-filter Characteristics SEI Message* | No | No | No | No | - | - |
| [*JVET-AC0062*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12264) | *AHG9: On NNPFC SEI Message* | No | No | No | No | - | - |
| [*JVET-AC0074*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12276) | *AHG9: On the VVC use of the NNPFC SEI message for picture rate upsampling* | No | No | No | No | - | - |
| [*JVET-AC0075*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12277) | *AHG9: On NNPFC and NNPFA SEI messages for picture rate upsampling post-filter* | No | No | No | No | - | - |
| [*JVET-AC0076*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12278) | *AHG9/AHG15: On the NNPFC SEI message for machine analysis* | No | No | No | No | - | - |
| [*JVET-AC0085*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12287) | *AHG9: On neural-network post-filter characteristics (NNPFC) SEI message for temporal upsampling towards machine vision* | No | No | No | No | - | - |
| [*JVET-AC0127*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12329) | *AHG9: Combination of picture rate upsampling with other NNPF purposes* | No | No | No | No | - | - |
| [*JVET-AC0128*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12330) | *AHG9: On the signalling of complexity information in NNPFC SEI message* | No | No | No | No | - | - |
| [*JVET-AC0129*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12331) | *AHG9: On the NNPFC SEI message update and activation* | No | No | No | No | - | - |
| [*JVET-AC0131*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12333) | *AHG9: On NNPFC SEI message repetition* | No | No | No | No | - | - |
| [*JVET-AC0132*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12334) | *AHG9: On design for region-based neural-network post-filter SEI message* | No | No | No | No | - | - |
| [*JVET-AC0134*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12336) | *AHG9: Signalling of NNPF quality improvement* | No | No | No | No | - | - |
| [*JVET-AC0135*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12337) | *AHG9: Separate activation of color components for neural-network post-filter* | No | No | No | No | - | - |
| [*JVET-AC0136*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12338) | *AHG9: On additional NNPF purposes* | No | No | No | No | - | - |
| [*JVET-AC0152*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12356) | *AHG9: Bit-masking based representation of nnpfc\_purpose* | No | No | No | No | - | - |
| [*JVET-AC0153*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12357) | *AHG9: Bitdepth increase indication in the NNPFC SEI message* | No | No | No | No | - | - |
| [*JVET-AC0154*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12358) | *AHG9: Miscellaneous cleanups of the neural-network post-filter SEI messages* | No | No | No | No | - | - |
| [*JVET-AC0174*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12378) | *AHG9: Multiple input pictures for neural-network post-processing filter* | No | No | No | No | - | - |
| [*JVET-AC0208*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12412) | *AHG9: Summary of Proposals Related to Neural-Network Post-Filter SEI Messages* | No | No | No | No | - | - |
| **Inter-Prediction** | | | | | | | |
| [JVET-AC0090](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12292) | AHG11: Neural Network-based Reference CU Quality Enhancement for Motion Compensation Prediction | No | No | No | No | BVI-DVC | - |
| [JVET-AC0114](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12316) | AHG11: Deep Reference Frame Generation for Inter Prediction Enhancement | Yes | Yes | Yes | No |  | Vinmeo |
| **E2E Methods** | | | | | | | |
| [JVET-AC0091](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12293) | AHG11: Fourier Series and Laplacian Noise-based Quantization Error Compensation for End-to-End Learning-based Image Compression | No | No | No | No | - | - |

1. **Input contributions**

There are 69 input contributions related to the AHG mandates. Forty-three of the contributions are part of the EE activity, while the remaining 26 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-AC0023](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12258) | EE1: Summary report 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) |
| **EE Technology** | | |
| [JVET-AC0051](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12236) | EE1-2.3: RPR-Based Super-Resolution Guided by Partition Information Combined with GOP Level Adaptive Resolution | [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), [J. Nam](mailto:junghak.nam@lge.com), [S. Yoo](mailto:sunmi.yoo@lge.com), [J. Lim](mailto:jaehyun.lim@lge.com), [S. H. Kim (LGE)](mailto:seunghwan3.kim@lge.com) |
| [JVET-AC0052](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12237) | EE1-2.4: CNN filter Based on RPR-based SR Combined with GOP Level Adaptive Resolution | [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), [J. Nam](mailto:junghak.nam@lge.com), [S. Yoo](mailto:sunmi.yoo@lge.com), [J. Lim](mailto:jaehyun.lim@lge.com), [S. H. Kim (LGE)](mailto:seunghwan3.kim@lge.com) |
| [JVET-AC0055](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12241) | EE1-1.11: Content-adaptive post-filter | [M. Santamaria](mailto:maria.santamaria_gomez@nokia.com), R. Yang, F. Cricri, J. Lainema, H. Zhang, R. G. Youvalari, M. M. Hannuksela (Nokia) |
| [JVET-AC0056](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12252) | EE1-3.1 CompressAI models integration using SADL | [F. Galpin](mailto:franck.galpin@interdigital.com), F. Lefebvre, F. Racape (InterDigital) |
| [JVET-AC0063](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12265) | EE1-1.3: 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-AC0064](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12266) | EE1-1.4: 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-AC0078](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12280) | EE1-2.1: Updates on RPR encoder and post-filter | [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) |
| [JVET-AC0089](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12291) | EE1-1.5: Combined intra and inter models for luma and chroma | [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-AC0106](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12308) | EE1-1.10: Complexity Reduction on Neural-Network Loop Filter | [J. N. Shingala](mailto:jay.shingala@ittiam.com), A. Shyam, A. Suneia, S. P. Badya (Ittiam), [T. Shao](mailto:tshao@dolby.com), A. Arora, [P. Yin](mailto:pyin@dolby.com), Sean McCarthy (Dolby) |
| [JVET-AC0116](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12318) | EE1-3.2 : 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) |
| [JVET-AC0118](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12320) | EE1-1.8: 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) |
| [JVET-AC0143](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12345) | EE1-1.6: NN chroma model without partitioning input | [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](mailto:ruoyang.yu@ericsson.com) |
| [JVET-AC0155](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12359) | EE1-1.9: Reduced complexity CNN-based in-loop filtering | [S. Eadie](mailto:seadie@qti.qualcomm.com), [M. Coban](mailto:mcoban@qti.qualcomm.com), [M. Karczewicz (Qualcomm)](mailto:martak@qti.qualcomm.com) |
| [JVET-AC0177](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12381) | EE1-1.7: Deep In-Loop Filter with Additional Input Information | [Y. Li](mailto:yue.li@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com) |
| [JVET-AC0194](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12398) | EE1-1.1: More refinements on NN based in-loop filter with a single model | [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-AC0195](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12399) | EE1-1.2: encoder-only optimization for NN based in-loop filter with a single model | [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-AC0196](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12400) | EE1-2.2: 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) |
| **EE Related** | | |
| [JVET-AC0066](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12268) | EE1-related: Improvement on EE1-1.7 | [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-AC0081](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12283) | EE1-related: A CNN Filter for RPR-Based Super-Resolution Using Wavelet Decomposition Combined with GOP Level Adaptive Resolution | [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), [J. Nam](mailto:junghak.nam@lge.com), [S. Yoo](mailto:sunmi.yoo@lge.com), [J. Lim](mailto:jaehyun.lim@lge.com), [S. H. Kim (LGE)](mailto:seunghwan3.kim@lge.com) |
| [JVET-AC0126](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12328) | EE1-related: Reduced complexity through channel redistribution in NN head | [P. Wennersten](mailto:per.wennersten@ericsson.com), [J. Ström](mailto:jacob.strom@ericsson.com), [D. Liu (Ericsson)](mailto:du.liu@ericsson.com) |
| [JVET-AC0156](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12360) | [EE1-related] RTNN: An In-loop Filter Based on Resblock and Transformer | [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-AC0178](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12382) | EE1-related: 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) |
| [JVET-AC0197](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12401) | EE1-1.1-related: More refinements on NN based in-loop filter | [R. Chang](mailto:renjiechang@tencent.com), [L. Wang](mailto:liqiangwang@tencent.com), [X. Xu](mailto:xiaozhongxu@tencent.com), [S. Liu (Tencent)](mailto:shanl@tencent.com) |
| **Cross Checks** | | |
| [JVET-AC0209](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12413) | Crosscheck of JVET-AC0177 (EE1-1.7: Deep In-Loop Filter with Additional Input Information) | [C. Zhou (vivo)](mailto:chuan.zhou@vivo.com) |
| [JVET-AC0215](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12419) | Crosscheck of JVET-AC0118 (EE1-1.8: QP-based loss function design for NN-based in-loop filter) | [C. Lin (Bytedance)](mailto:linchaoyi.cy@bytedance.com) |
| [JVET-AC0221](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12425) | Crosscheck of JVET-AC0196 (EE1-2.2: GOP Level Adaptive Resampling with CNN-based Super Resolution) | [D. Liu (Ericsson)](mailto:du.liu@ericsson.com) |
| [JVET-AC0222](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12426) | Crosscheck of JVET-AC0051 (EE1-2.3: RPR-Based Super-Resolution Guided by Partition Information Combined with GOP Level Adaptive Resolution) | [D. Liu (Ericsson)](mailto:du.liu@ericsson.com) |
| [JVET-AC0223](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12427) | Crosscheck of JVET-AC0052 (EE1-2.4: CNN filter Based on RPR-based SR Combined with GOP Level Adaptive Resolution) | [D. Liu (Ericsson)](mailto:du.liu@ericsson.com) |
| [JVET-AC0234](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12438) | Cross-check of JVET-AC0143 (EE1-1.6: NN chroma model without partitioning input) | [M. Santamaria (Nokia)](mailto:maria.santamaria_gomez@nokia.com) |
| [JVET-AC0238](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12442) | Crosscheck of JVET-AC0116 (EE1-3.2: neural network-based intra prediction with learned mapping to VVC intra prediction modes) | [Y. Li (Bytedance)](mailto:yue.li@bytedance.com) |
| [JVET-AC0241](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12445) | Crosscheck of JVET-AC0106 (EE1-1.10: Complexity Reduction on Neural-Network Loop Filter) | [Y. Li (Bytedance)](mailto:yue.li@bytedance.com) |
| [JVET-AC0254](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12459) | Crosscheck of JVET-AC0155 (EE1-1.9: Reduced complexity CNN-based in-loop filtering) | [Y. Li (Bytedance)](mailto:yue.li@bytedance.com) |
| [JVET-AC0257](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12462) | Crosscheck of JVET-AC0056(EE1-3.1 CompressAI models integration using SADL) | [T. Chujoh (Sharp)](mailto:chujoh.takeshi@sharp.co.jp) |
| [JVET-AC0270](https://jvet-experts.org/doc_end_user/current_document.php?id=12475) | Crosscheck of EE1-3.2 (JVET-AC0116 : neural network-based intra prediction with learned mapping to VVC intra prediction modes) | [M. Damghanian](mailto:mitra.damghanian@ericsson.com), [J. Ström](mailto:jacob.strom@ericsson.com) |
| [JVET-AC0271](https://jvet-experts.org/doc_end_user/current_document.php?id=12476) | Crosscheck of JVET-AC0063 (EE1-1.3: On chroma order adjustment in NNLF) | [R. Chang (Tencent)](mailto:renjiechang@tencent.com) |
| [JVET-AC0272](https://jvet-experts.org/doc_end_user/current_document.php?id=12477) | Crosscheck of JVET-AC0064 (EE1-1.4: On adjustment of residual for NNLF) | [R. Chang (Tencent)](mailto:renjiechang@tencent.com) |
| [JVET-AC0273](https://jvet-experts.org/doc_end_user/current_document.php?id=12478) | Crosscheck of JVET-AC0089 (EE1-1.5: Combined intra and inter models for luma and chroma) | [R. Chang (Tencent)](mailto:renjiechang@tencent.com) |
| [JVET-AC0274](https://jvet-experts.org/doc_end_user/current_document.php?id=12479) | Crosscheck of JVET-AC0116 (EE1-3.2.2: Low-complexity version of the neural network-based intra prediction mode in 16-bit signed integer) | [T. Shao (Dolby)](mailto:Tong.Shao@dolby.com) |
| [JVET-AC0278](https://jvet-experts.org/doc_end_user/current_document.php?id=12483) | Crosscheck of JVET-AC0194 (EE1-1.1: More refinements on NN based in-loop filter with a single model) | [Z. Xie(OPPO)](mailto:xiezhihuang@oppo.com) |
| [JVET-AC0279](https://jvet-experts.org/doc_end_user/current_document.php?id=12484) | Crosscheck of JVET-AC0195 (EE1-1.2: encoder-only optimization for NN based in-loop filter with a single model) | [Z. Xie(OPPO)](mailto:xiezhihuang@oppo.com) |
| [JVET-AC0286](https://jvet-experts.org/doc_end_user/current_document.php?id=12491) | Crosscheck of JVET-AC0064 (EE1-1.4: On adjustment of residual for NNLF) | [K. Jia (Bytedance)](mailto:%20jiake@bytedance.com) |

***Non-EE Input Contributions***

|  |  |  |
| --- | --- | --- |
| **Reporting** | | |
| [JVET-AC0011](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12251) | JVET AHG report: Neural network-based video coding (AHG11) | E. Alshina, S. Liu, A. Segall (co-chairs), F. Galpin, J. Li, T. Shao, H. Wang, Z. Wang, M. Wien, P. Wu (vice-chairs) |
| **Loop Filtering** | | |
| [JVET-AC0065](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12267) | Non-EE1: On flipping of input and output of model in NNVC filter set 0 | [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-AC0179](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12383) | AHG11: Swin-Transformer based In-Loop Filter for Natural and Screen Contents | [J. Li](mailto:lijunru@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [L. Zhang](mailto:lizhang.idm@bytedance.com), M. Wang (Bytedance) |
| **Post Filtering** | | |
| [*JVET-AC0047*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12232) | *AHG9: On bugfix for NNPFC and NNPFA SEI messages* | [*T. Chujoh*](mailto:chujoh.takeshi@sharp.co.jp)*, Y. Yasugi, T. Ikai (Sharp)* |
| [*JVET-AC0061*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12263) | *AHG9: Comments on Neural-network Post-filter Characteristics SEI Message* | [*S. Deshpande (Sharp)*](mailto:sdeshpande@sharplabs.com) |
| [*JVET-AC0062*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12264) | *AHG9: On NNPFC SEI Message* | [*S. Deshpande*](mailto:sdeshpande@sharplabs.com)*, A. Sidiya (Sharp)* |
| [*JVET-AC0074*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12276) | *AHG9: On the VVC use of the NNPFC SEI message for picture rate upsampling* | [*M. M. Hannuksela (Nokia)*](mailto:miska.hannuksela@nokia.com) |
| [*JVET-AC0075*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12277) | *AHG9: On NNPFC and NNPFA SEI messages for picture rate upsampling post-filter* | [*M. M. Hannuksela (Nokia)*](mailto:miska.hannuksela@nokia.com) |
| [*JVET-AC0076*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12278) | *AHG9/AHG15: On the NNPFC SEI message for machine analysis* | [*M. M. Hannuksela*](mailto:miska.hannuksela@nokia.com)*, F. Cricri, J. I. Ahonen, H. Zhang (Nokia)* |
| [*JVET-AC0085*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12287) | *AHG9: On neural-network post-filter characteristics (NNPFC) SEI message for temporal upsampling towards machine vision* | [*S. Wang*](mailto:srwang3-c@my.cityu.edu.hk)*,* [*J. Chen*](mailto:jiechen.cj@alibaba-inc.com)*,* [*Y. Ye (Alibaba)*](mailto:yan.ye@alibaba-inc.com)*,* [*S. Wang (CityU)*](mailto:shiqwang@cityu.edu.hk) |
| [*JVET-AC0127*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12329) | *AHG9: Combination of picture rate upsampling with other NNPF purposes* | [*Y.-K. Wang*](mailto:yekui.wang@bytedance.com)*,* [*J. Xu*](mailto:xujizheng@bytedance.com)*,* [*Y. Li*](mailto:yue.li@bytedance.com)*,* [*L. Zhang*](mailto:lizhang.idm@bytedance.com)*,* [*K. Zhang*](mailto:zhangkai.video@bytedance.com)*,* [*J. Li*](mailto:lijunru@bytedance.com)*,* [*C. Lin (Bytedance)*](mailto:linchaoyi.cy@bytedance.com) |
| [*JVET-AC0128*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12330) | *AHG9: On the signalling of complexity information in NNPFC SEI message* | [*Hendry*](mailto:dr.hendry@lge.com)*, S. Kim (LGE)* |
| [*JVET-AC0129*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12331) | *AHG9: On the NNPFC SEI message update and activation* | [*Hendry*](mailto:dr.hendry@lge.com)*,* [*J. Nam*](mailto:junghak.nam@lge.com)*, H. Jang, S. Kim, J. Lim (LGE)* |
| [*JVET-AC0131*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12333) | *AHG9: On NNPFC SEI message repetition* | [*Hendry*](mailto:dr.hendry@lge.com)*,* [*J. Nam*](mailto:junghak.nam@lge.com)*, H. Jang, S. Kim, J. Lim (LGE)* |
| [*JVET-AC0132*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12334) | *AHG9: On design for region-based neural-network post-filter SEI message* | [*Hendry*](mailto:dr.hendry@lge.com)*,* [*J. Nam*](mailto:junghak.nam@lge.com)*, H. Jang, S. Kim, J. Lim (LGE)* |
| [*JVET-AC0134*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12336) | *AHG9: Signalling of NNPF quality improvement* | [*C. Lin*](mailto:linchaoyi.cy@bytedance.com)*,* [*Y.-K. Wang*](mailto:yekui.wang@bytedance.com)*,* [*K. Zhang*](mailto:zhangkai.video@bytedance.com)*,* [*J. Li*](mailto:lijunru@bytedance.com)*,* [*Y. Li*](mailto:yue.li@bytedance.com)*,* [*L. Zhang (Bytedance)*](mailto:lizhang.idm@bytedance.com) |
| [*JVET-AC0135*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12337) | *AHG9: Separate activation of color components for neural-network post-filter* | [*C. Lin*](mailto:linchaoyi.cy@bytedance.com)*,* [*Y. Li*](mailto:yue.li@bytedance.com)*,* [*Y.-K. Wang*](mailto:yekui.wang@bytedance.com)*,* [*K. Zhang*](mailto:zhangkai.video@bytedance.com)*,* [*J. Li*](mailto:lijunru@bytedance.com)*,* [*L. Zhang (Bytedance)*](mailto:lizhang.idm@bytedance.com) |
| [*JVET-AC0136*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12338) | *AHG9: On additional NNPF purposes* | [*C. Lin*](mailto:linchaoyi.cy@bytedance.com)*,* [*Y.-K. Wang*](mailto:yekui.wang@bytedance.com)*,* [*K. Zhang*](mailto:zhangkai.video@bytedance.com)*,* [*Y. Li*](mailto:yue.li@bytedance.com)*,* [*J. Li*](mailto:lijunru@bytedance.com)*,* [*L. Zhang (Bytedance)*](mailto:lizhang.idm@bytedance.com) |
| [*JVET-AC0152*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12356) | *AHG9: Bit-masking based representation of nnpfc\_purpose* | [*J. Xu*](mailto:xujizheng@bytedance.com)*,* [*Y.-K. Wang*](mailto:yekui.wang@bytedance.com)*,* [*L. Zhang*](mailto:lizhang.idm@bytedance.com)*,* [*Y. Li (Bytedance)*](mailto:yue.li@bytedance.com) |
| [*JVET-AC0153*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12357) | *AHG9: Bitdepth increase indication in the NNPFC SEI message* | [*J. Xu*](mailto:xujizheng@bytedance.com)*,* [*Y.-K. Wang*](mailto:yekui.wang@bytedance.com)*,* [*L. Zhang (Bytedance)*](mailto:lizhang.idm@bytedance.com) |
| [*JVET-AC0154*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12358) | *AHG9: Miscellaneous cleanups of the neural-network post-filter SEI messages* | [*J. Xu*](mailto:xujizheng@bytedance.com)*,* [*Y.-K. Wang*](mailto:yekui.wang@bytedance.com)*,* [*L. Zhang*](mailto:lizhang.idm@bytedance.com)*,* [*C. Lin (Bytedance)*](mailto:linchaoyi.cy@bytedance.com) |
| [*JVET-AC0174*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12378) | *AHG9: Multiple input pictures for neural-network post-processing filter* | [*Y. Li*](mailto:yue.li@bytedance.com)*,* [*Y.-K. Wang*](mailto:yekui.wang@bytedance.com)*,* [*C. Lin*](mailto:linchaoyi.cy@bytedance.com)*,* [*J. Li*](mailto:lijunru@bytedance.com)*,* [*J. Xu*](mailto:xujizheng@bytedance.com)*,* [*K. Zhang*](mailto:zhangkai.video@bytedance.com)*,* [*L. Zhang (Bytedance)*](mailto:lizhang.idm@bytedance.com) |
| [*JVET-AC0208*](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12412) | *AHG9: Summary of Proposals Related to Neural-Network Post-Filter SEI Messages* | [*S. Deshpande (Sharp)*](mailto:sdeshpande@sharplabs.com)*,* [*Y.-K. Wang (Bytedance)*](mailto:yekui.wang@bytedance.com) |
| **Inter Prediction** | | |
| [JVET-AC0090](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12292) | AHG11: Neural Network-based Reference CU Quality Enhancement for Motion Compensation Prediction | [Yanhan Chu](mailto:yanhan_c@mail.sdu.edu.cn), [Zhikui Wang](mailto:wangzhikui@hisense.com), [Wen Zhang](mailto:zhangwen12@hisense.com), [Shuai Li(Hisense Visual Technology)](mailto:shuaili@sdu.edu.cn) |
| [JVET-AC0114](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12316) | 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), [X. Xu](mailto:xiaozhongxu@tencent.com), [S. Liu (Tencent)](mailto:shanl@tencent.com) |
| **E2E Methods** | | |
| [JVET-AC0091](file:////Users/shanliu-sl/Documents/contribution/jvet29ac/current_document.php%3fid=12293) | AHG11: Fourier Series and Laplacian Noise-based Quantization Error Compensation for End-to-End Learning-based Image Compression | [Shiqi Jiang](mailto:shiqijiang@mail.sdu.edu.cn), [Zhikui Wang](mailto:wangzhikui@hisense.com), [Wen Zhang](mailto:zhangwen12@hisense.com), [Shuai Li (Hisense Visual Technology)](mailto:shuaili@sdu.edu.cn) |

1. **Recommendations**

The AHG recommends:

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

[JVET-AC0012](https://jvet-experts.org/doc_end_user/current_document.php?id=12253) 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)]

1. **Activities**

The primary activity of the AHG was the “Exploration experiment on enhanced compression beyond VVC capability” (JVET-AB2024). The combined improvements of the ECM-7.0 over VTM-11.0ecm anchorfor AI, RA and LB configurations are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | All Intra Main10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -8.67% | -17.52% | -24.11% | 617% | 348% |
| Class A2 | -12.88% | -21.39% | -24.64% | 576% | 348% |
| Class B | -7.62% | -20.59% | -18.79% | 630% | 392% |
| Class C | -8.12% | -12.50% | -12.91% | 635% | 411% |
| Class E | -9.72% | -17.97% | -16.91% | 619% | 435% |
| Overall | **-9.13%** | **-17.98%** | **-19.03%** | **618%** | **387%** |
| Class D | -6.68% | -10.00% | -10.01% | 624% | 416% |
| Class F | -18.45% | -26.68% | -26.24% | 444% | 397% |
| Class TGM | -31.47% | -38.80% | -37.65% | 386% | 367% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Random Access Main 10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -19.46% | -21.91% | -29.05% | 583% | 646% |
| Class A2 | -21.45% | -28.23% | -31.25% | 586% | 894% |
| Class B | -17.65% | -28.04% | -25.93% | 614% | 830% |
| Class C | -19.26% | -22.56% | -22.39% | 605% | 892% |
| Class E |  |  |  |  |  |
| Overall | **-19.20%** | **-25.39%** | **-26.67%** | **600%** | **817%** |
| Class D | -20.25% | -23.80% | -24.06% | 605% | 959% |
| Class F | -22.05% | -29.05% | -29.34% | 520% | 556% |
| Class TGM | -30.00% | -35.63% | -35.75% | 556% | 438% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Low Delay B Main 10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -16.16% | -30.50% | -28.78% | 529% | 695% |
| Class C | -17.44% | -25.63% | -25.66% | 552% | 714% |
| Class E | -14.58% | -22.41% | -23.17% | 552% | 483% |
| Overall | **-16.19%** | **-26.85%** | **-26.34%** | **542%** | **640%** |
| Class D | -19.65% | -26.43% | -26.39% | 554% | 850% |
| Class F | -20.49% | -29.26% | -27.88% | 514% | 496% |
| Class TGM | -28.89% | -35.39% | -35.06% | 524% | 401% |

The rate reduction for natural sequences over VTM in RA configuration for {Y, U, V} increased from {-18.50%, -24.47%, -25.65%} to {-19.20%, -25.39%, -26.67%}. For SCC sequences (class TGM) the rate reduction for RA configuration increased from {-29.30%, -33.62%, -34.08%} to {-30.00%, -35.63%, -35.57%}.

1. **Contributions**

In addition to 32 EE2 contributions, 57 (comparing to 45 last meeting) EE2-related and ECM-related contributions were received which can be subdivided as follows:

***Intra (24)***

JVET-AC0048, "Non-EE: Modification of TIMD", Y. Yasugi, T. Ikai (Sharp)

JVET-AC0053, "AHG12: Simplified linear model solver", J. Lainema, P. Astola, A. Aminlou, R. G. Youvalari (Nokia)

JVET-AC0067, "Non-EE2: Modifications on template-based multiple reference line intra prediction", L. Xu, Y. Yu, H. Yu, D. Wang (OPPO)

JVET-AC0068, "Non-EE2: multi-candidate IntraTMP", F. Wang, L. Zhang, Y. Yu, H. Yu, D. Wang (OPPO)

JVET-AC0069, "Non-EE2: Intra Template-Matching Prediction Fusion", L. Zhang, F. Wang, Y. Yu, H. Yu, D. Wang (OPPO)

JVET-AC0070, "Non-EE2: combination of JVET-AC0068 and JVET-AC0069", F. Wang, L. Zhang, Y. Yu, H. Yu, D. Wang (OPPO)

JVET-AC0087, "Non-EE2: Intra TMP with half-pel precision", X. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba)

JVET-AC0097, "Non-EE2: SGPM combined with IntraTMP", C. Fang, S. Peng, D. Jiang, J. Lin, X. Zhang (Dahua)

JVET-AC0107, "AHG12: Fusion of Intra Template Matching", J. R. Arumugam, A. Natesan, V. Valvaiker, J. N. Shingala (Ittiam), T. Lu, P. Yin, F. Pu, T. Shao, A. Arora, S. McCarthy (Dolby)

JVET-AC0108, "Non-EE2: Extend search area in Intra Template Matching Prediction (IntraTMP)", J.-Y. Huo, H.-Q. Du, H.-L. Zhang, Z.-Y. Zhang, W.-H. Qiao, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.)

JVET-AC0109, "Non-EE2: Intra template matching (Intra TMP) based on linear filter model", J.-Y. Huo, W.-H. Qiao, X. Hao, Z.-Y. Zhang, H.-Q. Du, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.)

JVET-AC0110, "Non-EE2: A Fusion method of Intra Template Matching Prediction (Intra TMP)", J.-Y. Huo, H.-Q. Du, H.-L. Zhang, W.-H. Qiao, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.)

JVET-AC0111, "Non-EE2: Combination of JVET-AC0108, JVET-AC0109 and JVET-AC0110 for Intra TMP", J.-Y. Huo, W.-H. Qiao, H.-Q. Du, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.), H. Yuan (Shandong Univ.)

JVET-AC0120, "Non-EE2: template based intra MPM list construction", C. Zhou, Z. Lv, J. Zhang (vivo)

JVET-AC0121, "EE2-related: on GL-CCCM improvement (test 1.12a)", P Bordes, K Naser, F Galpin, E Francois (InterDigital), RG Youvalari, P Astola, J Lainema (Nokia)

JVET-AC0146, "AHG12: Filtered Template Matching based Intra Prediction (FTMP)", R. G. Youvalari, D. Bugdayci Sansli, P. Astola, J. Lainema (Nokia)

JVET-AC0148, "Non-EE2: CCCM using multiple downsampling filters", Y.-J. Chang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AC0161, "Non-EE2: IBC adaptation for coding of natural content", B. Ray, H. Wang, C.-C. Chen, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AC0170, "Non-EE2: Fuse intra template matching prediction with intra prediction", Y. Wang, K. Zhang, L. Zhang (Bytedance)

JVET-AC0172, "Non-EE2: IBC with fractional block vectors", W. Chen, X. Xiu, C. Ma, H.-J. Jhu, C.-W. Kuo, N. Yan, X. Wang (Kwai)

JVET-AC0175, "Non-EE2: Local-Boosting Cross-Component Prediction", K. Zhang, L. Zhang, Z. Deng (Bytedance)

JVET-AC0176, "Non-EE2: Non-Local Cross-Component Prediction", K. Zhang, L. Zhang, Z. Deng (Bytedance)

JVET-AC0191, "Non-EE2: Modification to MPM list derivation", H. Wang, Y.-J. Chang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AC0198, "Non-EE2: IntraTMP with multiple modes", P.-H. Lin, J.-L- Lin, V. Seregin, M. Karczewicz (Qualcomm)

***Inter (10)***

JVET-AC0093, "EE2-related: test 2.5a and PROF improvement", F. Galpin, A. Robert, K. Naser (InterDigital)

JVET-AC0103, "Non-EE2: AMVP mode with subblock-based temporal motion vector prediction", R.-L. Liao, J. Chen, Y. Ye, X. Li (Alibaba)

JVET-AC0164, "Non-EE2: Improvements on local illumination compensation in ECM7.0", X. Xiu, N. Yan, H.-J. Jhu, W. Chen, C.-W. Kuo, C. Ma, X. Wang (Kwai)

JVET-AC0182, "Non-EE2: High-Precision MV Refinement for BDOF", M. Salehifar, Y. He, K. Zhang, L. Zhang (Bytedance)

JVET-AC0186, "EE2-related: EE2-2.7 with further encoder optimizations", L. Zhao, K. Zhang, L. Zhang (Bytedance)

JVET-AC0187, "Non-EE2: Template matching-based subblock motion refinement", L. Zhao, K. Zhang, Z. Deng, L. Zhang (Bytedance)

JVET-AC0124, "Non-EE2: Bi-prediction with block-level only out-of-bound management", F. Le F. Le Léannec, M. Blestel, P. Andrivon, M. Radosavljević (Xiaomi)

JVET-AC0213, "Non-EE2: SbTMVP with MMVD", L.-F. Chen, R. Chernyak, X. Zhao, X. Xu, S. Liu (Tencent)

JVET-AC0239, "Non-EE2: Prediction of MVD magnitude suffix bins", A. Filippov, V. Rufitskiy, K. Suverov (Ofinno)

JVET-AC0276, "EE2-related: Complexity reduction for Decoder Side Control Point Motion Vector Refinement (test 2.3b)", M. Bestel, F. Le Léannec, P. Andrivon, M. Radosavljević (Xiaomi), H. Huang, Y. Zhang, Z. Zhang, C.-C. Chen, V. Seregim, M. Karczewicz (Qualcomm)

***Screen Content Coding (13)***

JVET-AC0059, "AHG12: TMP using Reconstruction-Reordered for screen content coding (RR-TMP)", Jung-Kyung Lee, Damian Ruiz Coll, Vikas Warudkar, (Ofinno)

JVET-AC0140, "EE2-related: Additional results for test EE2-3.3", A. Filippov, V. Rufitskiy (Ofinno)

JVET-AC0142, "Non-EE2: Cross-Component Discrete Mapping Model for Screen Content Coding", B. Vishwanath, K. Zhang, L. Zhang (Bytedance)

JVET-AC0159, "Non-EE2: IBC MBVD list derivation", Z. Zhang, P. Nikitin, H. Huang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AC0160, "Non-EE2: On the condition of OBMC", X. Li (Google)

JVET-AC0165, "EE2-related: Additional results to EE2-3.4", Damian Ruiz Coll, Jung-Kyung Lee, Vikas Warudkar, (Ofinno)

JVET-AC0167, "AHG12: Using block vector derived from IntraTMP as an IBC candidate for the current block", Jung-Kyung Lee, Damian Ruiz Coll, Vikas Warudkar, (Ofinno)

JVET-AC0169, "Non-EE2: Template Matching for RR-IBC", C.-C. Chen, H. Huang, Z. Zhang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AC0189, "Non-EE2: SGPM without blending", Z. Deng, L. Zhang, K. Zhang, L. Zhao (Bytedance)

JVET-AC0190, "Non-EE2: On OBMC", Z. Deng, K. Zhang, L. Zhang (Bytedance)

JVET-AC0192, "Non-EE2: Temporal block vector prediction", N. Zhang, K. Zhang, L. Zhang (Bytedance)

JVET-AC0193, "Non-EE2: Copy-Padding for IBC", N. Zhang, K. Zhang, L. Zhang (Bytedance)

JVET-AC0203, "AHG12: On LMCS luma mapping in template processing for IBC TM-based tools", A. Filippov, V. Rufitskiy (Ofinno)

***In Loop Filters (1)***

JVET-AC0173, "Non-EE2: ALF classification based on residual data", I. Jumakulyyev, N. Hu, Z. Zhang, V. Seregin, M. Karczewicz, H. Huang (Qualcomm)

***Transform (2)***

JVET-AC0163, "Non-EE2: On non-separable primary transform for intra blocks", N. Yan, X. Xiu, W. Chen, H.-J. Jhu, C.-W. Kuo, C. Ma, X. Wang (Kwai)

JVET-AC0205, "EE2-related: Modifications of MTS and LFNST for IntraTMP", R. G. Youvalari, D. Bugdayci Sansli, J. Lainema (Nokia)

***Entropy (1)***

JVET-AC0200, "AhG12 Dynamic CABAC models", F. Lo Bianco, F. Galpin, E. Francois (InterDigital)

***Other (6)***

JVET-AC0095, "AHG12: Fix related to pixel copy in motion compensation", K. Andersson, R. Yu (Ericsson)

JVET-AC0096, "AHG10/12: Suggestion for new CTC for RPR in VTM and ECM", K. Andersson, R. Yu (Ericsson)

JVET-AC0123, "MTT maximum depth correction for class B sequences at QP 22 in ECM", G. Laroche, P. Onno (Canon)

JVET-AC0138, "Adjusting luma/chroma BD-rate balance in ECM", E. Francois, Y. Chen, C. Salmon-Legagneur (InterDigital)

JVET-AC0149, "AHG10/12: Reduced I-frame QP for RA", K. Andersson, P. Wennersten (Ericsson)

JVET-AC0171, "AHG12: ECM-6 Tool Off Tests", X. Li (Google)

1. Recommendations

The AHG recommends to:

* To review all the related contributions.

It is verbally reported that, by shifting chroma gain to luma, more than 20% might be achieved in RA.

Most gain was achieved recently in intra (which is also increasing RA gain), and this seems again to be the tendency from the contributions at this meeting.

[JVET-AC0013](https://jvet-experts.org/doc_end_user/current_document.php?id=12254) 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)]

The group focused its efforts on updating the technical report during the intermeeting period in preparation of issuing a ballot. The editors met on a biweekly cadence throughout the period. The editorial process consisted of resolving many editorial issues without any major changes in the structure of the document.

1. **Related contributions**

Five contributions related to AHG13 were identified as of 1/11/2023:

* One was the AHG report:
  + JVET-AC0013 JVET AHG report: Film grain technologies (AHG13)
* Two were related to the technical report and explore film grain testing:
  + JVET-AC0151 Proposed text: Film grain synthesis technology for video applications (CD draft)
  + JVET-AC0043 AHG4: Report on AHG meeting on subjective test preparation for film grain synthesis
  + JVET-AC0043 AHG4: Report on AHG meeting on subjective test preparation for film grain synthesis
  + JVET-AC0199 On Film Grain Synthesis Subjective Evaluation

***Contributions***

**JVET-AC0151 Proposed text: Film grain synthesis technology for video applications (CD draft)**

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.

**JVET-AC0199 On Film Grain Synthesis Subjective Evaluation**

The initial version was incomplete (rejected by chair), but an update became available by the beginning of the meeting.

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.

[JVET-AC0014](https://jvet-experts.org/doc_end_user/current_document.php?id=12255) JVET AHG report: NNVC software development (AHG14) [S. Eadie, F. Galpin, Y. Li, L. Wang, Z. Xie (AHG chairs)]

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

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

NNVC-3.0 anchor at <https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/VVCSoftware_VTM> is used for NNVC performance evaluation.

***Software changes***

The following changes were integrated into NNVC-3.0:

* add missing license headers in training scripts [MR17]
* JVET-AB0083: EE1-1.8: More refinements on NN based in-loop filter with a single model (Test 1) [MR 18]
* JVET-AB0068: EE1-1.6: RDO Considering Deep In-Loop Filtering [MR 20]
* Update to SADL v3.0 [MR 19]
* JVET-AB0053: EE1-1.2: Removed attention branch and partitioning input from filter set 1. [MR 21] [MR 23]

The following fixes:

* Add missing license headers in training scripts
* refactor code of filter set #1: correct variable name, remove constant [MR 22]

NNVC 3.0 was tagged December 1st 2022.

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

NNVC 2.0 was tagged August 4th 2022 (add deblocking in RDO).

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

***CTC performance***

In this section, NNVC test results following NNVC CTC configuration descried in JVET-AB2016 are summarized.

**Filter set 0**

The following tables show NNVC-3.0 NN-based filter set #0 (int16 precision) performance over NNVC-3.0 anchor.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | |
|  | **BD-rate Over NNVC-3.0 anchor** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -9.26% | -13.15% | -18.91% | 160% | 56967% |
| Class A2 | -9.51% | -19.53% | -14.74% | 151% | 51368% |
| Class B | -8.59% | -19.05% | -20.78% | 158% | 57030% |
| Class C | -8.62% | -20.95% | -20.76% | 136% | 48210% |
| Class E |  |  |  |  |  |
| **Overall** | -8.91% | -18.47% | -19.19% | 151% | 53391% |
| Class D | -9.74% | -20.01% | -21.14% | 141% | 40159% |
| Class F | -3.79% | -12.18% | -11.09% | 191% | 23529% |
|  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | |
|  | **BD-rate Over NNVC-3.0 anchor** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -7.78% | -18.19% | -20.35% | 144% | 56136% |
| Class C | -8.26% | -20.07% | -20.45% | 124% | 45792% |
| Class E | -8.89% | -15.66% | -16.70% | 217% | 47327% |
| **Overall** | -8.22% | -19.03% | -20.39% | 152% | 50260% |
| Class D | -9.61% | -18.73% | -20.58% | 133% | 39906% |
| Class F | -4.19% | -10.69% | -8.72% | 195% | 23635% |
|  |  |  |  |  |  |
|  | **All Intra Main10** | | | | |
|  | **BD-rate Over NNVC-3.0 anchor** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -6.12% | -14.26% | -17.26% | 196% | 43391% |
| Class A2 | -5.73% | -15.81% | -12.98% | 168% | 33962% |
| Class B | -5.93% | -15.53% | -17.44% | 148% | 35682% |
| Class C | -6.40% | -16.42% | -18.16% | 129% | 21838% |
| Class E | -8.83% | -15.25% | -16.12% | 173% | 38475% |
| **Overall** | -6.51% | -15.52% | -16.61% | 158% | 33197% |
| Class D | -6.41% | -15.30% | -18.56% | 116% | 19535% |
| Class F | -4.02% | -11.81% | -11.71% | 128% | 29605% |

The following tables show NNVC-3.0 NN-based filter set #0 (int16 precision) performance over NCS-1.0 NN-based filter set #0.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | |
|  | **BD-rate Over NCS-1.0 filter set #0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -0.12% | 0.89% | -0.28% | 129% | 110% |
| Class A2 | -0.14% | -0.54% | -0.33% | 128% | 108% |
| Class B | -0.29% | -0.50% | -0.50% | 122% | 113% |
| Class C | -0.26% | -0.87% | -0.52% | 115% | 112% |
| Class E |  |  |  |  |  |
| **Overall** | -0.22% | -0.33% | -0.43% | 122% | 111% |
| Class D | -0.18% | -0.74% | -0.89% | 119% | 102% |
| Class F | -0.08% | -0.36% | -0.21% | 119% | 107% |
|  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | |
|  | **BD-rate Over NCS-1.0 filter set #0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -0.50% | -0.80% | -0.18% | 119% | 112% |
| Class C | -0.46% | 0.16% | -0.07% | 113% | 115% |
| Class E | -0.50% | 1.70% | -0.26% | 126% | 135% |
| **Overall** | -0.49% | -0.37% | -0.13% | 118% | 118% |
| Class D | -0.44% | 0.03% | -1.36% | 114% | 107% |
| Class F | -0.05% | -0.56% | 0.17% | 126% | 111% |
|  |  |  |  |  |  |
|  | **All Intra Main10** | | | | |
|  | **BD-rate Over NCS-1.0 filter set #0** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | 0.00% | 0.00% | 0.00% | 111% | 110% |
| Class A2 | 0.00% | 0.00% | 0.00% | 124% | 113% |
| Class B | 0.00% | 0.00% | 0.00% | 113% | 112% |
| Class C | 0.00% | 0.00% | 0.00% | 106% | 102% |
| Class E | 0.00% | 0.00% | 0.00% | 114% | 106% |
| **Overall** | 0.00% | 0.00% | 0.00% | 113% | 109% |
| Class D | 0.00% | 0.00% | 0.00% | 105% | 100% |
| Class F | 0.00% | 0.00% | 0.00% | 112% | 109% |

**Filter Set 1**

The following tables show NNVC-3.0 NN-based filter set #1 (int16 precision) performance over NNVC-3.0 anchor.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | |
|  | **BD-rate Over NNVC-3.0 anchor** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -8.89% | -15.92% | -18.72% | 189% | 39242% |
| Class A2 | -10.23% | -20.25% | -17.44% | 177% | 35733% |
| Class B | -9.05% | -23.01% | -21.66% | 197% | 39779% |
| Class C | -9.92% | -21.77% | -22.30% | 166% | 36427% |
| Class E |  |  |  |  |  |
| **Overall** | -9.49% | -20.71% | -20.40% | 183% | 37928% |
| Class D | -11.48% | -23.64% | -23.95% | 167% | 32995% |
| Class F | -5.02% | -11.79% | -10.59% | 256% | 16584% |
|  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | |
|  | **BD-rate Over NNVC-3.0 anchor** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -7.83% | -15.97% | -13.81% | 177% | 38000% |
| Class C | -9.14% | -15.25% | -14.54% | 154% | 35569% |
| Class E | -8.79% | -17.14% | -17.21% | 307% | 30032% |
| **Overall** | -8.51% | -15.65% | -14.14% | 194% | 35048% |
| Class D | -10.82% | -19.44% | -18.51% | 156% | 30052% |
| Class F | -4.94% | -9.35% | -6.37% | 249% | 16689% |
|  |  |  |  |  |  |
|  | **All Intra Main10** | | | | |
|  | **BD-rate Over NNVC-3.0 anchor** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -6.59% | -17.96% | -19.97% | 173% | 28224% |
| Class A2 | -6.84% | -20.30% | -17.54% | 148% | 21252% |
| Class B | -6.82% | -21.91% | -21.71% | 142% | 22586% |
| Class C | -7.61% | -18.66% | -21.87% | 131% | 14756% |
| Class E | -10.20% | -20.84% | -20.64% | 151% | 24992% |
| **Overall** | -7.52% | -20.08% | -20.58% | 147% | 21469% |
| Class D | -7.46% | -19.01% | -21.05% | 118% | 12569% |
| Class F | -5.28% | -11.80% | -10.96% | 124% | 19230% |

The following tables show NNVC-3.0 NN-based filter set #1 (int16 precision) performance over NCS-1.0 NN-based filter set #1.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Random access Main10** | | | | |
|  | **BD-rate Over NCS-1.0 filter set #1** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | 0.03% | 0.04% | -0.25% | 111% | 105% |
| Class A2 | -0.09% | -0.05% | -0.16% | 112% | 105% |
| Class B | -0.05% | 0.06% | 0.08% | 114% | 108% |
| Class C | -0.11% | 0.01% | 0.04% | 103% | 103% |
| Class E |  |  |  |  |  |
| **Overall** | -0.06% | 0.02% | -0.04% | 110% | 105% |
| Class D | -0.05% | -0.18% | -0.09% | 100% | 105% |
| Class F | -0.35% | -0.10% | 0.09% | 105% | 95% |
|  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | |
|  | **BD-rate Over NCS-1.0 filter set #1** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -0.03% | 0.02% | 0.24% | 107% | 102% |
| Class C | -0.04% | -0.06% | 0.40% | 108% | 109% |
| Class E | -0.18% | 0.13% | -0.13% | 111% | 107% |
| **Overall** | -0.07% | -0.02% | 0.31% | 108% | 106% |
| Class D | -0.05% | -0.80% | -0.56% | 101% | 104% |
| Class F | 0.00% | -0.26% | -0.21% | 110% | 107% |
|  |  |  |  |  |  |
|  | **All Intra Main10** | | | | |
|  | **BD-rate Over NCS-1.0 filter set #1** | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -0.14% | -0.05% | -0.13% | 98% | 88% |
| Class A2 | -0.37% | 0.19% | 0.33% | 106% | 86% |
| Class B | -0.32% | -0.11% | -0.22% | 103% | 86% |
| Class C | -0.24% | 0.23% | -0.15% | 94% | 82% |
| Class E | -0.37% | 0.01% | 0.03% | 97% | 85% |
| **Overall** | -0.29% | 0.05% | -0.06% | 100% | 85% |
| Class D | -0.15% | 0.55% | 0.42% | 88% | 82% |
| Class F | -1.24% | -0.04% | -0.02% | 104% | 107% |

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

1. **Discussions**

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

* The setting of a better solution for models’ storage in order to avoid repository excess size.
* Upgrade the VTM version to benefit from bug corrections from AhG3.
* To enforce adopted contributions to merge into existing code, including for the training scripts if they exist (see issue with [MR 21/23]).
* Issue raised on the encoder invariance: likely due to an aggressive optimization compilation option affecting all float operations in the encoder.

1. **Recommendations**

The AHG recommends to:

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

It was pointed out that training scripts should be aligned/merged with the base software, and not come as separate packages.

It was suggested to develop more formal “software guidelines” for NNVC (new mandate).

[JVET-AC0015](https://jvet-experts.org/doc_end_user/current_document.php?id=12257) JVET AHG report: Optimization of encoders and receiving systems for machine analysis of coded video content (AHG15) [C. Hollmann, S. Liu, S. Wang, M. Zhou (AHG chairs)]

There are nine input contriubtions related to AHG15 mandates submitted to this meeting. They are listed below.

***Common Test Conditions***

Draft common test conditions (CTC) for AHG15: Optimization of encoders and receiving systems for machine analysis of coded video content, are summarized in document JVET-AC0073. This document includes detailed descriptions of test datasets, anchor software and configurations, anchor generation processes, machine task networks used, test and training conditions, evaluation methodologies and metrics. A reporting template in Excel format and a set of configuration files are also included in the same contribution package.

**Test datasets**

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

The SFU-HW (SFU-HW-objects-v1) dataset is a video dataset consisting of 17 sequences which are known from previous standardization efforts in JCT-VC and JVET. The sequences can be found on <ftp://hevc@mpeg.tnt.uni-hannover.de>. The annotations are available at <https://data.mendeley.com/datasets/hwm673bv4m/1>.

The Tencent Video Dataset (TVD) is a video dataset consisting of three sequences in 1920x1080 resolution used for object tracking with lengths of 3000, 636 and 2334 frames, respectively. An overview of the three sequences can be found in Table 3 of JVET-AC0073. The dataset with corresponding annotations is available at <https://multimedia.tencent.com/resources/tvd>.

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

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

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

More detailed information about these test datasets can be found in JVET-AC0073.

**Anchor software**

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

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

**Anchor configuration**

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

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

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

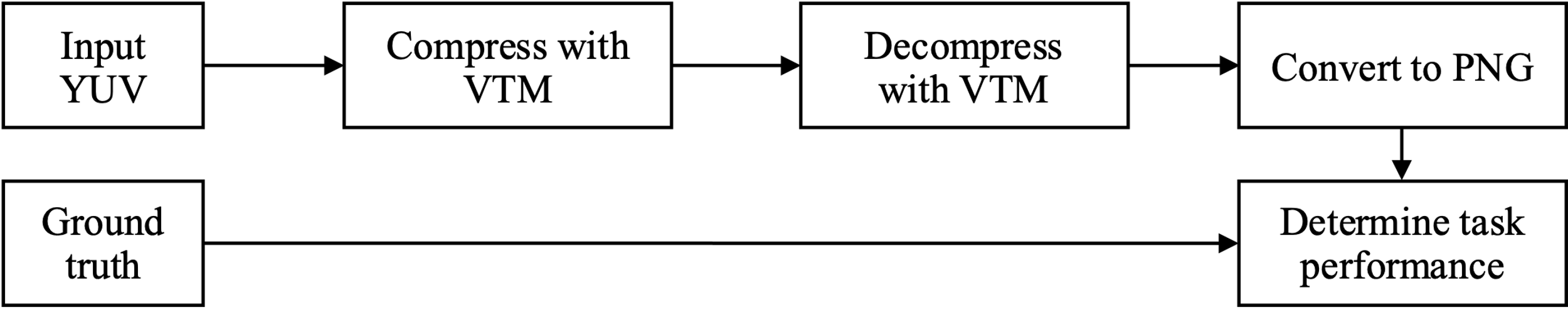


Figure 1. Anchor generation pipeline

The process to generate anchor results is described in Figure 1. Detailed description about anchor generation process is referred to JVET-AC0073. It is worthwhile to mention that for machine consumption, machine task networks are applied to compressed and reconstructed videos to obtain corresponding machine task performance. The task network Faster R-CNN X101-FPN which is part of Detectron2 (<https://github.com/facebookresearch/detectron2>) is used to evaluate object detection (SFU-HW dataset and image dataset), the model is available [here](https://dl.fbaipublicfiles.com/detectron2/COCO-Detection/faster_rcnn_X_101_32x8d_FPN_3x/139173657/model_final_68b088.pkl). The task network JDE-1088x608 which is part of the Towards Realtime MOT framework (<https://github.com/Zhongdao/Towards-Realtime-MOT>) is used to evaluate object tracking (TVD dataset), the model is available [here](https://drive.google.com/open?id=1nlnuYfGNuHWZztQHXwVZSL_FvfE551pA). More details about these machine task networks and how they are used for AHG15 anchor generation are in referred to JVET-AC0073.

**Evaluation methodology and metrics**

Proposed technologies are evaluated based on their compression performance, measured by bitrate, PSNR, mAP and MOTA, as well as encoding and decoding runtime to reflect the complexity of the proposed technology, to some extent. Definitions and detailed descriptions of these metrics can be found in JVET-AC0073. It is noted that the mean Average Precision (mAP) is used to measure object detection performance, and Multiple Object Tracking Accuracy (MOTA) is used to measure object tracking performance.

For the purpose of reporting encoding and decoding running times, the anchor and proposal should be simulated on the same platform, e.g. similar CPU and GPU configuration, to have reliable time comparison. Parallel encoding and decoding as described in JVET-B0036 may be applied for RA configurations.

In addition, relevant inference and training information should be reported if the proposed technology consists of learning-based components, such as network structure, number of network parameters, precision of network parameters, number of multiply–accumulate operations (MAC) per pixel, patch size, batch size, epoch, training time, training datasets, lost function, number of iterations, optimizer, and any pre- and post-processings used. More details can be found in JVET-AC0073.

**Reporting template**

A reporting template in Excel file format has been prepared to illustrate results following CTC and evaluation methodology described in JVET-AC0073. It is attached in JVET-AC0073 package.



Figure 2. Reporting template summary

***Anchor Results***

Anchor results were generated following the CTC described in JVET-AC0073 and crosschecked by Tencent and Ericsson.

***Technical Report Preparation***

An initial draft of technical report has been prepared and uploaded as JVET-AC0049.

***AHG Coordination***

Offline discussions were conducted about harmonization and difference on testing materials and conditions between this AHG and the VCM AHG in WG4.

**Input contributions**

There are nine input contriubtions related to AHG15 mandates. They are listed below.

|  |  |  |
| --- | --- | --- |
| **Report** | | |
| JVET-AC0015 | JVET AHG report: Optimization of encoders and receiving systems for machine analysis of coded video content (AHG15) | C. Hollmann, S. Liu, S. Wang, M. Zhou (AHG chairs) |
| **Proposal** | | |
| JVET-AC0049 | [AHG15] Draft technical report on optimizations for encoders and receiving systems for machine analysis of coded video content | [C. Hollmann (Ericsson)](mailto:christopher.hollmann@ericsson.com), [S. Liu (Tencent)](mailto:shanl@tencent.com), [S. Wang (Alibaba)](mailto:shurun.wsr@alibaba-inc.com) |
| JVET-AC0073 | AHG15: On common test conditions for optimization of encoders and receiving systems for machine analysis of coded video content | [S. Liu (Tencent)](mailto:shanl@tencent.com), [C. Hollmann (Ericsson)](mailto:christopher.hollmann@ericsson.com) |
| JVET-AC0076 | AHG9/AHG15: On the NNPFC SEI message for machine analysis | [M. M. Hannuksela](mailto:miska.hannuksela@nokia.com), F. Cricri, J. I. Ahonen, H. Zhang (Nokia) |
| JVET-AC0077 | AHG9/AHG15: On bitstreams that are potentially suboptimal for user viewing | [M. M. Hannuksela](mailto:miska.hannuksela@nokia.com), F. Cricri, H. Zhang (Nokia) |
| JVET-AC0079 | [AHG15] Effect of the perceptual QP adaptation (QPA) on machine task performance | C. Kim, D. Gwak, J. Lim (LGE) |
| JVET-AC0086 | AHG15: Feature based Encoder-only algorithms for the Video Coding for Machines | [B. Li](mailto:libinzhe.lbz@alibaba-inc.com), [S. Wang](mailto:shurun.wsr@alibaba-inc.com), [Y. Ye (Alibaba)](mailto:yan.ye@alibaba-inc.com), [S. Wang (CityU)](mailto:shiqwang@cityu.edu.hk) |
| JVET-AC0092 | AHG15: Investigations on the common test conditions of Video Coding for Machines (VCM) | [S. Wang](mailto:srwang3-c@my.cityu.edu.hk), [J. Chen](mailto:jiechen.cj@alibaba-inc.com), [Y. Ye (Alibaba)](mailto:yan.ye@alibaba-inc.com), [S. Wang (CityU)](mailto:shiqwang@cityu.edu.hk) |
| **Information** | | |
| JVET-AC0050 | [AHG15] Information about datasets used in VCM | [C. Hollmann (Ericsson)](mailto:christopher.hollmann@ericsson.com) |

**Recommendations**

The AHG recommends:

* Review all input contributions.
* Discuss and refine test conditions, evalution and reporting procedures.
* Discuss and refine anchor generation processes and results.
* Discuss existing and continue collecting new test materials.
* Discuss and establish software development and experiment environment.
* Continue investigating non-normative technologies and their suitability for machine analysis applications.
* Continue developing draft technical report on optimization of encoders and receiving systems for machine analysis of coded video content.

# Project development (25)

## Deployment and advertisement of standards (2)

Contributions in this area were discussed in session 27 at 2145–XXXX UTC on Thursday 19 Jan. 2023 (chaired by JRO).

[JVET-AC0020](https://jvet-experts.org/doc_end_user/current_document.php?id=12354) Deployment status of the HEVC standard [G. J. Sullivan (SC 29 chair)]

This information contribution contains a survey of deployed products and services using the HEVC standard and the formal specifications in which it is supported, along with a brief introduction to the standard written for broad readership. Revision marking is included to show changes relative to JVET-AB0020-v1 of October 2022.

As of September 2022, a developer survey (conducted from June to September 2022) by Bitmovin with 424 respondents from over 80 countries reported:

a. For live coding applications

i. 49% of video developers “currently using” HEVC in production

ii. 43% of video developers “planning to implement” HEVC in the next 12–24 months

b. For video-on-demand coding applications

i. 42% of video developers “currently using” HEVC in production

ii. 40% of video developers “planning to implement” HEVC in the next 12–24 months

[JVET-AC0021](https://jvet-experts.org/doc_end_user/current_document.php?id=12355) Deployment status of the VVC standard [G. J. Sullivan (SC 29 chair)]

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

Revision marking is included to show changes relative to JVET-AB0021-v2 of October 2022.

Bitmovin reported the following statistics in December 2022 from a video developer survey of June to September 2022 with 424 respondents from over 80 countries:

a. For live coding:

i. 19% of developers “currently using” VVC in production

ii. 20% of developers “planning to implement” VVC within 12–24 months

b. For video-on-demand coding:

i. 15% of developers currently using VVC

ii. 29% of developers “planning to implement” VVC within 12–24 months

Version 1.7.0 of VVenC was released in December 2022. The main changes for VVenC v1.7.0 since version 1.6.1 included faster presets (~20% for “faster”, ~5% for other presets), coding efficiency improvement by adding block importance mapping from the VTM (~1% for “faster”, ~2% for other presets), improved support for ARM processors with SIMDe, content-adaptive placement of random-access points, the possibility to add a logo overlay to encoded video, and some other features and cleanups.

The **Pentonic 1000** chipset announced in November 2022 (for availability in the first quarter of 2023) also supports VVC with similar capabilities to the Pentatonic 700.

**Philips** and **Sony** were reported in January 2023 to be expected to use the Pentonic chipsets in televisions shipping in 2023, with a formal announcement for 4K Philips OLED televisions with the Pentonic 1000 chipset expected in the following month.

**Skyworth** announced the **A63** television around the beginning of 2023, with 4K 188 Hz HDR-10 VVC playback capability.

**MainConcept** (a subsidiary of Endeavor Streaming) announced a real-time VVC encoder and SDK with real-time 8K 10-bit capability in August 2022.

**Jongbel Media Solutions** released its VVC/H.266 Video ES Viewer in 2022 for high-level syntax parsing and analysis, with limited bitstream editing features.

## Text development and errata reporting (3)

Contributions in this area were discussed in session 14 at 1600–1620 UTC on Monday 16 Jan. 2023 (chaired by JRO), in session 21 at 1320–1400 UTC on Wednesday 16 Jan. 2023 (chaired by JRO), and in session 22 at 1620–1700 UTC on Wednesday 16 Jan. 2023 (chaired by JRO).

[JVET-AC0275](https://jvet-experts.org/doc_end_user/current_document.php?id=12480) AHG9: Comments on VVC 2nd Ed. DAM1 [M. M. Hannuksela, M. Santamaria (Nokia)] [late]

This document contains comments that did not find their way into the official ballot summary of responses.

It was asserted that all detailed comments would also be covered either by similar comments of other NBs, or by general comments of other NBs. The content of JVET-AC0275 was disposed as follows:

FI\_01 – also suggested by other comments and already agreed to delay.

FI\_02 – agreed

FI\_03 – was discussed and decided in BoG JVET-AC0324.

FI\_04 – was discussed and decided in BoG JVET-AC0324.

FI\_05 – was discussed and resolved somewhat different (item 1 in JVET-AC0324)

FI\_06 – corresponding USNB comment exists

FI\_07 – agreed.

FI\_08 – first aspect (monochrome) shall be disallowed. Second aspect should be kept as in HEVC, disallow.

[JVET-AC0311](https://jvet-experts.org/doc_end_user/current_document.php?id=12516) On number of conformance tests [Hendry (LGE), Y. Sanchez (HHI), [Y-K. Wang (Bytedance)](mailto:yekui.wang@bytedance.com)] [late]

It is asserted that the equation for calculating the number of conformance tests in VVC specification has the following problems:

1. It is described that the n5 variable corresponds to conformance tests for AU based conformance and DU-based when general\_du\_hrd\_params\_present\_flag is equal to 1. In such case, then the equation cannot be correct as n5 should be used as multiplication factor, not just addition, to the number of IRAP pictures (i.e., associated with n2, n3, and n4).
2. The equation is asserted to calculate the total number of conformance tests for a given bitstreamToDecode. However, it is questioned if it is possible to do considering that when the bitstreamToDecode has more than one CVS, some parameters that apply to a CVS may not apply to other CVSs. For example, the CVS-level constraint “When bp\_alt\_cpb\_params\_present\_flag is equal to 1, the value of bp\_du\_hrd\_params\_present\_flag shall be equal to 0” affects the calculation based on the equation, however, the values of bp\_alt\_cpb\_params\_present\_flag and bp\_du\_hrd\_params\_present\_flag is CVS specific and may be different from one CVS to other CVS.

It is suggested that the JVET experts first discuss and decide the approach / direction of how the number of conformance test should be calculated. Particularly for problem #2 above, it may be difficult to calculate the number of conformance tests for the whole bitstreamToDecode and it may be better to consider modifying the specification text to provide the equation to calculate the number of conformance tests for a CVS. Modified equation and text implementation in this approach / direction is provided.

This contribution is related to specification bug ticket #1572

It was commented that AVC did not specify a necessary number of conformance tests, no equation. The equation was introduced in HEVC (but it was different, only five parameters instead of six in VVC), and introduced a necessary number of conformance tests. Problem #2 applies to both VVC and HEVC.

Options:

1. Remove equations and related text from both VVC and HEVC (but leave the statements about testing the buffer behaviour in HRD similar as in AVC)
2. Fix the equation for VVC (problem #1), which would however only apply to bitstreams containing a single CVS.
3. Fix the equation in VVC (problem #1), and change the scope from bitstream to CVS in both VVC and HEVC (problem #2)

It is noted that all these aspects are purely informative, equations are intended to help implementing conformance tests of HRD. Instead, they may rather be confusing as they are now.

Decision: Adopt option 1, both HEVC FDIS and VVC prelim. FDIS

[JVET-AC0346](https://jvet-experts.org/doc_end_user/current_document.php?id=12551) Draft 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 (editors)]

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 (items that are to be confirmed are explicitly indicated). This document also provides publication status backgrounds of these standards.

Incorporated items at the JVET-AB meeting:

* For VVC, incorporated the following items:
  + JVET-AB0223: Text improvement for Timing / DU information SEI message in HEVC and VVC
  + Added a NOTE in the HRD text to mention that "leading pictures" associated with a DRAP picture may not be correctly decodable, per the discussion of JVET-AB0055 (On leading pictures design in DRAP SEI Message).
  + Added a NOTE to mention that the definition of IRAP in VSEI covers the case of GDR with ph\_recovery\_poc\_cnt equal to 0, per the discussion of JVET-AB0057 (On the associated IRAP for DRAP and EDRAP pictures).
  + Tickets [#1564](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1564), [#1568](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1568), [#1569](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1569), and [#1572](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1572) (it was noted that the fix suggested for #1564 is editorial)
  + JVET-AB0120: Addition of interpolation for "initial CPB removal delay offset"
* For HEVC, incorporated the following items:
  + JVET-AB0223: Text improvement for Timing / DU information SEI message in HEVC and VVC
  + On conformance indication of the Multiview Main, Scalable Main, Scalable Main 10, and 3D Main profiles (resulted from a discussion between Alexis Tourapis and Gary Sullivan – thanks!)
  + On inference of high\_precision\_offsets\_enabled\_flag (resulted from a discussion between [Andy Fu](mailto:andy.fu@allegrodvt.com) and Gary Sullivan – thanks!)
  + On a redundant sentence in the definitionn of next\_bits( n ) from Gary Sullivan
  + On addition the parenthetic abbreviation TB to the text in clause 9.3.3.1 (suggested by [Cliff Reader](mailto:cliff@reader.com)– thanks!)
  + On indention of item 3 in the paragraph just below Equation 8-14 in clause 8.4.1 (reported and suggested by [Cliff Reader](mailto:cliff@reader.com) – thanks!)
  + On redudnant call to clause 8.6.1 (reported by [Cliff Reader](mailto:cliff@reader.com) – thanks!)
  + JVET-AB0120: Addition of interpolation for "initial CPB removal delay offset"
* For AVC, incorporated the following items:
  + JVET-AB0120: Addition of interpolation for "initial CPB removal delay offset"

Incorporated items at the JVET-AA 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:
  + Added changes on motion vector value ranges, based on JVET meeting minutes noted under JVET-AA0222.
  + Added a pointer of some HEVC errata items that also apply for AVC in Subsection 4.2.

Incorporated items at the JVET-Z meeting:

* For AVC, corrections to texts related to alpha blending corresponding to the changes adopted to VSEI from JVET-Z0119.
* Updated the summary and publication status for VVC.
* Updated the summary and publication status for VSEI.
* Updated the summary for HEVC.

Incorporated items at the JVET-Y meeting:

* Removed all errata items for VVC and VSEI, due to that for both VVC and VSEI the second edition texts are being prepared by both ITU-T and ISO/IEC, and all the errata items have been incorporated therein, including fixes for bug tickets [#1533](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1533) and [#1534](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1534) for VVC (fixes for these two tickets have not been included in previous versions of the errata report).
* Removed errata items for AVC that have been resolved in both the latest ITU-T text (08/2021) and the new submission (MDS21132\_WG05\_N00098) of the ISO/IEC 10th edition text.

Incorporated items at the JVET-X meeting:

* Updated the publication status of the standards
* For VVC (the changes are included in an attachment to this document):
  + Some changes resulted from reports/suggestions by Yue Yu and Frank Bossen (thanks!)
  + Various changes resulted from reports/suggestions by Peter de Rivaz (huge thanks!)
* For HEVC (the changes are included both below and in an attachment to this document):
  + Some changes resulted from reports/suggestions by Yue Yu, Frank Bossen, and Cliff Reader (thanks!)
  + Some changes to the semantics of the entry point offset syntax element, resulted from the discussion of JVET-X0050
  + Changes to the semantics of the alpha channel information SEI message resulted from the discussion of JVET-X0059
* For AVC (the changes are included in an attachment to this document):
  + Some changes to the semantics of the alpha blendnig related syntax elements in clause 7.4.2.1.2, message resulted from the discussion of JVET-X0059

Incorporated items at the JVET-W meeting:

* For VVC (the changes are included in an attachment to this document):
  + Fix for tickets [#1488](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1488), [#1511](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1511), [#1491](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1491), [#1493](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1493), [#1494](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1494)
* For HEVC:
  + Added the attachment for HEVC from JVET-W0187, which includes the following fixes:
    - Correcting an “off by one” error in the loop counter for ar\_num\_object\_updates, which indicates the number of updated annotated objects for the annotated regions SEI message.
    - Adding a description of the purpose of some variables that are derived in the semantics of the annotated regions SEI message
    - Adding a missing descriptor to specify the type of coding for the depth\_nonlinear\_representation\_model[ i ] syntax element (the same as used in another specification, so clearly what was intended) for the depth representation information SEI message
    - Adding missing range specifications for a few syntax elements of the depth representation information SEI message
    - Miscellaneous very small cleanups of consistency, grammar and clarity, and correction of typographical errors.
* For AVC:
  + Added the attachment for HEVC from JVET-W0187, which includes the same list of fixes as above for HEVC.

Incorporated items at the JVET-V meeting:

* For VVC (the changes are included in an attachment to this document):
  + Fix for ticket [#1486](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1486)
* For HEVC
  + Removed the item on the absence of “persistence flag” for the annotated regions SEI message (this was agreed to be removed at the JVET-T meeting in October 2020, but the decision was overlooked)

Incorporated items at the JVET-U meeting:

* For VVC (the changes are included in an attachment to this document):
  + Some minor editorial corrections and improvements (JVET-U0073, JVET-U0085, and one item from Hendry)
  + Fixes for tickets [#1416](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1416), [#1428](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1428), [#1432](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1432), [#1454](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1454), and [#1469](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1469)
* For AVC: Added a reference to JVET-T1006-v1 for the errata changes, and removed the actual errata text changes from this document. (JVET-U0049)

Decision: All items of JVET-AC0346 confirmed. Items remain in JVET-AC1004 until corresponding changes are done in both ITU-T and ISO/IEC versions.It is intended to include all clarified items into the HEVC FDIS and VVC preliminary FDIS.

[JVET-AC0352](https://jvet-experts.org/doc_end_user/current_document.php?id=12557) Comments on ITP-PQ-C2 [W. Husak (Dolby)] [late]

The ITP color representation was introduced during the 28th JVET meeting in document JVET-AB00172. This document updates the matrix coefficients to reflect the proper values for RGB→LMS conversions. The conversions are different from the existing coefficients to reflect the proper amount of hue crosstalk to improve linearity.

Presented in session 23 at 2215 UTC (chaired by JRO)

This provides the necessary information to make the ITP description precise and self-contained.

It was pointed out that another sequence of numbers (having ITP ahead of YCoCg-R) might be preferable in the table, to have a logical grouping of color types.

Decision: Adopt JVET-AC0352 for a new version of JVET-AC1008. Also put a note that a re-numbering of code points might be desirable.

## Test conditions (4)

Contributions in this area were discussed in session 22 at 1700–1730 UTC on Wednesday 18 Jan. 2023 (chaired by JRO), and in session 23 at 2100–2145 UTC on Wednesday 18 Jan. 2023 (chaired by JRO).

[JVET-AC0041](https://jvet-experts.org/doc_end_user/current_document.php?id=12226) [AHG4] On visual volumetric video-based coding (V3C) testing conditions [S. Schwarz, M. M. Hannuksela, P. Rondao Alface, L. Kondrad, L. Ilola (Nokia)]

This contribution proposes to study coding tools for video content utilised in visual volumetric video-based coding (V3C). It includes an overview of V3C content characteristics, potential test content, and testing procedures.

It was reported that the purpose would be development of “light weight” coding tools.

It was commented that inclusion of new tools would in any case require definition of new profiles.

Why not just use the CTC used in WG 7 and report evidence about potential benefit of tools that are related to PCC, and also report the performance on conventional content. Could also be ECM

[JVET-AC0044](https://jvet-experts.org/doc_end_user/current_document.php?id=12229) [AHG4] Occupancy-only PSNR calculations for V3C V-PCC coding evaluation [S. Schwarz, M. M. Hannuksela (Nokia)]

No need for presentation. Related to AC0041.

[JVET-AC0096](https://jvet-experts.org/doc_end_user/current_document.php?id=12298) AHG10/12: Suggestion for new CTC for RPR in VTM and ECM [K. Andersson, R. Yu (Ericsson)]

Reference Picture Resampling (RPR) is a tool that allows referencing pictures encoded at higher or lower resolution which can be beneficial for adapting to transmission capabilities or for achieving better coding efficiency when coding at lower bitrates. It is noted that the current RPR CTC for VTM and ECM only applies to low-delay B configuration (LDB) and it is asserted that the current setup can lead to uncomfortable viewing experiences. In this contribution we suggest changing the RPR CTC for better testing of RPR-related coding tools and for better viewing experiences. The suggested changes contain the following four elements. Firstly, introduce more variation of the encoded resolutions in the same sequence. This enables better exercising of variants of RPR filtering cases. Secondly, reduce the QP when encoding at a lower resolution to maintain similar bitrate as when encoding at full resolution. This enables better testing of the full range of QPs when encoding in different resolutions. The third change is to always measure PSNR values of the decoded output sequences in the full resolution since the output will anyway be viewed in the full resolution. The fourth change is to also include random access (RA) configuration since RPR can also be useful for RA which has been demonstrated by GOP based RPR.

v2: complete results for ECM and partial results for ECM with fixes from JVET-AC0095 also including patches for VTM and ECM respectively.

v3: more results from ECM with fixes from JVET-AC0095.

What is the reasoning behind keeping rate constant when switching resolution? Is this a realistic scenario in RPR applications?

What would be necessary to add to the software to test under these conditions?

Is the purpose of the CTC to demonstrate benefit? No, the main intent is to exercise various switching scenarios for functionality testing. It was commented that this is usually not the main purpose of CTC.

It was pointed out that a software able to execute these CTC would be beneficial in developing additional conformance testing for RPR.

It was commented that currently not many tool investigations around RPR are performed.

It was suggested that functionality testing should not be mixed up with the usual purpose of CTC, i.e. the benefit of tools.

Next step from this contribution would be a software update. Further study on that.

Beyond that, while the current CTC on RPR are still valid, it should be emphasized that these are not very realistic for a real-life RPR application. On the other hand, the current proposal may also not be more realistic. Further study on this.

Decision(SW): Adopt the software from JVET-AC0096 for both VTM and ECM. Also include description in an update JVET-AC2002

[JVET-AC0329](https://jvet-experts.org/doc_end_user/current_document.php?id=12534) Crosscheck of JVET-AC0096 (AHG10/12: Suggestion for new CTC for RPR in VTM and ECM) [J. Nam (LGE)] [late]

[JVET-AC0137](https://jvet-experts.org/doc_end_user/current_document.php?id=12339) AHG3: On Affine AMVR setting for VTM CTC [X. Li (Google)]

In the current VTM CTC, affine AMVR is enabled for RA but is disabled for LDB and LDP. It is reported that affine AMVR actually shows good tradeoff for LDP, i.e., 0.41% luma gain with 105% encoding time. It is proposed to turn on affine AMVR for LDP in VTM CTC.

Would this also have effect for ECM? Likely not.

One expert commented that 0.4% luma gain vs. 5% encoder runtime might not be a good tradoff.

Why higher gain for class D? Has not been analysed.

Decrease of decoding time? No, numbers are not reliable.

It was commented that increasing the performance of LP (when it does not come almost for free) is not of primary importance. Better target would be reduction of runtime without loss.

LDP is by purpose a mode with less encoder complexity than LDB. Therefore, changes with not as good tradeoff should not be made.

No action.

[JVET-AC0256](https://jvet-experts.org/doc_end_user/current_document.php?id=12461) Cross-check of JVET-AC0137: On Affine AMVR setting for VTM CTC [A. Henkel (HHI)] [late]

## Subjective quality testing and verification testing (5)

Contributions in this area were discussed in Joint meeting at 1620–1745 UTC on Monday 16 Jan. 2023 (chaired by JRO and MW).

[JVET-AC0042](https://jvet-experts.org/doc_end_user/current_document.php?id=12227) AHG4: Report on AHG meeting on verification test preparation for VVC multilayer coding [M. Wien (RWTH)]

No need for presentation, was covered in AHG4 report

[JVET-AC0043](https://jvet-experts.org/doc_end_user/current_document.php?id=12228) AHG4: Report on AHG meeting on subjective test preparation for film grain synthesis [M. Wien (RWTH)]

No need for presentation, was covered in AHG4 report

[JVET-AC0145](https://jvet-experts.org/doc_end_user/current_document.php?id=12347) AHG4: experiments in preparation of dual-layer VVC visual tests [P. de Lagrange (InterDigital)]

This document discusses the quality ladder category of the VVC multi-layer visual tests planned in JVET-AB2021. It proposes a recommended quality ladder for VVC, and reports about experiments of dual-layer coding various pairs in the quality ladder. Using one example sequence, an example combination of single-layer and dual-layer coding for the full ladder is said to save about 28% storage space, at the expense of some impact on end user for some rate points (around 7 to 11% bitrate increase and 7-10% quality degradation).

This is in the category of spatial and quality scalability of the test plan.

A “quality ladder” is proposed which consists of a set of two-layer arrangements (each base+enhancement) to get rate/quality adaptation ov a wide range. This saves storage relative to only using single layer points in that ladder, but the quality point of the enhancement layers has higher rate than if it was by itself coded as single layer (base layer points are identical to their corresponding single layer points). De facto, the values of “bit rate increase” and “quality degradation” add somewhat up to the increase of enhancement layer rate compared to corresponding single-layer points.

Next steps: Taking elements from this document, define test sequences, select useful rate points, update of test plan.

It was asked which filters were used for downscaling, and if luma and chroma positions are properly aligned (for which it would not be appropriate using same filters for luma and chroma.

[JVET-AC0181](https://jvet-experts.org/doc_end_user/current_document.php?id=12385) AHG4: experiments in preparation of film grain visual tests [P. de Lagrange (InterDigital)]

This document reports about some experiments in preparation of visual tests planned in JVET-AB2022 for FGC SEI visual tests on top of VVC, more specifically the first category which goal is to assess the visual improvement brought by film grain synthesis over a wide range of coding quality. Three sequences that originally contain grain have been encoded with the VTM at every QP from 12 to 51, and checked visually to find when grain is just preserved, when grain is fully removed by encoding alone, and the QP range useful for visual assessment.

The goal was to find:

* Where grain just starts to be impaired by encoding (or is “just preserved”), with rating close to 10
* Where grain is removed, but the underlying content hopefully very well preserved (rating close to 9), the encoder acting as a denoiser
* The QP range useful for visual quality assessment (rating typically from 2 to 9)

From the joint meeting on Monday:

It was commented that content from the SVT sequences may not be very representative, and that for other kind of material just using “VTM as denoiser” may not be appropriate.

Having clean material and adding different levels of noise might be a viable approach (provided that an appropriate artistic synthesis was available).

It was further suggested to divide the test into two classes of “light grain” and “heavy grain”. It was commented that the latter could either be artificially generated or coming from old film or generated by purpose with heavy-grain film materials (or even be natural grain strengthened by artificial grain).

A follow-up joint meeting was held on Wednesday 1530-1620:

* Clarify what the goal of a first round of subjective test would be: Content as currently available in our test sets, removal for purpose of simpler coding, synthesis after decoding, and comparison against original (this would be the main focus of the current draft), could be done with naïve subjects.
* Acquire wider variety of material, and extend the test in the bullet point above. If this includes more heavy film grain, this might require further development on noise removal and parameter derivation in the software (still with naïve subjects).

For the second bullet, it might be considered to use cropped material if tha can easier be acquired to circumvent copyright problems. Update the draft test plan from the contribution JVET-AC0181, and also include initial plans on the second bullet.

It had been commented before that superimposing synthetically generated film grain (done in a studio) over a clean sequence (which could be some of our exising test sequences) might give some advantage, as it would allow using the clean sequence as ground truth, e.g.

* Testing how good a film grain removal algorithm is, as this allows mimicking “perfect removal”
* Testing if film grain synthesis superimposed upon the coded sequence can hide coding artifacts
* Testing how close film grain synthesis generated via SEI works to reproduce similar grain as the one orginally superimposed.

How to implement this is for further discussion.

It was mentioned that in this context, the problem exists that reproducing “real” film grain is sophisticated. On the other hand, most film grain present today may be synthetic.

[JVET-AC0199](https://jvet-experts.org/doc_end_user/current_document.php?id=12403) AHG13: On Film Grain Synthesis Subjective Evaluation [X. Meng, W. Zhang, S. Labrozzi (Disney Streaming)] [late]

Initial version rejected as “placeholder”.

See notes under 6.3

## Test material (0)

This section is kept as a template for future use.

## Quality assessment methodology (1)

Contributions in this area were discussed in plenary session at 1300–XXXX UTC on Friday 20 Jan. 2023 (chaired by JRO).

[JVET-AC0267](https://jvet-experts.org/doc_end_user/current_document.php?id=12472) Training Methods in Visual Assessment: Potential Improvements for Expert Viewing Tests [M. Wien (RWTH), V. Baroncini (VABTech)] [late]

This document reports results of a subjective assessment which was carried out in addition to the tests reported in JVET-AB2029 at the Mainz meeting site with a small group of experts and a modified training procedure. The goal of this additional subjective test is to investigate the effectiveness of the modified training procedure and the feasibility of a smaller group of viewers for the expert viewing scenario.

The scatter plots seem to indicate a consistent behavior of the MOS scores resulting from the two tests. This applies to the resulting MOS values as well as to the corresponding confidence intervals. These results suggest that a comparable confidence may have been reached with half the number of expert viewers. Considering the effort of conducting expert viewing tests, this result is taken as an encouraging indication to further study and potentially refine the modified training method.

The MOS-rate plots resulting from the reported tests reveal a benefit of the VVenC bitstreams compared to the VTM bitstreams for about half of the test sequence. For the other half, the results do not indicate a clear trend towards the one or the other encoder. It is noted that the plots for the Campfire, CatRobot1, and DaylightRoad2 sequences show saturation effects for the highest rate point each. These effects were not observed in the results reported in JVET-AB2029.

Contribution for information.

The training consisted in demonstrating examples and explaining to the experts which range of MOS would be expected to be allocated with those examples, in order to better “calibrate” the votes of the individuals.

Useful for future expert viewing tests.

## Conformance test development (0)

This section is kept as a template for future use.

## Software development (3)

Contributions in this area were discussed in session 26 at 1530–1630 UTC on Thursday 19 Jan. 2023 (chaired by Y. Ye).

[JVET-AC0057](https://jvet-experts.org/doc_end_user/current_document.php?id=12256) AhG14 SADL v4 changes [F. Galpin, T. Dumas, P. Bordes, E. François (InterDigital)] [late]

This contribution presents updates in the Small AdHoc Deep Learning (SADL) library. The update contains both some performance improvement and new features.

New features include support for sparse matrix-vector multiplication for network using dense layers, slicing layer, grouped convolutions and separable convolutions.

New features in SADL version 4:

* Sparse matrix-vector multiplication, no need to change model, useful for intra prediction
* Separable convolutions: useful for model complexity reduction
* Grouped convolution: useful for model complexity reduction. This changes the model format, and it is recommended for existing models to be updated to this new model format.
* Unit testing: useful for regression testing on models, and layers within a model
* New layer for slicing tensor: this was not listed in the contribution but included in the SADL update

It was reported that it would be desirable to automate the regression testing of merge requests. This is currently not done for NNVC. It was commented by software coordinator that some additional testing may be accommodated, but if excessive testing is performed, there could be computing resource issues on the git server. This is delegated to the AHG3 and AHG14 software coordinators to find a reasonable solution to meet the request.

The proposed actions are discussed as follows:

* It was proposed to update the SADL version in NNVC to the version 4 for the next software version. Currently SADL version 3 is used, and it was commented that it would be beneficial to use the latest SADL version in NNVC. The AHG14 software coordinator commented that there are no known issues with this proposed SADL version update (other than the need to update models to new format as noted above)
* It was proposed to make the SADL repository public to be aligned with the NNVC repository. Currently, SADL is not public whereas NNVC is public. This means that to clone SADL repo, MPEG credentials would be needed, whereas cloning the NNVC repo does not require credentials. Due to the fact that NNVC has dependency on SADL, this discrepancy creates some access difficulty for participants. It was commented that there did not seem to be specific reasons for keeping SADL private. It was commented that making resource publicly available seems to be a good practice in general. It was also commented that other than SADL, all other software repos on the git server are public (NNVC was made public in the last meeting). This was agreed. K. Suehring will take the action to make SADL public.

Decision(SW): use SADL version 4 in the next NNVC software release.

[JVET-AC0166](https://jvet-experts.org/doc_end_user/current_document.php?id=12370) AHG3: Guidelines for VVC reference software development [F. Bossen, X. Li, K. Sühring (software coordinators)]

An update to JVET-N1003 “Guidelines for VVC reference software development” is proposed to reflect common code development practices and recent development tools versions.

It was commented that the last version of this guideline document was updated more than four years ago. The proposed guideline updates were reviewed, including the following (not an exhaustive list):

* Copyright changed from 2019 to 2023
* C++ 11 upgraded to C++ 17, with the latter providing more features beneficial for software development
* Updating to more up-to-date compiler versions
* Screenshot examples are updated
* Other general editorial improvements (clarification etc.)

The guidelines on using the formatting tool to automatically format the source code are not mandatory. This was previously discussed and considered best to keep as optional. It was commented that given we have very active work ongoing, making this mandatory could create burden.

It was commented that the guidelines are very good and should be applied to other software development, and in fact, ECM software development follows these guidelines. It was commented that the C++ portion of the NNVC software should also follow these guidelines. This was agreed.

It was suggested to revise the title of the document to include these exploration software packages. The new title should be “Guidelines for VTM-based software development”.

These suggested changes are agreed to be included in the revised output document JVET-AC2003.

[JVET-AC0204](https://jvet-experts.org/doc_end_user/current_document.php?id=12408) AHG3: Guidelines for HEVC reference software development [K. Sühring, F. Bossen, X. Li (software coordinators)]

An update to JCTVC-H1001 “HEVC software guidelines” is proposed to reflect common code development practices and recent development tools versions.

It was commented that the last version of this guideline document was provided in 2012 as a JCTVC document. Changes similar to the JVET-AC0166 are included in this document as software development guidelines.

Following the discussion regarding JVET-AC0166, it was commented that these guidelines apply to other HM-based software packages, such as SHM (SHVC reference software), SCM (SCC extension reference software), HTM (Multiview-HDVC and 3D-HEVC reference software). It was therefore suggested to update the title to “Guidelines for HM-based software development”. It was suggested to add X. Li as editor of this document, and remove D. Flynn.

## These suggestions were discussed and agreed to be included in the revised output document JVET-AC1001.Implementation studies and complexity analysis (1)

Contributions in this area were discussed in session 26 at 1630–1650 UTC on Thursday 19 Jan. 2023 (chaired by Y. Ye).

[JVET-AC0266](https://jvet-experts.org/doc_end_user/current_document.php?id=12471) 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)] [late]

This document provides updated information on features, coding efficiency and runtime for version 1.7.0 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. Furthermore, 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 v1.7.0 since version 1.6.1 include:

* Faster presets: ~20% for faster, ~5% for other presets
* Coding efficiency improvement by adding block importance mapping from VTM:
  + ~1% BD-rate gain for faster
  + ~2% BD-rate gain for other presets
* New features and improvements:
  + Support for ARM through SIMD everywhere (SIMDe)
  + Adaptive intra period (content adaptive placement of random-access points)
  + Noise-based QP clipping in rate control
  + Improved Screen Content Coding detector to ignore black borders
  + Using MCTF to collect statistics for QPA (even if MCTF is disabled)
  + Possibility to add logo overlay to encoded video
* Various bugfixes and cleanups

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 | -11.2% | 290x | 2000x | -15.2% | 330x | 2100x | -13.4% | 310x | 2000x |
| FAST | -24.5% | 120x | 840x | -26.2% | 140x | 890x | -25.5% | 130x | 870x |
| MEDIUM | -33.7% | 33x | 220x | -36.1% | 43x | 270x | -35.0% | 38x | 240x |
| SLOW | -37.4% | 9.7x | 65x | -39.5% | 13x | 84x | -38.5% | 12x | 75x |
| SLOWER | -40.6% | 2.4x | 16x | -42.8% | 3.4x | 21x | -41.8% | 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 | -21.9% | 280x | 1900x | -20.8% | 320x | 1700x | -18.6% | 300x | 1800x |
| FAST | -32.7% | 120x | 830x | -31.2% | 140x | 740x | -30.1% | 130x | 790x |
| MEDIUM | -38.8% | 32x | 250x | -39.7% | 41x | 200x | -37.8% | 37x | 220x |
| SLOW | -42.0% | 9.6x | 81x | -43.0% | 13x | 58x | -41.1% | 11x | 69x |
| SLOWER | -44.3% | 2.4x | 20x | -46.0% | 3.4x | 14x | -43.9% | 2.9x | 17x |

No new version of VVdeC has been released since v1.6.0.

It was commented that better speed can be achieved using native SIMD rather than SIMD everywhere. It was also commented that assembly for entropy coding seemed to provide speed benefit for other open source software such as x264. These could be considered in the future development plan of VVEnc/VVDec.

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

Contributions in this area were discussed in session 26 at 1650–1725 UTC on Thursday 19 Jan. 2023 (chaired by Y. Ye).

[JVET-AC0180](https://jvet-experts.org/doc_end_user/current_document.php?id=12384) AHG 7: Reference frame padding for GDR [T. Poirier, F. Aumont (InterDigital)]

This contribution proposes to pad the dirty area using samples from the clean area instead of setting unavailable samples to 2*BD-1* for reference pictures used by CUs in the clean (refreshed) area.

Simulations were conducted with ECM-7.0 under low-delay B configuration specified by AHG 7. The results show a BD-rate gain of 0.2% overall, 0.2% in class C and 0.4% in class E.

Cross checker confirmed the simulation results.

Normative GDR-related aspects in ECM differ from normative GDR-related aspects in VVC as follows:

* Handling of dirty and clean regions
* Asymmetric deblocking across the dirty/clean boundaries
* Virtual boundary processing

It was commented that the GDR-related deblocking filter adoption from last meeting (JVET-AB0171) has not yet been integrated into ECM7. And it was asked if that aspect would have interaction with the current proposal. The proponent thinks there could be some interactions between the two.

Integration of JVET-AB0171 has encountered the following issue and been delayed:

* when virtual boundary is used (this is outside of CTC), 135-degree SAO would cause the dirty area and clean area to be mixed, the patch to fix the issue has been submitted and merged recently (after ECM7 has been released)

Since then, patch for JVET-AB0171 has been submitted, it is pending review and not yet integrated into ECM. It was confirmed that JVET-AB0171 will be integrated soon and should be available as part of ECM8.

It was commented that the testing of this proposed low-level change would not affect the CTC, and would need to use the GDR-specific conditions defined by AHG7.

It was commented that during AHG report reviews, it was decided to stop AHG7 and put the AHG7 activities requiring normative changes into AHG12.

It was commented that GDR-specific testing conditions have not been documented, and need to be clearly specified in the EE2 descriptions.

Investigate in EE, using GDR-specific testing conditions.

[JVET-AC0338](https://jvet-experts.org/doc_end_user/current_document.php?id=12543) Cross check of JVET-AC0180 (AHG 7: Reference frame padding for GDR) [S. Hong (Nokia)] [late]

## AHG10: Encoding algorithm optimization (3)

Contributions in this area were discussed in session 26 at 1725–1745 UTC on Thursday 10 Jan. 2023 (chaired by Y. Ye), and in session 27 at 2100–2145 UTC on Thursday 10 Jan. 2023 (chaired by JRO).

[JVET-AC0138](https://jvet-experts.org/doc_end_user/current_document.php?id=12340) Adjusting luma/chroma BD-rate balance in ECM [E. François, Y. Chen, C. Salmon-Legagneur (InterDigital)]

This contribution reports experiments based on ECM-7.0 aiming at obtaining a better repartition of BD-rate variations between luma and chroma, when using VTM-11.0ecm6 as reference. As a result, the contribution tests adjustments of the configuration parameters “CbQpOffset”, “CrQpOffset” and “LMCSOffset” for the configurations of the ECM CTCs for SDR and HDR contents.

Several encoder configurations are tested for SDR contents, leading to the following AI/RA/LB BD-rate PSNR results over VTM-11.0ecm6:

* ECM-7.0: AI: -9.1%/ -18.0% / -19.0%; RA: -19.2%/ -25.4%/ -26.7%; LB: -16.2%/ -26.9%/ -26.3%
* config. 1: AI: -10.1%/ -10.2% / -11.7%; RA: -19.4%/ -22.9%/ -24.3%; LB: -16.4%/ -22.3%/ -22.1%
* config. 2: AI: n/a ; RA: -20.4%/ -15.4%/ -17.5%; LB: -16.7%/ -16.1%/ -14.7%

One encoder configuration is tested for HDR contents, leading to the following AI/RA results over VTM-11.0ecm6 (provided numbers are BD-rate wPSNR for H1, BD-rate PSNR for H2):

* ECM-7.0: AI/H1: -7.37% / -25.17% / -30.93%; AI/H2: -7.23% / -17.05% / -20.56%;
* config: AI/H1: -8.65% / -8.52% / -15.25%; AI/H2: -7.93% / -7.75% / -11.82%;
* ECM-7.0: RA/H1: -14.35% / -33.89% / -39.81%; RA/H2: -13.31%/ -26.59%/ -30.97%\*
* config: RA/H1: -15.39% / -18.47% / -24.97%; RA/H2: -13.89%/ -14.79%/ -20.05%\*

\* H2 (HLG) RA results are not yet fully available – one missing rate point (SunsetBeach, QP 22) was replaced by H2 RA results previously generated, but with wrong config files (per\_class classA\_randomaccess.cfg omitted)

It was commented that a similar contribution (JVET-Y0223) had been brought previously (which included an additional knot over “QPOutValCb” to control the luma/chroma balance), and concerns were expressed at the time regarding the potential impact to subjective quality. It was mentioned by the proponents that subjective quality had been checked on a few sequences, and no obvious artefacts were found.

It was discussed that the current imbalance could be due to the adoption of more chroma tools (e.g. cross component prediction tools) in ECM.

It was commented that in some configurations/content, there seemed to be a rather large drop in chroma gains and in return a relatively small increase in luma gains.

It was commented that this contribution was intended as for information only, in order to show how more-balanced gains between luma and chroma in the current ECM could look like, and to show if 20+% luma gain is to be achieved, how much chroma gains could still remain.

Further study encouraged, including other methods to achieve better balance between luma and chroma (such as using QPCMappingTable).

[JVET-AC0236](https://jvet-experts.org/doc_end_user/current_document.php?id=12440) Cross-check of JVET-AC0138 [F. Le Léannec (Xiaomi)] [late]

[JVET-AC0139](https://jvet-experts.org/doc_end_user/current_document.php?id=12341) AHG10: Encoder Optimization of VTM Merge Functions [X. Li (Google)]

In this contribution, an encoder only optimization on merge is proposed for VTM software. 97%, 99% and 92% encoding time is reported with 0.02%, 0.01% and 0.03% luma BD-rate loss for RA, LDB and LDP configurations, respectively. By replacing three merge functions on regular merge, sub-block merge and geometric merge with one unified merge function, redundant code is reportedly removed.

Results confirmed, and confirmation that the code changes are appropriate, given by cross-checker.

Could this optimization also be applied to ECM? In principle yes, but there are more merge candidates, such that it is not a straightforward code copy.

Decision(SW): Adopt JVET-AC0139 (also CTC with the suggested settings)

[JVET-AC0319](https://jvet-experts.org/doc_end_user/current_document.php?id=12524) Crosscheck of JVET-AC0139 (AHG10: Encoder Optimization of VTM Merge Functions) [F. Bossen (CA)] [late]

[JVET-AC0149](https://jvet-experts.org/doc_end_user/current_document.php?id=12351) AHG10/12: Reduced I-frame QP for RA [K. Andersson, P. Wennersten (Ericsson)]

This proposal shows results for a config file change reducing the I-frame QP by one in the RA configuration. It is reported that BD-rate is improved by -0.28%/-2.62%/-2.76% for Y/Cb/Cr in VTM. Similar numbers are reported for ECM. Subjective benefits are reported for static content, no noticeable subjective problems are reported. It is proposed that the change be adopted into the CTC for both VVC and ECM.

v2: complete results for ECM and correction for class F and aligned class F results for VTM with excel and accurate timing for class B, C and D for VTM.

v3: results for HDR for VTM and ECM added. Thanks to Interdigital for help with HDR for ECM. Also added MSSSIM comparison for SDR.

It was commented that the results are varying per sequence, there are several sequences with losses over various classes. The best setting is likely to be sequence dependent, but in CTC identical settings should be used.

Encoding time variations reported in the contribution are not reliable. Should not change the encoding time. By tendency, the bit rate should be higher (sometimes >5%)

The proponent reports that detail is increased, and no pumping is visible.

It was commented that CTC should not be changed too frequently.

Could it be done the other way round, i.e. keep the QP for I pictures fixed and increase for the other pictures? There was some difficulty, loss was observed in particular at QP22. It was asserted that this is caused by the QP toggling feature which was introduced for low QP range.

Further study.

[JVET-AC0255](https://jvet-experts.org/doc_end_user/current_document.php?id=12460) Cross-check of JVET-AC0149: AHG10/12: Reduced I-frame QP for RA [A. Henkel (HHI)] [late]

## Profile/tier/level specification (0)

Contributions in this area were discussed in session 25 at 1440–1450 UTC on Thursday 19 Jan. 2023 (chaired by JRO).

[JVET-AC0339](https://jvet-experts.org/doc_end_user/current_document.php?id=12544) New higher level for multilayer coding [S. Iwamura, S. Nemoto, A. Ichigaya (NHK)] [late]

This contribution proposes addition of a higher level for multilayer coding. In VVC version 2, maximum level 6.3 is defined for 4K/8K multilayer coding as well as single layer coding of 12K 60Hz. However, max luma sample rate (MaxLumaSr) of level 6.3 is not sufficient for three-layer coding use case such as 2K/4K/8K multilayer coding with a frame rate of 60Hz. Therefore, new higher level 6.4 for typical three-layer coding is proposed as follows:

Table A.2 – General tier and level limits

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Level** | **general\_level\_idc value\*** | **Max luma picture size MaxLumaPs (samples)** | **Max CPB size MaxCPB (CpbVclFactor or CpbNalFactor bits)** | | **Max slices per AU MaxSlicesPerAu** | **Max # of tiles MaxTilesPerAu** | **Max # of tile columns MaxTileCols** |
| **Main tier** | **High tier** |
| **6.3** | 105 | 80 216 064 | 240 000 | 1 600 000 | 1 000 | 990 | 30 |
| **6.4** | 108 | 80 216 064 | 240 000 | 1 600 000 | 1 000 | 990 | 30 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Level** | **Max luma sample rate MaxLumaSr (samples/sec)** | **Max bit rate MaxBR (BrVclFactor or BrNalFactor bits/s)** | | **Min compression ratio MinCrBase** | |
| **Main tier** | **High tier** | **Main tier** | **High tier** |
| **6.3** | 4 812 963 840 | 320 000 | 1 600 000 | 8 | 4 |
| **6.4** | 9 625 927 680 | 320 000 | 1 600 000 | 8 | 4 |

It was asked if not also the maximum bit rate should be increased?

For future consideration. As currently an “unconstrained” level is under development which could e.g. be used for demonstrations of future products, definition of higher levels becomes relevant when those products are foreseeable to be coming to the market. It would be interesting to get information what timelines are expected for this.

## Proposed modification of system interface (0)

This section is kept as a template for future use.

## AHG15: Optimization of encoders and receiving systems for machine analysis of coded video content (8)

Contributions in this area were discussed in session 6 at 1530–1750 UTC on Thursday 12 Jan. 2023 (chaired by JRO) and in session 13 at 1325–1440 UTC on Monday 16 Jan. 2023 (chaired by JRO).

[JVET-AC0351](https://jvet-experts.org/doc_end_user/current_document.php?id=12556) BoG report on non-normative optimization for machine [C. Hollmann, S. Liu (BoG coordinators)]

This document contains the report of the BoG on non-normative optimization for machine. The BoG met on Wednesday 18 January 2023 at 1520-1720 UTC, discussing subjects related to non-normative optimization for machine consumption of coded video content, including:

* Common test conditions
  + Test sequences
  + RA/LD/AI configurations
  + QP points and BD-rate calculation
* Cross-check procedure
* Reference software
  + Where to setup the repository
  + What to be included in the initial version reference software
* Other aspects

The BoG recommends:

* Align common test conditions with WG 4 VCM (including test sequences, test configurations, QP points, etc.)
* Further discuss on QP point selection for calculation of the BD-rate used for tool evaluation.
* Align crosscheck procedure with WG 4 VCM.
* Further discuss on crosscheck procedure for neural network-based encoding as well as processing tools.
* Setup repository for software development and test.
* Integrate technologies proposed in JVET-AB0275 and JVET-AC0086 to the reference software (VTM12.0) and share them in git.

2.1 Common test conditions

JVET-AC0073 AHG15: On common test conditions for optimization of encoders and receiving systems for machine analysis of coded video content [S. Liu (Tencent), C. Hollmann (Ericsson)]

This document describes common test conditions (CTC), reference software and configurations, reporting and other information to be used for experimenting and evaluating technologies for encoders and receiving systems for machine consumption of coded video content.

2.1.1 Test sequences

Discussion:

WG4 VCM AHG decided to remove class E (SFU-HW dataset) and Kimono sequence from CTC, and truncated (or split) long TVD sequences into 10sec clips.

Question was asked whether experiments can be conducted using original (un-shortened) TVD sequences. – yes, but performance might be different to truncated ones.

Related input contribution m62326.

Recommendation: align with WG4.

2.1.2 RA/LD/AI configurations

Recommendation: align with WG4.

- Include RA and LD (LDB) as mandatory configurations and AI as optional.

- Do not use temporal subsampling for AI configuration.

2.1.3 QP points and BD-rate calculation

Recommendation: align with WG4. Further discuss about the key BD-rate used for tool evaluation.

- QP points are adjusted to avoid saturation problem.

- BD-rate of low-4 and high-4 QPs are calculated in addition to all (6) QP BD-rate calculation.

It was commented that maybe adding more QPs would help.

It was then commented that 6 QPs are sufficient.

It was commented that multiple BD-rate measurements based on the same R-D data may not be beneficial, instead may bring confusion. Will bring this question to WG4 – the key question is which BD-rate to look at when comparing performance and making decisions.

2.2 Cross-check procedure

Recommendation: align with WG4. Further discuss crosscheck criteria for neural network-based encoding algorithms.

It was previously discussed and decided to require identical bitstreams for crosscheck and allow small variance for mAP and MOTA (1% relative to the performance score).

It was commented that if a neural network-based encoder algorithm (e.g. for motion estimation) is used, then the bitstreams may not be identical. This issue will be further investigated and bring to WG4 for discussion.

Question was asked about crosscheck on training. We will start from crosschecking inference for now, while training information such as dataset and process should be described in the contribution. This aspect will be further discussed and bring to WG4 as well.

2.3 Reference software

2.3.1 Where to setup the repository

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

jvet-ahg-mccvc (?) name to be decided in JVET.

2.3.2 What to be included in the initial version reference software

JVET-AC0049 [AHG15] Draft technical report on optimizations for encoders and receiving systems for machine analysis of coded video content [C. Hollmann (Ericsson), S. Liu (Tencent), S. Wang (Alibaba)]

This contribution contains a draft for the technical report on optimizations for encoders and receiving systems for machine analysis of coded video content that JVET decided on working on during its 28th meeting.

Some items could be included in the initial reference software:

- Spatial and temporal QP configuration /adjustment on VTM (JVET-AB275)

- Pre-processing (RoI based analysis and processing) technology (JVET-AC0086)

Recommendation: Anchor remains VTM12.0. Above items will be implemented on top of VTM12.0 and shared in git for investigation.

2.4 Other aspects

It was asked whether other VTM versions could be used for proposals. It is recommended to use VTM12.0 for both anchor and proposals for now. It is encouraged to investigate the benefit of using of newer VTM versions and bring input contributions to the next meeting.

Was presented Thu. 19 Jan. 1300

The first version of the software codebase will still be based on VTM12 (anchor of CfP), but it was emphasized that it is highly desirable to convert to newest VTM. It was argued that the performance of newer VTM has improved and it would be difficult to compare with results of WG 4 where VTM 12 is still used. However, most of the improvements are likely due to CTC modfications (which were more aligned with those of ECM), and the investigation on VCM could use their own CTC which is more aligned with the one historically used in VTM12. This should be investigated in the AHG.

It is planned to produce output documents for draft 1 of TR (preliminary WD of WG 5), and CTC.

[JVET-AC0049](https://jvet-experts.org/doc_end_user/current_document.php?id=12234) [AHG15] Draft technical report on optimizations for encoders and receiving systems for machine analysis of coded video content [C. Hollmann (Ericsson), S. Liu (Tencent), S. Wang (Alibaba)]

This contribution contains a draft for the technical report on optimizations for encoders and receiving systems for machine analysis of coded video content that JVET decided on working on during its 28th meeting.

It was suggested to include the concept of keyframe identification and analysis. Visual features conained in keyframes give important clues on the content of a video for machine vision tasks and should not be destroyed by coding. This could also be combined with dynamically changing the frame rate based on amount of temporal changes.

It is suggested to also mention conventional (e.g. VVC interpolation) filters for spatial upsampling rather than bicubic.

It was commented that the current version of the technical report is very high-level, and needs to be filled by more concrete technology implementation for the given purposes, e.g in terms of pre and post processing, and related encoder control.

It would also be important to discuss the video content properties which need to be retained for the machine vision tasks.

It might be good to identify a selection of machine vision tasks and describe how a video needs to be encoded in order to not destroy the performance relative to the encoded video.

The purpose should be to generate a bitstream without knowing which machine vision task would be applied after decoding. Encoding for exactly one task could hypothetically be much more optimized, but might not be a realistic application scenario where compression would be needed.

Which types of metadata are existing in our standards that could be useful for machine vision task? For example, there are many ways of describing camera position, movement, etc.

Our standards also contain ways to code depth maps, for example

A comment was made about possible low-latency requirements.

The report requires a clear focus, and that should come from the application perspective – this focus should be the same for the normative activities of WG 4 and non-normative of the TR. JVET experts have experience about how to use and how to optimize encoders for low-level tools.

Further discuss in joint meeting with WG 2 and WG 4.

[JVET-AC0050](https://jvet-experts.org/doc_end_user/current_document.php?id=12235) [AHG15] Information about datasets used in VCM [C. Hollmann (Ericsson)]

(include abstract)

For information – no need for presentation

[JVET-AC0073](https://jvet-experts.org/doc_end_user/current_document.php?id=12275) AHG15: On common test conditions for optimization of encoders and receiving systems for machine analysis of coded video content [S. Liu (Tencent), C. Hollmann (Ericsson)]

This document describes common test conditions (CTC), reference software and configurations, reporting and other information to be used for experimenting and evaluating technologies for encoders and receiving systems for machine consumption of coded video content.

It was pointed out that rate matching is not useful for measuring BD rates. It was used in the CfP, whre also BDmAP and BD-MOTA were criteria.

The proposal is to use classes A-E with similar conditions as in JVET-CTC (were also used in CfP), as well as longer sequences (Tencent). It is also using LB configuration which had not been used in CfP. LB could be a problem with longer sequence.

It was asked if usage of class E is useful?

Tasks proposed to be investigated are object detection and object tracking (bounding box and object ID), In the CfP (and in WG 2), another task is image segmentation.

For RA and LB, both object detection and object tracking are investigated.

In AI, only object detection is investigated, and additional results are obtained from annotated image databases which were also in the CfP. This is relevant to assess performance with regard to analysing keyframes from a video.

CTC should be kept aligned with those that are used with WG 4. This was further discussed in the joint meeting on Wednesday 18 Jan., and the following agreements were made:

* Usage of class E – to be removed
* Kimono – to be removed
* Temporal subsampling AI – not good for object tracking
* Usage of low-delay configuration – LDB to be used
* Length of the Tencent sequences – will be investigated in shorter chunks
* Omit rate points which are too high, such that performance is saturated
* Relevance of instance segmentation task using only image database – to be dropped from CTC, due to lack of relevant data
* Diversity of content in general, e.g. object tracking is only performed by tracking persons – improvement of the selection of machine vision tasks should be considered, also that the results are improved for the uncoded video (which might also lead to saturation only at higher rate points).
* What are useful ranges of scores (also at low end)? This requires further consideration, and also relates to the question how useful BD criteria based on MaP are here.

[JVET-AC0076](https://jvet-experts.org/doc_end_user/current_document.php?id=12278) AHG9/AHG15: On the NNPFC SEI message for machine analysis [M. M. Hannuksela, F. Cricri, J. I. Ahonen, H. Zhang (Nokia)]

Was reviewed in BoG JVET-AC0324.

[JVET-AC0077](https://jvet-experts.org/doc_end_user/current_document.php?id=12279) AHG9/AHG15: On bitstreams that are potentially suboptimal for user viewing [M. M. Hannuksela, F. Cricri, H. Zhang (Nokia)]

The contribution reviews some earlier contributions on preprocessing and encoding methods targeted at machine analysis and concludes that spatial and temporal quality fluctuations resulting from such methods may result into decoded video that looks subjectively unpleasant.

The contribution proposes a suboptimal user viewing indication SEI message, which indicates that the video resulting by decoding the current CLVS is not optimized for user viewing and may therefore look incoherent. The proposed SEI message can be used to avoid accidental displaying of machine-targeted bitstreams.

It was pointed out that at systems layer in file format specification there is an indicator that a video is not suitable for viewing (entitled as “restricted video”)

In the discussion, specifically in the context of VCM, it was suggested to clarify with WG 4 how metadata for machine consumption would be conveyed, which will likely be needed also in the case when die video stream itself is a legacy stream.

The contributors explain that the usage of the proposed SEI message is not specific to the VCM application. The message also contains indicators which convey if only some temporal sublayers are suboptimal. It was asked if in such concept it might also be beneficial to indicate if some regions are or are not useful. It was noted that the annotated regions SEI might fulfill such purpose (see also discussion in BoG JVET-AC0324).

It was further noted that concept of VUI indication exists to indicate that a bitstream has some special meaning. It might be more consistent to align the intended purpose with that.

Further study recommended.

[JVET-AC0079](https://jvet-experts.org/doc_end_user/current_document.php?id=12281) [AHG15] Effect of the perceptual QP adaptation (QPA) on machine task performance [C. Kim, D. Gwak, J. Lim (LGE)]

This contribution reports the impact of the perceptual QP adaptation (QPA) on machine task performance. VTM 12.0 with QPA enabled configuration is used for encoding open image dataset that was used in the evaluation of the VCM CfP and object detection was performed using the decoded results. Experimental results reportedly show 15.17% BD rate gain for object detection.

Mainly effective for high QPs. It was reported by another participant that similar experiments were performed with other data sets, where in some cases gains, but sometimes also losses occurred.

It was suggested to study how this would combine with other approaches of VCM-related encoder optimization. In the current contribution, the VTM anchor was improved.

Further study recommended.

It was suggested to have a BoG (C. Hollmann, S. Liu) after the joint meeting with WG 4 which should further discuss testing condition and a plan for software development.

[JVET-AC0086](https://jvet-experts.org/doc_end_user/current_document.php?id=12288) AHG15: Feature based Encoder-only algorithms for the Video Coding for Machines [B. Li, S. Wang, Y. Ye (Alibaba), S. Wang (CityU HK)]

On the basis of the joint adaptive spatial temporal (JAST) system, this document describes a feature-based encoder-only algorithm. The algorithm makes use of the deep learning information of the visual task network to guide the encoder-only pre-processing. These results are reported in this contribution: The proposed algorithm could achieve up to -63.49% and -63.38% BD-rate savings for object detection task on TVD and FLIR dataset respectively. The proposed encoder-only algorithm also achieves obvious performance improvement for other machine tasks and test datasets.

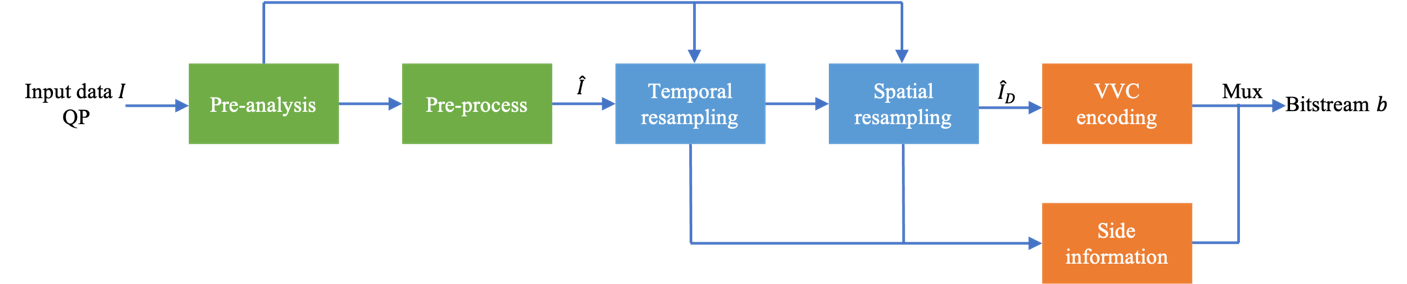


Fig. 1 The high-level block diagram of the JAST encoding process

It was asked why the performance is much better for classes A and B. It is asserted by the proponents that the effect is more significant for higher detail which is more apparent in higher resolutions.

Candidate for software (further discussion in BoG)

It was suggested that a cross-check procedure should be established for inclusion in the software package.

[JVET-AC0092](https://jvet-experts.org/doc_end_user/current_document.php?id=12294) AHG15: Investigations on the common test conditions of Video Coding for Machines (VCM) [S. Wang, J. Chen, Y. Ye (Alibaba), S. Wang (CityU HK)]

This contribution illustrates the non-monotonic issue for machine task performance in the common test condition (CTC) of video coding for machines (VCM), which affects the ability for us to compare the performance of different proposals in an apple-to-apple manner. Specifically, the non-monotonic rate distortion behavior of various machine vision tasks are observed and illustrated in both some of the proposed technologies and the anchor. In this contribution, we investigate the CTC of VCM and propose three possible solutions to solve/alleviate this issue.

The proposal is to better focus on BD quality rather than BD rate, but several experts suggested that BD rate should also be considered.

Non-monotonic behaviour happens at low rates (high QP) where perhaps the encoding simplifies the video such that it is for the benefit of the machine vision task – avoiding high QP could potentially resolve this problem

The quality-over-rate graphs have the tendency of a “threshold-like” behaviour (i.e. breakdown below a certain rate point, constant performance above) – how useful are BD metrics in that case?

How useful is it to include rate points where the machine vision performance becomes unacceptable, e.g. mAP in the low range of 10-20?

Is a part of the problem, that in some cases the performance of the machine vision task is not good even for the uncoded video? It was reported that some of the algorithms were not really trained for the type of video that are used.

Further discuss in joint meeting with WG 4 how the evaluation methodology could be improved.

## General aspects of standards development and applications of standards (3)

Contributions in this area were discussed in session 25 at 1415–1440 UTC on Thursday 19 Jan. 2023 (chaired by JRO), unless noted otherwise.

[JVET-AC0316](https://jvet-experts.org/doc_end_user/current_document.php?id=12521) ISO 22028-5 impact on CICP [N. Bonnier, D. Concion, D. Podborski, J. Roland, A. Tourapis (Apple)] [late]

ISO TC42 is currently developing specification ISO 22028-5, which defines a set of colour image encodings for use in storage, transmission, and display of HDR and WCG digital still images. The purpose of this contribution is to inform JVET group on the impact of this work to the CICP specification.

This is still under development (at WD?), with very different transfer characteristics than currently supported by CICP. This will likely be a TS rather than. Contribution for information. It is to be determined if this might have future impact on new definitions in CICP.

[JVET-AC0317](https://jvet-experts.org/doc_end_user/current_document.php?id=12522) Registration authority for CICP video colour representation code points [D. Podborski, [A. Tourapis](mailto:atourapis@apple.com) (Apple)] [late]

This contribution discusses the idea of handling CICP video colour representation code points through a registration authority, similar to FourCC registration with MP4RA.

It is reported that in WG 3, a similar mechanism is used for various standards (e.g. FF, OMAF), but it is still approved by WG 3 on a case-by-case basis if a code point is considered.

It was commented that this would require some legal negotiation between ISO and the corresponding RA, which might not be easy to achieve. Further, same would be necessary with ITU, as CICP is a joint standard.

It was further commented that with such an approach CICP would not be self contained, e.g. not including the complete equations. Further, the range of code points is limited (8 bit).

No action at this moment. For further consideration.

[JVET-AC0334](https://jvet-experts.org/doc_end_user/current_document.php?id=12539) Summary of the Evidences on Lenslet Video Coding [X. Jin (THU), M. Teratani (ULB)] [late]

This contribution provides evidence of designing coding tools dedicated to lenslet videos. Lenslet video is a general format for dense light fields that can be captured by the plenoptic cameras, e.g. Lytro and Ratrix cameras, or rearranged from dense two-dimensional multiview data. It shows high potential in both storage-oriented and communication-oriented applications. We demonstrate that designing specific coding tools dedicated to lenslet videos can improve compression efficiency obviously compared to the existing standards. We hope the summarization of lenslet video coding evidence is helpful to support the creation of a video codec extension.

Was presented in joint meeting with WG 4 on Wednesday 18 Jan. 1500-1525 UTC.

Currently 6% improvement are found for lenslet content by a combination of low-level tools in VVC (in Intra and RA conditions). These would need to be disabled for normal video content.

It was commented that this could be more relevant for a future standard than extending VVC by block-level tools in a new profile. A pure HLS in VVC might be OK, but it appears that not much gain could be achieved by that.

It was commented that 6% gain would probably not justify putting new tools and defining a profile giving benefit only for a very specific application case. This would also contradict the idea of versatility.

Adding a test condition in JVET was requested, but it was commented that test conditions are already existing in WG 4, and in the context of an exploration, JVET software can be used and extended for a given exploratory purpose. It was suggested that also ECM could be used in the WG 4 exploration to investigate if it is better suitable for compressing lenslet content.

No specific action in JVET, further study in WG 4 recommended.

# Low-level tool technology proposals

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

This section is kept as a template for future use. The possibility of removing this section and discontinuing AHG8 was considered, but AHG8 was reconstituted for another meeting cycle.

## AHG11: Neural network-based video coding (29)

### Summary, BoG reports, and information documents

Contributions in this area were discussed in session 3 at 2100–2315 UTC and in session 4 at 2335–2355 UTC on Wednesday 11 Jan. 2023 (chaired by JRO).

[JVET-AC0023](https://jvet-experts.org/doc_end_user/current_document.php?id=12258) EE1: Summary report of exploration experiments on neural network-based video coding [E. Alshina, F. Galpin, Y. Li, M. Santamaria, J. Ström, H. Wang, L. Wang, Z. Xie (EE coordinators)]

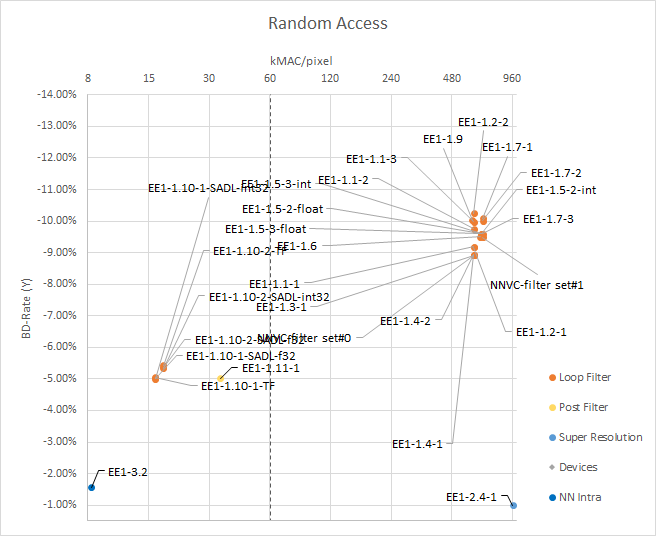
This document summarizes Exploration Experiment 1 (EE1) tests performed between the JVET-AB and JVET-AC meetings to evaluate **Neural Network-based Video Coding (**NNVC) technologies, analyze their performance and complexity aspects. EE1 conducted tests in three categories: NN-based in-loop and post-filters, adaptive resolution coding with “classical” and NN-based filters, NN-based Intra coding. Average across JVET test set compression gain over VVC anchor achieved in those categories for random access configuration is 10.2% (filters), 2.1% (adaptive resolution coding with NN-re-sampler), 1.6% (NN-based Intra).

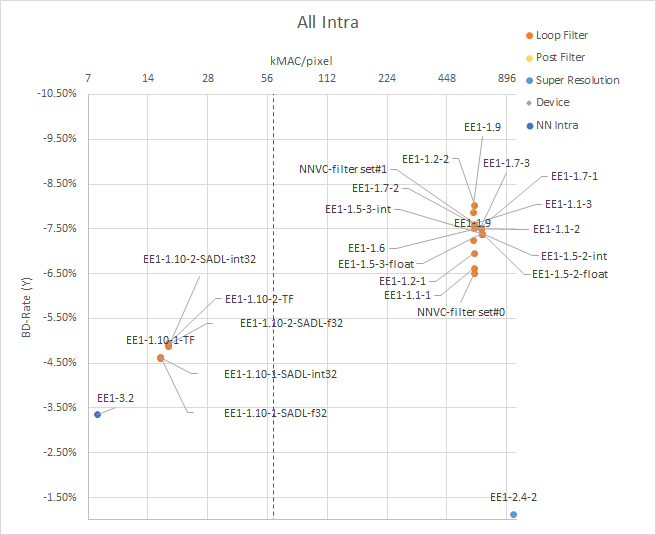
NNVC common SW base (***NNVC***) and training scripts are highly encouraged to be used by all EE1 proponents. All proponents were asked to report results relatively to the AhG11 anchor NNVC-3.0 [1] and use up-to-date version of results reporting template recommended by AhG11. It was the case for all EE1 test except test on end-to-end AI picture coding (which was tested using very different test conditions). Some proponents still don’t report MS-SSIM metric, even support for it was added long time ago.

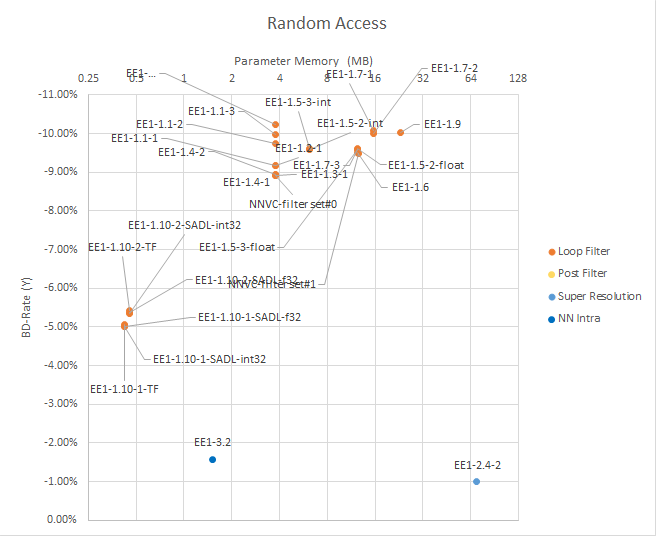
During this round of Exploration Experiment tests were massively cross-checked, which indicates that more and more JVET members become familiar with testing and training NN-based video coding technologies. Massive cross-check was performed for both model training and for inference (coding).

In testing the performance difference up to 0.1% BD-rate is reported. It must be asserted that with NN-based filters enabled in NNVC SW same level of performance deviation is observed (even for quantized to int16 models). It was explained by NNVC SW coordinators that issue is known and likely called by usage of different compilers. Hopefully this issue will be solved in AhG14.

The most promising technologies recommended by JVET **undergo procedure of cross-check for the training**. If NN model was re-trained by cross-checker and then in test performance difference between model from proponent and cross-checker is even larger (up to 0.5 %).





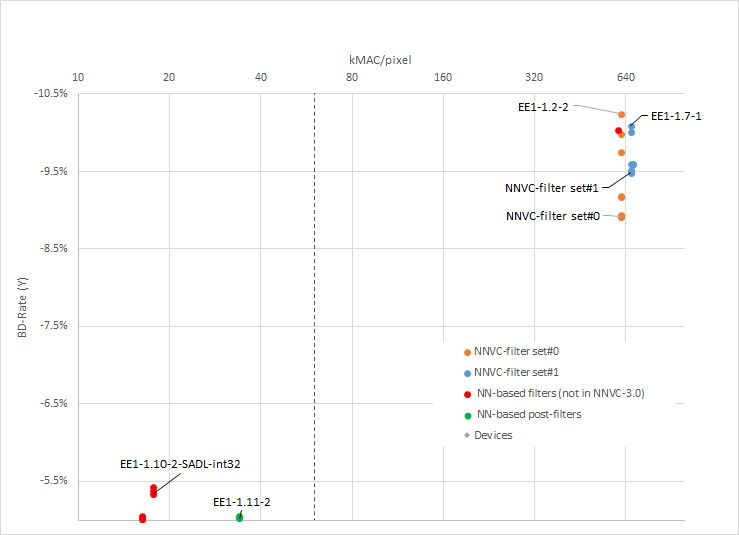


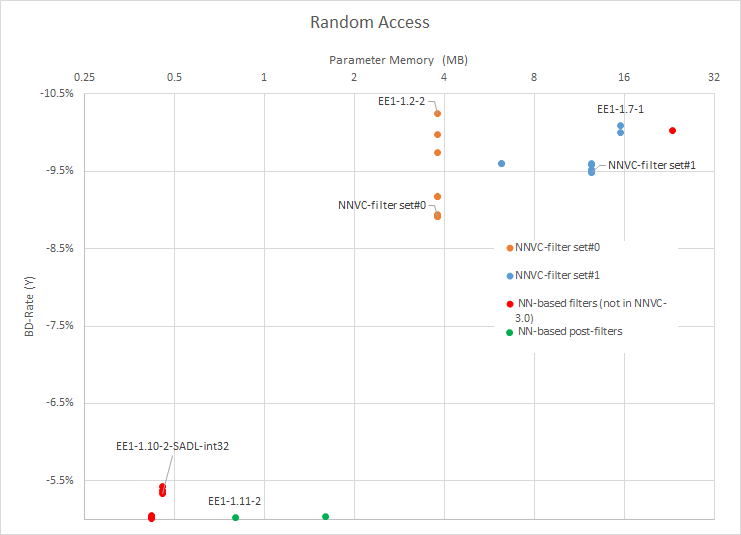
*Figure 1 Complexity performance analysis of all EE1 test: Random Access cfg BD-rate vs kMAC/pxl (top), All intra cfg. BD-rate vs kMAC/pxl (middle), Random Access cfgBD-rate vs kMAC/pxl (bottom).*

Complexity for all EE1 test has been analyzed and summarized according to AhG 11 recommendations. BD-rate vs computational complexity (kMAC/pxl) and total memory for model parameters is shown in Figure 1.

1. **Exploration experiments on Enhancement filters**

There are two technologies in this category, which are part of NNVC SW, known as ***filter set #0*** and ***filter set#1***. Those technologies are clustering around level of complexity 600~700 kMAC/pxl and demonstrate BD-rate gain vs VVC anchor 9~10%. ***Filter set#0*** filters family got slightly higher performance improvement during this EE1 round (compare to ***filter set#1*** family) w/o increment kMAC/pxl and memory.

**



*Figure 2 Typical representatives of exploration experiment tests on NN-based enhancement filters: BD-rate gain over VVC in random access configuration vs kMAC/px (top); vs total memory size for all models (bottom).*

***NNVC filter set#0 based***

*Table 1 EE1 tests for filter set#0 architecture*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| EE1- Test | document | Random access, AhG11 CTC | | | All Intra, AhG11 CTC | | |
| Y | Cb | Cr | Y | Cb | Cr |
| **NNVC-filter set#0** | [**JVET-AC0014**](https://jvet-experts.org/doc_end_user/current_document.php?id=12255) | **-8.9%** | **-18.5%** | **-19.2%** | **-6.5%** | **-15.5%** | **-16.6%** |
| EE1-1.1-1 | [JVET-AC0194](https://jvet-experts.org/doc_end_user/current_document.php?id=12398) | -9.2% | -19.6% | -19.2% | -6.6% | -15.3% | -16.4% |
| EE1-1.1-2 | [JVET-AC0194](https://jvet-experts.org/doc_end_user/current_document.php?id=12398) | -9.8% | -21.1% | -20.6% | -7.5% | -16.8% | -17.4% |
| EE1-1.1-3 | [JVET-AC0194](https://jvet-experts.org/doc_end_user/current_document.php?id=12398) | -10.0% | -21.4% | -20.9% | -7.6% | -16.7% | -17.5% |
| EE1-1.2-1 | [JVET-AC0195](https://jvet-experts.org/doc_end_user/current_document.php?id=12399) | -9.2% | -18.5% | -19.2% | -7.0% | -15.6% | -16.7% |
| EE1-1.2-2 | [JVET-AC0195](https://jvet-experts.org/doc_end_user/current_document.php?id=12399) | -10.2% | -21.4% | -20.9% | -8.0% | -16.8% | -17.6% |
| EE1-1.3-1 | [JVET-AC0063](https://jvet-experts.org/doc_end_user/current_document.php?id=12265) | -8.9% | -19.7% | -19.7% |  |  |  |
| EE1-1.4-1 | [JVET-AC0064](https://jvet-experts.org/doc_end_user/current_document.php?id=12266) | -8.9% | -19.4% | -19.4% |  |  |  |
| EE1-1.4-2 | [JVET-AC0064](https://jvet-experts.org/doc_end_user/current_document.php?id=12266) | -8.9% | -20.2% | -19.9% |  |  |  |

In all tests NN parameters precision is Int 16, total number of parameters is 1.9 Million, computational complexity 615 kMAC/pxl (assuming block wise processing).

**EE1-1.1** is about filter set#0 model refinement to address change in NNVC CTC (enabling EncDbOpt). Performance of filter set#0 improved by 1.2% (all configuration) in test EE1-1.1-3 called real-iterative training. This training hint looks useful and deserves to be presented and explained to the group in details.

**EE1-1.2** consideration of filter application is RDO process. In RDO simplified model is used. Similar encoder only technology is already part of NNVC for ***filter set #1***. Performance improvement from encoder only change is 0.3% (RA cfg). The combination with “normative” modifications from EE1-1.1-3 improves performance of ***filter set#0*** by 1.5%. The best in ***filter set#0*** family test EE1-1.2-2 achieves 10.2% gain over VVC anchor; computational complexity 615 kMAC/pxl, only one model with 1.9 M parameters.

In test **EE1-1.3** Chroma components are swapped prior to start NN-based filter, so filter trained for “Cr” component is used for “Cb” (and vice versa). Experiment was done for both filter architectures in NNVC. BD-rate gain is mostly visible for Chroma component (1...2%).

In test **EE1-1.4** residual adjustment by reducing residual magnitude by value {1, 2} (content adaptively). Performance improvement looks not that impressive for random access configuration, but in low-delay B configuration the gain of this tool is 0.9% (on top of ***NNVC filter set#0***). This fact indicates the lack of NNVC filters adaptively or not sufficient enough training of them for low-delay use case.

***NNVC filter set#1 based***

*Table 2 1 EE1 tests for filter set#1 architecture*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| EE1-test | document | Total params, M | kMAC/pxl | Random access, AhG11 CTC | | | All Intra, AhG11 CTC | | |
| Y | Cb | Cr | Y | Cb | Cr |
| **filter set#1** | [**JVET-AC0014**](https://jvet-experts.org/doc_end_user/current_document.php?id=12255) | **6.2** | **664** | **-9.5%** | **-20.7%** | **-20.4%** | **-7.5%** | **-20.1%** | **-20.6%** |
| 1.5-2 float | [JVET-AC0089](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0089-v1.zip) | 3.1 | 677 | -9.6% | -21.3% | -22.0% | -7.4% | -18.2% | -20.0% |
| 1.5-2 int16 | [JVET-AC0089](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0089-v1.zip) | 3.1 | 677 | -9.6% | -21.0% | -22.1% | -7.4% | -18.3% | -20.1% |
| 1.5-3 float | [JVET-AC0089](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0089-v1.zip) | 3.1 | 673 | -9.6% | -21.0% | -21.5% | -7.4% | -18.2% | -19.9% |
| 1.5-3 int16 | [JVET-AC0089](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0089-v1.zip) | 3.1 | 673 | -9.6% | -20.9% | -21.5% | -7.4% | -18.3% | -19.9% |
| 1.6 int16 | [JVET-AC0143](https://jvet-experts.org/doc_end_user/current_document.php?id=12345) | 6.2 | 664 | -9.5% | -20.7% | -20.4% | -7.5% | -20.0% | -20.8% |
| 1.7-1 int16 | [JVET-AC0177](https://jvet-experts.org/doc_end_user/current_document.php?id=12381) | 7.8 | 664 | -10.1% | -20.7% | -20.4% | -7.5% | -20.1% | -20.6% |
| 1.7-2 int16 | [JVET-AC0177](https://jvet-experts.org/doc_end_user/current_document.php?id=12381) | 7.8 | 664 | -10.0% | -20.7% | -20.4% | -7.5% | -20.1% | -20.6% |
| 1.7-3 int16 | [JVET-AC0177](https://jvet-experts.org/doc_end_user/current_document.php?id=12381) | 6.2 | 664 | -9.6% | -20.7% | -20.3% | -7.5% | -20.1% | -20.6% |
| 1.8 int16 | [JVET-AC0118](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0118-v1.zip) | 6.2 | 664 |  |  |  | -7.5% | -20.6% | -21.3% |

In test **EE1-1.5** same model is used for intra and inter (both Luma and Chroma), NNVC filter set#1 architecture slightly modified by removing attention modules. So, total size of memory needed for NNVC filter set #1 is reduced by half, computational complexity (kMAC/pxl) about the same; compression performance of ***NNVC filter set#1*** preserved (or even slightly better). Architecture with and w/o IBP information as input to Chroma NNVC filter network was tried, performance difference is very small. Cross-check for training is still running. Inference results have been cross-checked. Modification is recommended to be considered for adoption to NNVC (need to complete training cross-check).

Test **EE1-1.6** shows that partitioning information can be safely (w/o performance drop) removed from Intra Chroma models in ***NNVC filter set#1***. In the inter chroma model, there is already no partitioning input. The training and inference were cross-checked. After re-training, the BD-rates between proponent and cross-checker deviate up to 0.05% in luma and 0.20% in chroma.

Test **EE1-1.7** studies two aspects: 1) adding reference frame as extra input to ***NNVC filter set#1***; 2) flipping input to ***NNVC filter set#1.*** Aspect#1 provides gain of 0.6% (but extra model is used and so total memory size for filter set#1 increases (6.2 🡪 7.8 million parameters). Aspect#2 provides 0.1% (RA) /0.2% (LDB) gain, doesn’t increase number of multiplications or memory for filter parameters. The best in ***filter set#1*** family test EE1-1.7.1 achieves 10.1% gain over VVC anchor; computational complexity 664 kMAC/pxl, five models with 7.8 M parameters (instead of four in the original filter set#1).

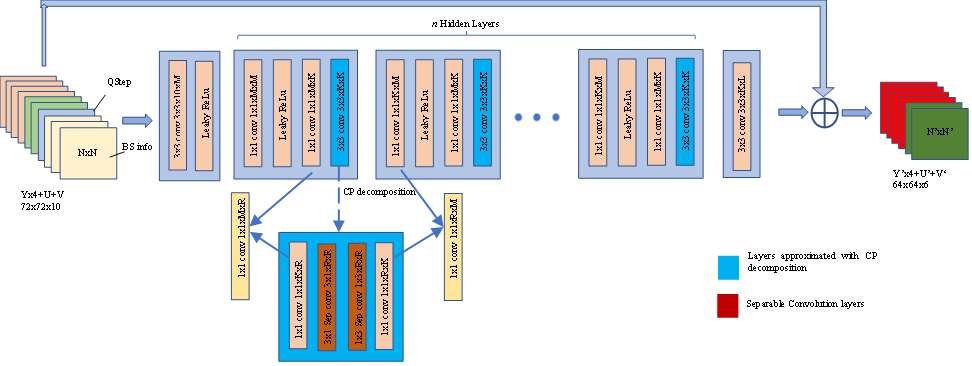
In test **EE1-1.8** loss function in training was modified: weighted sum of L1 and L2 was used, weight depends on QP. Some minor Chroma gain is observed. (no results for RA)

***Not in NNVC NN-filters***

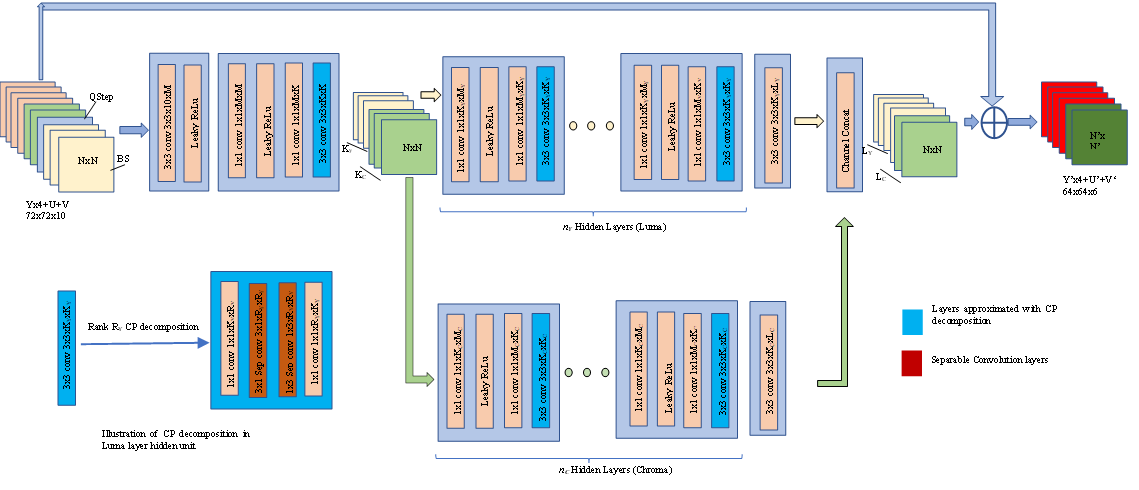
Experiment **EE1-1.9** is trying to combine network architecture design elements of filter ***filter set #0*** and ***filter set#1***. Gain and complexity are at the same level with filters already included into NNVC.

Filters tested in experiment **EE1-1.10** operate at ~30 times smaller computational complexity and ~20 times smaller model size compared to filters in NNVC SW, demonstrating promising 5% BD-rate gain over VVC anchor.

This filter is not yet cross-checked. Key elements of two variants of low complexity NN-based filter design in EE1-1.10 are decomposition if 2D convolution (3×3) to the sequence of two 1D convolution (3×1 and 1×3) separable convolutions with same number of channels (R=K) coupled with linear fusion of adjacent 1x1 convolutions and split of Luma and Chroma layers (Chroma layers uses less channels). Details of simple NN-based filter in EE1-10 are shown of figures below. Note that 3x1 and 1x3 convolutions are both spatial and depth-wise separable.



*Figure 3 CP Decomposition + fusion of 1x1 conv layers of JVET-X0140 Baseline Model*



*Figure 4 Split luma and chroma model + CP Decomposition + fusion of 1x1 conv layer*

**Test 1.10.1** Test the model with CP decomposition and fusing adjacent 1x1 convolution with following model parameters:

* Number of hidden layers: n = 11
* Feature maps and rank: K = 24, R = 24, M = 72
* Model complexity = 16.2 kMAC/pixel

**Test** **1.10.2** Test the model with CP decomposition, fusing adjacent 1x1 convolution and split architecture for luma and chroma components with following model parameters:

* Number of hidden layers for luma and chroma: nY = nC = 10
* Feature maps and rank for luma and chroma split (24L, 8C):
  + Luma: KY = 24, RY = 24, MY = 72
  + Chroma: KC = 8, RC = 8, MC = 24
* Model complexity: 17.7 kMAC/pixel

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

Post-filter in experiment **EE1-1.11** was recommended for the cross-check of training. Retraining was done by proponent. Performance deviation from original proposal after re-training is up 0.5%, this issue is under investigation jointly by proponent and cross-checker.

It was verbally reported that a part of the problem coul be the way of extracting training data (e.g., 8- to 10-bit conversion for some of the data).

Recommendations from the summary report:

* Proponents should ensure smooth integration of adopted NN-based tools to NNVC SW (including well described training scripts)

1. Adopt improvements in EE1-1.2 for NNVC filter set #0 verified though this round of EE1 (a total of 1.5% gain by a combination of RDO optimization which gives approx. 0.3% and iterative training which gives the remaining gain)– After review of contribution JVET-AC0195 it was decided that the method EE1-1.2-2 is not sufficiently verified. The best fully verified approach that uses the new modified training procedures is EE1-1.1-2. Decision: Adopt JVET-AC0194 method 1.1-2. It was noted that further improvement could be expected by invoking RDO, but it is not known if the RDO decision method from EE1-1.2-1 that was optimized for the original filter #0 would work with the new model.

* Adopt improvements EE1-1.5 and EE1-1.7 for NNVC filter set #1 verified though this round of EE1; EE-1.5 reduces number of models from 4 to 2 without performance degradation, but has not been fully cross-checked; EE-1.7 gives gain, but also increases the number of models from 4 to 5. It might not be a good tradeoff standalone, but might be combined with EE-1.5, but there is no proposal for such a combination. Later such a proposal was made and adopted, see notes under JVET-AC0310.
* Establish training cross-check for of EE1-1.10 (simple NN-based filter) in the next round of EE – this could be attractive as another point of lower-complexity operation in the NNVC software.
* Adopt EE1-1.4 for the next version of NNVC SW as generic technique improving performance of NN-based filters (it does not give relevant gain for RA and AI, but gives gain for LB, but this could be due to the fact that no LB data were used in training) – further investigate in EE, also on top of EE1-1.1.2, also on top of filter set #1
* It was recommended to adopt EE1-1.11 for the next version of NNVC SW as an example of post-filter (was decided after confirmation of cross-check)

As a general remark, it would be desirable to have only one filter architecture in the range of “High complexity/high gain” in NNVC. Currently, the combination idea for both filter sets (EE1-1.9) is not competitive with the best-performing versions of the standalone sets from the EE. It should be further studied in the next EE if some of the aspects such as RDO optimization and iterative training would be beneficial for this method as well. Contribution JVET-AC0155 was reviewed later.

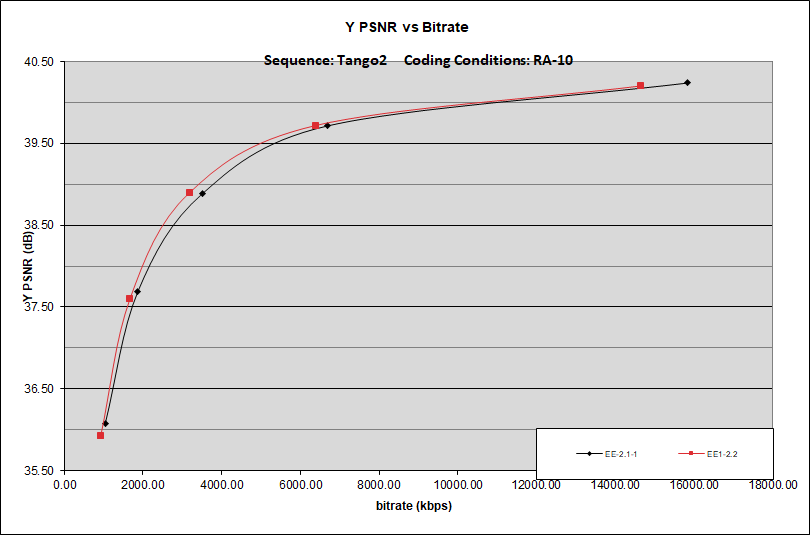
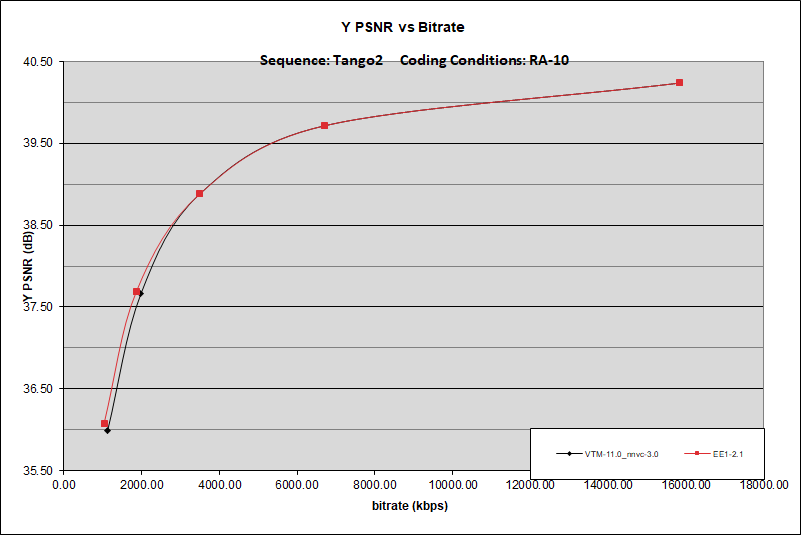
**Adaptive resolution coding and super-resolution NN-based post-filters**

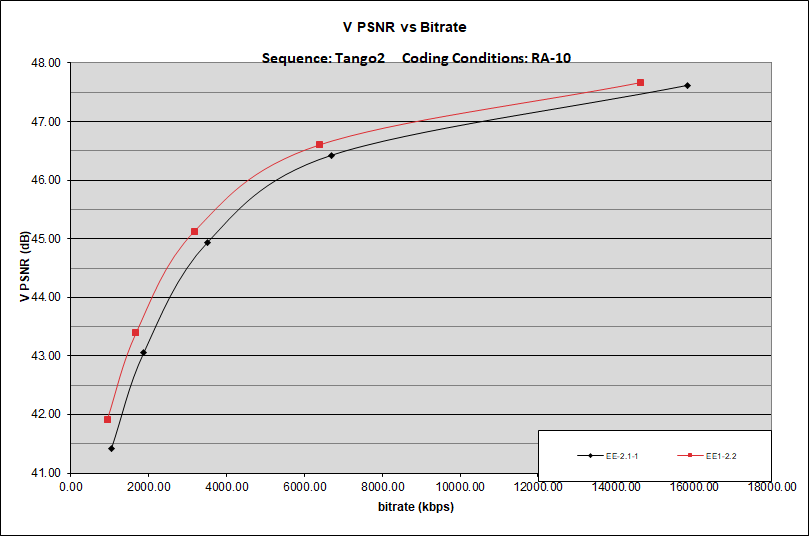
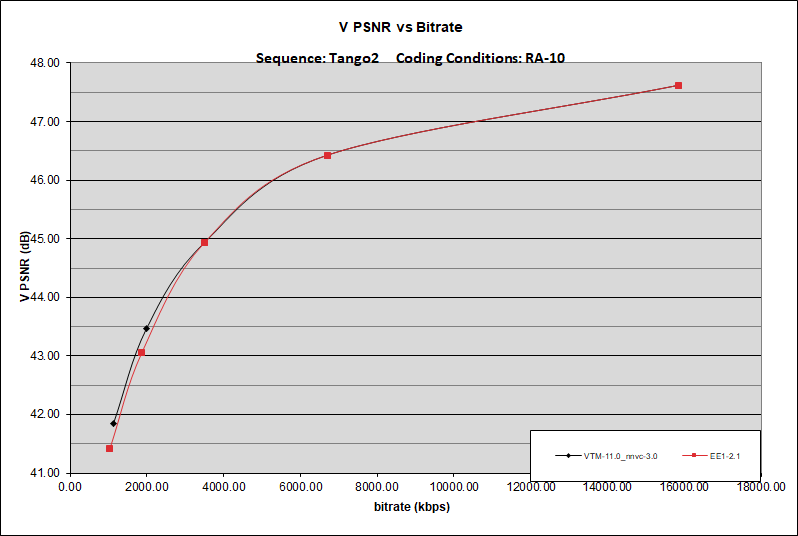
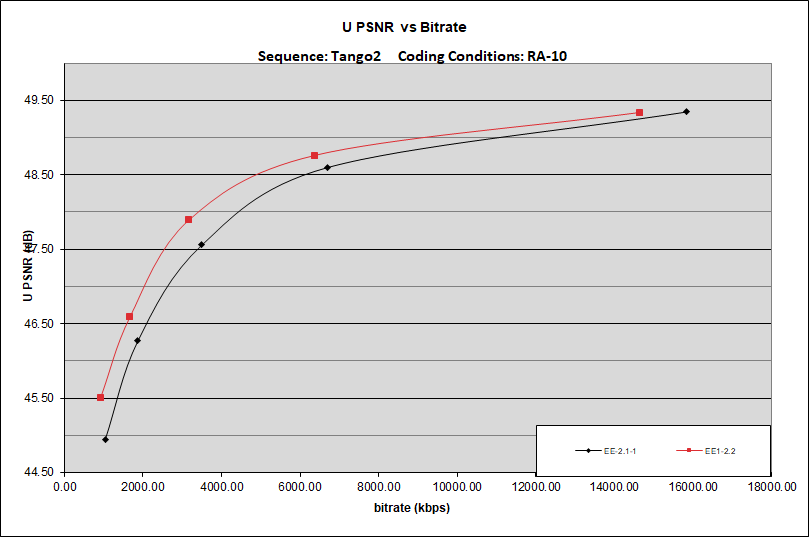
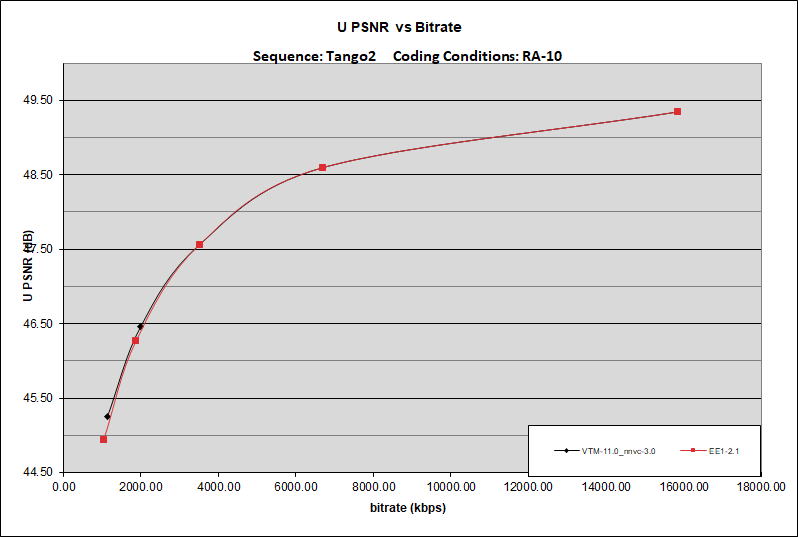
In **EE1-2.1** adaptive resolution coding is used in combination with “no NN-based” filters: RPR filters and long-tap post filter (8-tap for luma, 6-tap for chroma). Coding resolution is selected adaptively w/o multi-pass coding. At least 0.5% gain can be achieved by extending coded picture selection by scaling factor 1.25 and 1.5 additionally to 1.0 and 2.0.

Other three tests in this category use NN-based re-sampling filters. Test results for RA configuration (average among all resolutions, not only 4K) are summarized in Table 4. Figure 5 shows example of RD-curves for VVC full resolution coding, adaptive scaling factors 1.0 or 2.0 with RPR and “NN-based filters. Curves are (almost) not crossing, quality level is close enough, bit-rate saving is obvious.

*Table 4 Adaptive resolution coding with “non-NN based” (“classical”) and NN-based filters.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| EE1 test | scale | filter | BD-rate-Y | BD-rate-U | BD-rate-V |
| **2.1-1** | 1.0, 2.0 | RPR | -0.7% | 0.5% | 0.6% |
| **2.1-2** | 1.0, 2.0, 1.25, 1.5 | RPR | -1.2% | 3.2% | 2.7% |
| **2.1-3** | 1.0, 2.0, 1.25, 1.5 | RPR & long | -1.2% | 2.9% | 2.4% |
| **2.2** | 1.0, 2.0 | NN-based, 469kMAC/pxl | -2.1% | -1.0% | -0.8% |
| **2.3** | 1.0, 2.0 | NN-based, 926kMAC/pxl | -0.2% | 1.3% | 1.8% |
| **2.4** | 1.0, 2.0 | NN-based, 964kMAC/pxl | -1.0% | -0.2% | 0.5% |



**

*Figure 5 Example of RD-curves for RA, Tango: VTM anchor (full size coding), EE1-2.1 (adaptive scale 1.0 or 2.0), RPR filters and EE2.2 (adaptive scale 1.0 or 2.0), NN-based filters*

It is noted that the average results reported above are over all sequences in all classes. In case of 4K content, gains are higher.

2.2 has the most attractive gain, better than RPR filters and less complex/better than other NN based proposals. Both training and inference have been cross-checked.

Recommendation from EE1 summary report:

* Adopt EE1-2.2 for the next version of NNVC SW as an example of adaptive resolution coding with NN-based re-sampler

Decision: Adopt JVET-AC0196 into NNVC 4.0 (EE1-2.2)

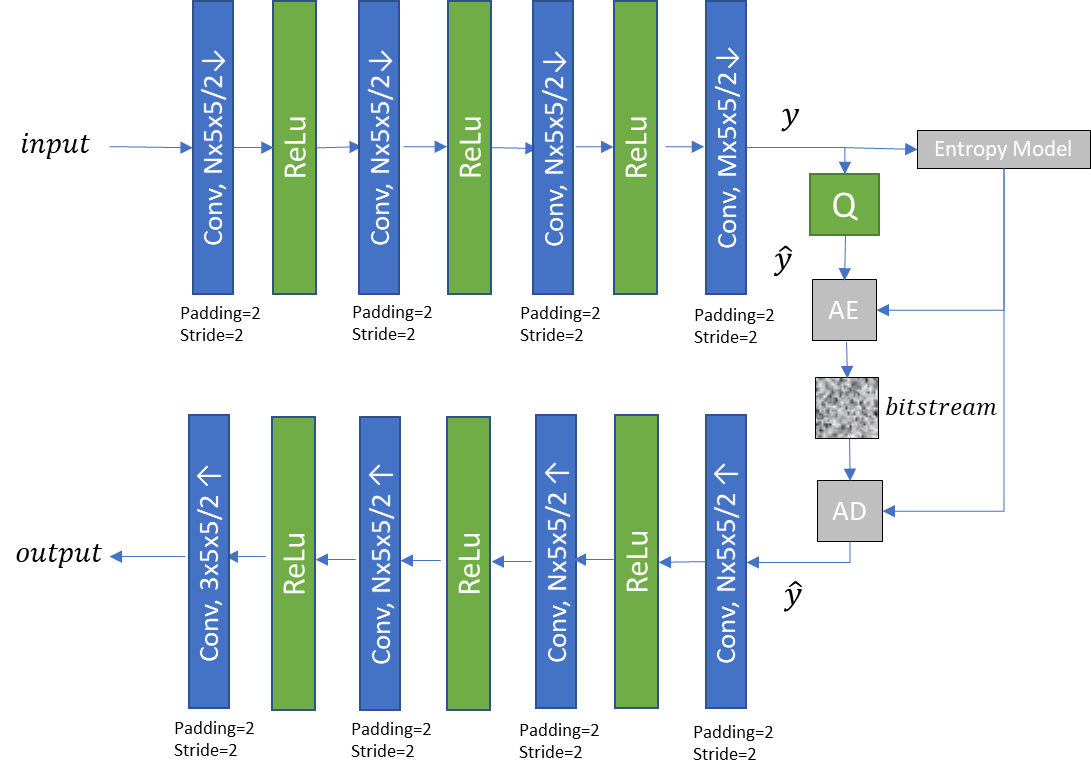
It was pointed out that the usage of RPR filters as a post processing upsampler (with GOP-wise switching) is already implemented in VTM 19 (from JVET-AB0080), and an equivalent implementation should be incuded in the NNVC software as an additional comparison point. This should be done as part of the integration of EE1-2.2.

It was also requested to provide the option of an integer implementation using SADL.

A comment was made that even better performance might be achieved when a downsampler that matches the upsampler (e.g., based on NN technology) would be used, i.e. non-normative preprocessing

1. **NN-Intra coding**

In test **EE1-3.1** one well known end-to-end AI based picture coding (Figure 6) was implemented with SADL, tested together with RDOQ. Test set is KODAK, anchor is not AhG 11 anchor, so results of this test are not shown in EE1 summary. Through this test it was shown that quantization aware training allows solving performance problem at high rate, which shows up if static quantization is used.



*Figure 6 Overview of the model architecture. N=128 and M=192.*

Test **EE1-3.2** studies NN-based Intra, block level NN-tool incorporated into VVC hybrid coding. Inference cross-check for this technology was done few meetings ago. This is the time for training cross-check (JVET-AC0290) and combination test with NNVC filter set#0 and NNVC filter set #1. The cross check was divided into three evaluation points. Two are finished and the last one is still running.

Test results are summarized in Table 5. There is no conflict in performance between NN-based Intra and filters, gain is more or less additive. This tool can be recommended to adoption to NNVC if training cross-check is successful.

*Table 5 Combination results for NN-based Intra and NNVC filters.*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| NN-Intra | NN-filter | # params | kMAC/pxl | BD-rate vs VVC (RA cfg) | | | BD-rate vs VVC (all intra cfg) | | |
| **Y** | U | V | **Y** | U | V |
| YES | no | 1.5 | 7.8 | **-1.6%** | -1.1% | -1.2% | **-3.4%** | -3.6% | -3.7% |
| NO | filter set#0 | 1.9 | 615 | **-8.9%** | -18.5% | -19.2% | **-6.5%** | -15.5% | -16.6% |
| YES | filter set#0 | 3.4 | 623 | **-10.3%** | -19.4% | -20.2% | **-9.6%** | -18.6% | -19.7% |
| NO | filter set#1 | 6.2 | 682 | **-9.5%** | -20.7% | -20.4% | **-7.5%** | -20.1% | -20.6% |
| YES | filter set#1 | 7.7 | 690 | **-10.9%** | -21.6% | -21.4% | **-10.5%** | -22.9% | -23.3% |

Recommendation from EE summary:

* Adopt EE1-1.3.2 for the next version of NNVC SW as an example of NN-based Intra tool

Candidate for adoption. See further notes unde JVET-AC0116.

### EE1 contributions: Neural network-based video coding (17)

Beyond the EE summary report JVET-AC0023, selected contributions in this area were discussed at 2100–2300 UTC on Thursday 12 Jan. 2023 (chaired by JRO), and at 2320–0020 UTC on Friday 13 Jan. 2023 (chaired by JRO).

For actions decided to be taken, see section 5.2.1, unless otherwise noted.

[JVET-AC0051](https://jvet-experts.org/doc_end_user/current_document.php?id=12236) EE1-2.3: RPR-Based Super-Resolution Guided by Partition Information Combined with GOP Level Adaptive Resolution [Q. Han, C. Jung (Xidian Univ.), Y. Liu, M. Li (OPPO), J. Nam, S. Yoo, J. Lim, S. H. Kim (LGE)]

This contribution reports the EE1-2.3 test results, which is a combination of JVET-AB0076 and JVET-JVET-AB0102 test 2.1.1. At each GOP, the encoder can adaptively select a scale factor from ×1.0 and ×2.0 and CNN-based super-resolution is utilized for the latter case. Compared with VTM-11.0-NNVC-3.0, the experimental results show {-4.78%, 2.36%, 4.53%} and {-0.72%, 2.71%, 4.27%} BD-rate changes on average (A1 and A2) for {Y, Cb, Cr} under AI and RA configurations, respectively.

[JVET-AC0222](https://jvet-experts.org/doc_end_user/current_document.php?id=12426) Crosscheck of JVET-AC0051 (EE1-2.3: RPR-Based Super-Resolution Guided by Partition Information Combined with GOP Level Adaptive Resolution) [D. Liu (Ericsson)] [late]

[JVET-AC0052](https://jvet-experts.org/doc_end_user/current_document.php?id=12237) EE1-2.4: CNN filter Based on RPR-based SR Combined with GOP Level Adaptive Resolution [S. Huang, C. Jung (Xidian Univ.), Y. Liu, M. Li (OPPO), J. Nam, S. Yoo, J. Lim, S. H. Kim (LGE)]

This contribution reports the EE1-2.4 test results, which is a combination of JVET-AB0093 and JVET-Z0065 test 2.1.1. At each GOP level, the encoder can adaptively select a scale factor from 1.0x and 2.0x and CNN-based super-resolution is utilized for is the latter case. Compared with VTM-11.0-nnvc-2.0, the test 2.4.1 experimental results show {-4.14%(Y), -0.33%(U), -0.22%(V)} and {-3.73%(Y), -3.07%(U), -1.15%(V)} and the test 2.4.2 experimental results show {-4.98%(Y), 0.07%(U), -0.78%(V)} and {-3.73%(Y), -3.07%(U), -1.15%(V)} BD-rate gains on average (A1 and A2 classes), under AI and RA configurations.

As a general comment, the constraint of 10% rate matching and showing PSNR graphs allows much better interpretation of SR results.

It was also commented that it would be beneficial to provide SSIM resuls in SR proposals.

[JVET-AC0223](https://jvet-experts.org/doc_end_user/current_document.php?id=12427) Crosscheck of JVET-AC0052 (EE1-2.4: CNN filter Based on RPR-based SR Combined with GOP Level Adaptive Resolution) [D. Liu (Ericsson)] [late]

[JVET-AC0055](https://jvet-experts.org/doc_end_user/current_document.php?id=12241) EE1-1.11: Content-adaptive post-filter [M. Santamaria, R. Yang, F. Cricri, J. Lainema, H. Zhang, R. G. Youvalari, M. M. Hannuksela (Nokia)]

This contribution reports the results of tests EE1-1.11, related to the content-adaptive CNN post-filter presented in JVET-AB0048. The signalling involves the NN post-filter characteristics and activation SEI messages in JVET-Z0244. The inference was done using float and int16 precision and SADL library. EE1-1.11 is evaluated against VTM 11.0 NNVC 3.0 in the RA configuration. It is reported that the overall coding gains (for int16 inference) are -5.03% (Y), -18.51% (Cb), -16.87% (Cr).

Overfitting is used (based on entire sequence characteristics, and sending an update over a base filter).

Status of cross-check: Several issues have been identified. It is not possible using parallel encoding in RA, which also causes some issues with extracting training data.

Decision: Adopt JVET-AC0055 to NNVC 4.0 software and description document JVET-AC2019.

[JVET-AC0331](https://jvet-experts.org/doc_end_user/current_document.php?id=12536) Crosscheck of JVET-AC0055 (EE1-1.11: Content-adaptive post-filter) [J. Ström (Ericsson)] [late]

It was reported during session 27 that the cross-check was successfully completed.

[JVET-AC0056](https://jvet-experts.org/doc_end_user/current_document.php?id=12252) EE1-3.1 CompressAI models integration using SADL [F. Galpin, F. Lefebvre, F. Racapé (InterDigital)]

Already discussed in context of EE report.

[JVET-AC0257](https://jvet-experts.org/doc_end_user/current_document.php?id=12462) Crosscheck of JVET-AC0056 (EE1-3.1 CompressAI models integration using SADL) [T. Chujoh (Sharp)] [late]

[JVET-AC0063](https://jvet-experts.org/doc_end_user/current_document.php?id=12265) EE1-1.3: On chroma order adjustment in NNLF [Z. Dai, Y. Yu, H. Yu, D. Wang (OPPO)]

Already discussed in context of EE report.

[JVET-AC0271](https://jvet-experts.org/doc_end_user/current_document.php?id=12476) Crosscheck of JVET-AC0063 (EE1-1.3: On chroma order adjustment in NNLF) [R. Chang (Tencent)] [late]

[JVET-AC0064](https://jvet-experts.org/doc_end_user/current_document.php?id=12266) EE1-1.4: On adjustment of residual for NNLF [Z. Dai, Y. Yu, H. Yu, D. Wang (OPPO)]

Already discussed in context of EE report.

[JVET-AC0272](https://jvet-experts.org/doc_end_user/current_document.php?id=12477) Crosscheck of JVET-AC0064 (EE1-1.4: On adjustment of residual for NNLF) [R. Chang (Tencent)] [late]

[JVET-AC0286](https://jvet-experts.org/doc_end_user/current_document.php?id=12491) Crosscheck of JVET-AC0064 (EE1-1.4: On adjustment of residual for NNLF) [K. Jia (Bytedance)] [late]

[JVET-AC0078](https://jvet-experts.org/doc_end_user/current_document.php?id=12280) EE1-2.1: Updates on RPR encoder and post-filter [J. Nam, S. Yoo, J. Lim, S. Kim (LGE)]

Already discussed in context of EE report.

[JVET-AC0290](https://jvet-experts.org/doc_end_user/current_document.php?id=12495) Cross-check of JVET-AC0078 (EE1-2.1: Updates on RPR encoder and post-filter) [K. Andersson (Ericsson)] [late]

[JVET-AC0089](https://jvet-experts.org/doc_end_user/current_document.php?id=12291) EE1-1.5: Combined intra and inter models for luma and chroma [D. Liu, J. Ström, M. Damghanian, P. Wennersten, K. Andersson (Ericsson)]

[JVET-AC0273](https://jvet-experts.org/doc_end_user/current_document.php?id=12478) Crosscheck of JVET-AC0089 (EE1-1.5: Combined intra and inter models for luma and chroma) [R. Chang (Tencent)] [late]

[JVET-AC0106](https://jvet-experts.org/doc_end_user/current_document.php?id=12308) EE1-1.10: Complexity Reduction on Neural-Network Loop Filter [J. N. Shingala, A. Shyam, A. Suneia, S. P. Badya (Ittiam), T. Shao, A. Arora, P. Yin, S. McCarthy (Dolby)]

This contribution reports the experimental results of EE1-1.10. The models for complexity reduction in this EE test are originally proposed in JVET-AA0080 and JVET-AB0136, with the introduction of parallel fusion of deblocked samples and NNLF outputs for further improvement. In EE1-1.10.1, the model with CP decomposition plus fusing adjacent 1x1 convolution is tested. In EE1-1.10.2, the model with CP decomposition, fusing adjacent 1x1 convolution and split architecture for luma and chroma components is tested. The models are implemented on top of NNVC common software NNVC-3.0. They are also implemented with fixed point using SADL. BD-Rate results are reported compared to NNVC-3.0 anchor. While the JVET-X0140 low complexity model achieves about 5.16% luma gain for AI and 4.98% for RA with worse case block level complexity of 33.6 KMAC/Pixel, the proposed models can achieve better trade-off with lower complexity. The AI and RA results of {Y, Cb, Cr} BD-Rate gain are as follows:

EE1-1.10.1: CP-decomposition with fusing, worst case block level complexity of 16.2 KMAC/Pixel

* LibTensorflow, fp32: AI {-4.65%, -5.66%, -5.03%}, RA {-5.05%, -6.29%, -4.46%}
* SADL, fp32: AI {-4.64%, -5.62%, -4.97%}, RA {-5.02%, -6.36%, -4.38%}
* SADL, int32: AI {-4.62%, -5.62%, -4.96%}, RA {-5.01%, -6.24%, -4.32%}

EE1-1.10.2: CP-decomposition with fusing and split luma chroma (24L, 8C), worst case block level complexity of 17.7 KMAC/Pixel

* LibTensorflow, fp32: AI {-4.94%, -7.76%, -7.69%}, RA {-5.43%, -7.79%, -6.57%}
* SADL, fp32: AI {-4.91%, -7.21%, -6.97%}, RA {-5.37%, -7.61%, -6.31%}
* SADL, int32: AI {-4.88%, -7.24%, -6.99%}, RA {-5.34%, -7.64%, -6.34%}

Was presented in session 12.

The method uses 4 model, each of size about 57K parameters, 17.7 kMAC/Pix

Study in next EE:

* Training crosscheck
* Int16 implementation
* If possible, SIMD implementation in SADL

Possibility of further reduction of number of computations should be considered.

[JVET-AC0241](https://jvet-experts.org/doc_end_user/current_document.php?id=12445) Crosscheck of JVET-AC0106 (EE1-1.10: Complexity Reduction on Neural-Network Loop Filter) [Y. Li (Bytedance)] [miss] [late]

[JVET-AC0116](https://jvet-experts.org/doc_end_user/current_document.php?id=12318) EE1-3.2: neural network-based intra prediction with learned mapping to VVC intra prediction modes [T. Dumas, F. Galpin, P. Bordes (InterDigital)]

This contribution reports the results of EE1-3.2.2. In EE1-3.2.2, VTM-11-NNVC with Filter-Set-1 [4] activated, Filter-Set-0 [2, 3] deactivated, and the low-complexity version of the neural network-based intra prediction mode activated [1], all the neural networks being in 16-bit signed integer, must be run. This run aims at showing that the BD-rate gains of the low-complexity version of the neural network-based intra prediction mode inside VTM-11-NNVC are almost equivalent when this neural network-based intra prediction mode is combined with either Filter-Set-0 or Filter-Set-1.

It is reported that, on top of VTM-11-NNVC with Filter-Set-1 activated and Filter-Set-0 deactivated, the low-complexity version of the neural network-based intra prediction mode yields -3.21%, -3.52%, -3.38% and -1.54%, -1.01%, -1.25% of average BD-rate gains in AI and RA respectively.

On top of VTM-11-NNVC with Filter-Set-1 deactivated and Filter-Set-0 activated, the low-complexity version of the neural network-based intra prediction mode yields -3.35%, -3.60%, -3.65% and -1.57%, -1.08%, -1.20% of average DB-rate gains in AI and RA respectively.

Was presented in session 12.

Performance in combination with filter sets #0 and #1 almost identical.

Both training and inference cross-checks confirmed the results

Decision: Adopt JVET-AC0116

[JVET-AC0238](https://jvet-experts.org/doc_end_user/current_document.php?id=12442) Crosscheck of JVET-AC0116 (EE1-3.2: neural network-based intra prediction with learned mapping to VVC intra prediction modes) [Y. Li (Bytedance)] [late]

[JVET-AC0270](https://jvet-experts.org/doc_end_user/current_document.php?id=12475) Crosscheck of EE1-3.2 (JVET-AC0116: neural network-based intra prediction with learned mapping to VVC intra prediction modes) [M. Damghanian, J. Ström (Ericsson)] [late]

This document is about training crosscheck. It was confirmed during session 27 that the crosscheck was completed successfully (version 4 of JVET-AC0270)

[JVET-AC0274](https://jvet-experts.org/doc_end_user/current_document.php?id=12479) Crosscheck of JVET-AC0116 (EE1-3.2.2: Low-complexity version of the neural network-based intra prediction mode in 16-bit signed integer) [T. Shao (Dolby)] [late]

[JVET-AC0298](https://jvet-experts.org/doc_end_user/current_document.php?id=12503) Cross-check of EE1-3.2 (Neural network-based intra prediction with learned mapping to VVC intra prediction modes) [M. Abdoli (IRT b-com)] [late]

[JVET-AC0118](https://jvet-experts.org/doc_end_user/current_document.php?id=12320) EE1-1.8: QP-based loss function design for NN-based in-loop filter [C. Zhou, Z. Lv, J. Zhang (vivo)]

Already discussed in context of EE report.

[JVET-AC0215](https://jvet-experts.org/doc_end_user/current_document.php?id=12419) Crosscheck of JVET-AC0118 (EE1-1.8: QP-based loss function design for NN-based in-loop filter) [C. Lin (Bytedance)] [late]

[JVET-AC0143](https://jvet-experts.org/doc_end_user/current_document.php?id=12345) EE1-1.6: NN chroma model without partitioning input [J. Ström, D. Liu, K. Andersson, P. Wennersten, M. Damghanian, R. Yu (Ericsson)]

No need for presentation – th adopted EE1-1.5 removes the intra model entirely there is no need to discuss removing certain inputs for the chroma intra model.

[JVET-AC0234](https://jvet-experts.org/doc_end_user/current_document.php?id=12438) Cross-check of JVET-AC0143 (EE1-1.6: NN chroma model without partitioning input) [M. Santamaria (Nokia)] [late]

[JVET-AC0155](https://jvet-experts.org/doc_end_user/current_document.php?id=12359) EE1-1.9: Reduced complexity CNN-based in-loop filtering [S. Eadie, M. Coban, M. Karczewicz (Qualcomm)]

In this contribution, the floating- and fixed- point, SADL-based inference results are reported for the simplified N=24 backbone block model introduced in JVET-AB0164. The BD-rate savings for Y, Cb, Cr components in RA and AI configurations, respectively, are {-10.03%, -20.26%, -19.97%} and {-7.88%, -17.08%, -18.33%} for floating-point inference, and {-10.19%, -20.65%, -20.47%} and {-7.24%, -17.88%, -19.10% } for fixed-point, 16 bit inference. Performance improvement in RA which is untypical; generally, the performance difference between float and integer is larger than for many other architectures. The reasons for this should be further studied.

The models were not re-trained after quantization, just trained as float and then rounded to integer.

Two different models for B and I pictures are used (as in filter set #1)

Partition information is used as input, which is likely unnecessary as per experience made with filter set 1.

Further simplifications (including reduction to ome model), performance improvements, and training crosscheck to be investigated in EE.

[JVET-AC0254](https://jvet-experts.org/doc_end_user/current_document.php?id=12459) Crosscheck of JVET-AC0155 (EE1-1.9: Reduced complexity CNN-based in-loop filtering) [Y. Li (Bytedance)] [miss] [late]

[JVET-AC0177](https://jvet-experts.org/doc_end_user/current_document.php?id=12381) EE1-1.7: Deep In-Loop Filter with Additional Input Information [Y. Li, [K. Zhang](mailto:zhangkai.video@bytedance.com), L. Zhang (Bytedance)]

This contribution presents the EE test results of JVET-AB0073. The proposed technique includes two aspects:

Aspect #1: Feeding pictures from reference picture lists into the CNN filter as additional information.

Aspect #2: Flipping the input/output samples of the CNN filter.

The proposed CNN model is implemented on top of NNVC-3.0 using SADL int16 precision.

BD-rate changes of {Y, Cb, Cr} on top of NNNC-3.0 with filter set #1 enabled and NNVC-3.0 anchor (NN filters are disabled) are reportedly summarized as below:

EE1-1.7.1 (aspect #1 + aspect #2):

Compared with NNVC-3.0 filter set #1, RA: {-0.67%, 0.03%, 0.05%}, EncT 92%, DecT 90%

Compared with NNVC-3.0 anchor, RA: {-10.09%, -20.68%, -20.36%}, EncT 168%, DecT 34064%

EE1-1.7.2 (aspect #1):

Compared with NNVC-3.0 filter set #1, RA: {-0.58%, -0.02%, -0.04%}, EncT 91%, DecT 89%

Compared with NNVC-3.0 anchor, RA: {-10.01%, -20.72%, -20.43%}, EncT 167%, DecT 33870%

EE1-1.7.3 (aspect #2):

Compared with NNVC-3.0 filter set #1, RA: {-0.12%, 0.03%, 0.07%}, EncT 84%, DecT 92%

Compared with NNVC-3.0 anchor, RA: {-9.59%, -20.68%, -20.34%}, EncT 154%, DecT 34758%

Was presented in session 12.

JVET-AC0310 is a related contribution targeting the reduction of the number of models with a similar approach as had been used in EE1-1.5

It was asked why the input is the unfiltered reconstructed picture (before deblocking)? The proponent replies that the difference in performance is minor, but extraction of data for training is simpler.

[JVET-AC0209](https://jvet-experts.org/doc_end_user/current_document.php?id=12413) Crosscheck of JVET-AC0177 (EE1-1.7: Deep In-Loop Filter with Additional Input Information) [C. Zhou (vivo)] [late]

[JVET-AC0194](https://jvet-experts.org/doc_end_user/current_document.php?id=12398) EE1-1.1: More refinements on NN based in-loop filter with a single model [R. Chang, L. Wang, X. Xu, S. Liu (Tencent)]

Already discussed in context of EE report.

[JVET-AC0278](https://jvet-experts.org/doc_end_user/current_document.php?id=12483) Crosscheck of JVET-AC0194 (EE1-1.1: More refinements on NN based in-loop filter with a single model) [Z. Xie (OPPO)] [late]

[JVET-AC0195](https://jvet-experts.org/doc_end_user/current_document.php?id=12399) EE1-1.2: encoder-only optimization for NN based in-loop filter with a single model [R. Chang, L. Wang, X. Xu, S. Liu (Tencent)]

In this contribution, an encoder-only tool for EE1-1.1 is studied on top of NNVC-3.0 (filter set #0). The filters in the RDO process are retained based on the network structure of in-loop filter in EE1-1.1. Then, the well-trained models are used to further enhance the solution in EE1-1.1. Based on the NNVC-3.0, the test results are shown in order of RA and AI configurations as follows.

Test 1.2.1 (Base model from NNVC-3.0 filter set #0):

RA: -9.17% -18.51% -19.19% EncT: 159% DecT: 55431%

AI : -6.96% -15.56% -16.67% EncT: 173% DecT: 31788%

Test 1.2.2 (Base model from EE1-1.1.3):

RA: -10.24% -21.40% -20.91% EncT: 163% DecT: 55880%

AI : -8.03% -16.77% -17.63% EncT: 173% DecT: 31640%

Compared with NNVC-3.0 (filter set #0) under RA and AI configurations, additional 0.26% (from 8.91% to 9.17%) and 0.45% (from 6.51% to 6.96%) BD-rate luma gains are observed. Compared with EE1-1.1.3 under RA and AI configurations, additional 0.26% (from 9.98% to 10.24%) and 0.42% (from 7.61% to 8.03%) BD-rate luma gains are observed.

In terms of decoder side, there is no any increase in both MAC result and memory size. For the runtime at encoder side of Test 1.2.1, there are about 9% (from 150% to 159%) and 25% (from 148% to 173%) increase under RA and AI configurations, respectively. As for the runtime at encoder side of Test 1.2.2, there are about 9% (from 154% to 163%) and 26% (from 147% to 173%) increase under RA and AI configurations, respectively.

Was presented in session 7.

Benefit from more extensive training (with fast data loading of pre-generated training patches), modifying the training data set, adjusting balance between I and B frames for which the same model is used, and using weighted combination L1/L2 loss function. For tests 1.1.3 and 1.2.2 (iterative training), no training cross-check was performed.

The real difference of 1.2.2 relative to 1.2.1 is not precisely described in the contribution. From the explanation given, this means that the model is re-trained using data that were filtered by a first trained model in a second pass. This is however a common approach that has been applied to other architectures

Total size of data set that was used? Approx. 700 GB storage size is needed.

As a general remark, it was suggested to make it mandatory to use implementation in NNVC in EE1 (unless for cases e.g. end-to-end architectures that would not be compatible with NNVC)

[JVET-AC0279](https://jvet-experts.org/doc_end_user/current_document.php?id=12484) Crosscheck of JVET-AC0195 (EE1-1.2: encoder-only optimization for NN based in-loop filter with a single model) [Z. Xie (OPPO)] [late]

[JVET-AC0196](https://jvet-experts.org/doc_end_user/current_document.php?id=12400) EE1-2.2: GOP Level Adaptive Resampling with CNN-based Super Resolution [R. Chang, L. Wang, X. Xu, S. Liu (Tencent)]

Already discussed in context of EE report.

[JVET-AC0221](https://jvet-experts.org/doc_end_user/current_document.php?id=12425) Crosscheck of JVET-AC0196 (EE1-2.2: GOP Level Adaptive Resampling with CNN-based Super Resolution) [D. Liu (Ericsson)] [late]

### EE1 related contributions: Neural network-based video coding (8)

Contributions in this area were discussed in session 12 at 0025+1–0130 UTC on Friday 13 Jan. 2023 (chaired by JRO), and in session 19 at 2100–2215 UTC on Tuesday 17 Jan. 2023 (chaired by JRO).

[JVET-AC0066](https://jvet-experts.org/doc_end_user/current_document.php?id=12268) EE1-related: Improvement on EE1-1.7 [Z. Xie, Y. Yu, H. Yu, D. Wang (OPPO)]

In EE1-1.7, two aspects were proposed to improve the coding performance of NNVC-3.0 filter set #1. First, it is proposed to feed collocated blocks from reference frames into the CNN-based in-loop filtering. Second, the input samples of CNN can be flipped and the output samples of CNN can be flipped back. This contribution proposes a slice-level model switching method to further improve the performance of filter set #1. Compared with EE1-1.7.1, the performance is reported as below:

RA: Y 0.00%, U 0.00%, V 0.00%, EncT 100%, DecT 100%;

LDB: Y -0.93%, U 0.01%, V 0.06%, EncT 103%, DecT 100%.

For RA, the method is not used.

The flipping is already used in EE1-1.7

The “normative” change would be a slice-level flag for signalling the model, and at the decoder the POC needs to be checked. This would be some kind of HLS control. Otherwise, the proposal is more about encoder optimization on top of EE1-1.7.

Several experts supported this as an interesting idea to improve the performance in LDB.

This could be a candidate for adoption provided that EE1-1.7 would be adopted. It was reported verbally during session 27 that the gains are also retained (or with partial results slightly higher) for LDB when the method id implemented on top of the combination 1.1.5/1.1.7 (as per JVET-AC0310). It was requested providing these results in a new input (see JVET-AC0355).

[JVET-AC0355](https://jvet-experts.org/doc_end_user/current_document.php?id=12561) EE1-related: Improvement over combination of EE1-1.5 and EE1-1.7 Z. Xie, Y. Yu, H. Yu, D. Wang (OPPO)

This contribution reports the results of proposed method in JVET-AC0066 on top of JVET-AC0310, which is a combination test of EE1-1.5 and EE1-1.7. EE1-1.5.3 proposes to reduce model memory by using a shared model for the in-loop filtering of different types of slices. EE1-1.7.1 proposes to improve the in-loop filtering performance by introducing an additional CNN model. JVET-AC0066 proposes an adaptive slice-level model switching method to further enhance the performance of EE1-1.7. On top of the combination of EE1-1.5 and EE1-1.7, the results of the proposed method can be summarized as below:

RA class D: {Y 0.00%, U 0.01%, V 0.00%, EncT 101%, DecT 99%};

LB class C: {Y -1.48%, U 0.54%, V -0.24%, EncT 110%, DecT 100%};

LB class E: {Y -1.55%, U -0.30%, V -0.39%, EncT 106%, DecT 114%};

LB class D: {Y -1.35%, U 0.50%, V -0.81%, EncT 110%, DecT 100%}.

Class B not finished yet

Results on the other classes are consistent, and the gain in case of LDB has become even higher than in the combination on top of 1-1.7 standalone, however only for classes C and D (class E more or less unchanged). This might be due to the fact that class E does not have much motion, such that the on/off switch of temporal input is not needed

A cross-check is currently running. The cross-checker confirms that the integration is straightforward.

The method would be turned off by default (as 1-1.7 is).

It was suggested to be investigate in EE if the cross-check is finished successfully, investigate more about the origin of the additional gain and adopt it as part of 1-1.7 at the next meeting.

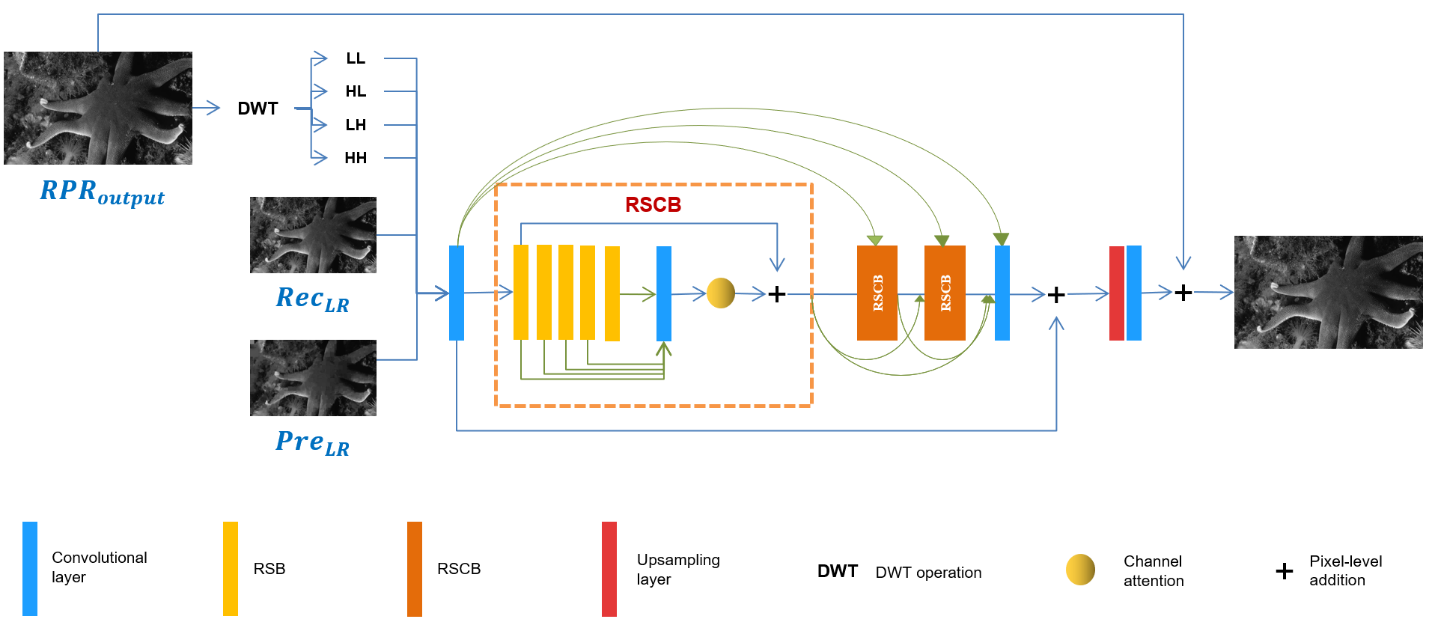
[JVET-AC0304](https://jvet-experts.org/doc_end_user/current_document.php?id=12509) Crosscheck of JVET-AC0066 (EE1-related: Improvement on EE1-1.7) [R. Chang (Tencent)] [late]

[JVET-AC0308](https://jvet-experts.org/doc_end_user/current_document.php?id=12513) Crosscheck of JVET-AC0066 (EE1-related: Improvement on EE1-1.7) [[Y. Li (Bytedance)](mailto:yue.li@bytedance.com)] [miss] [late]

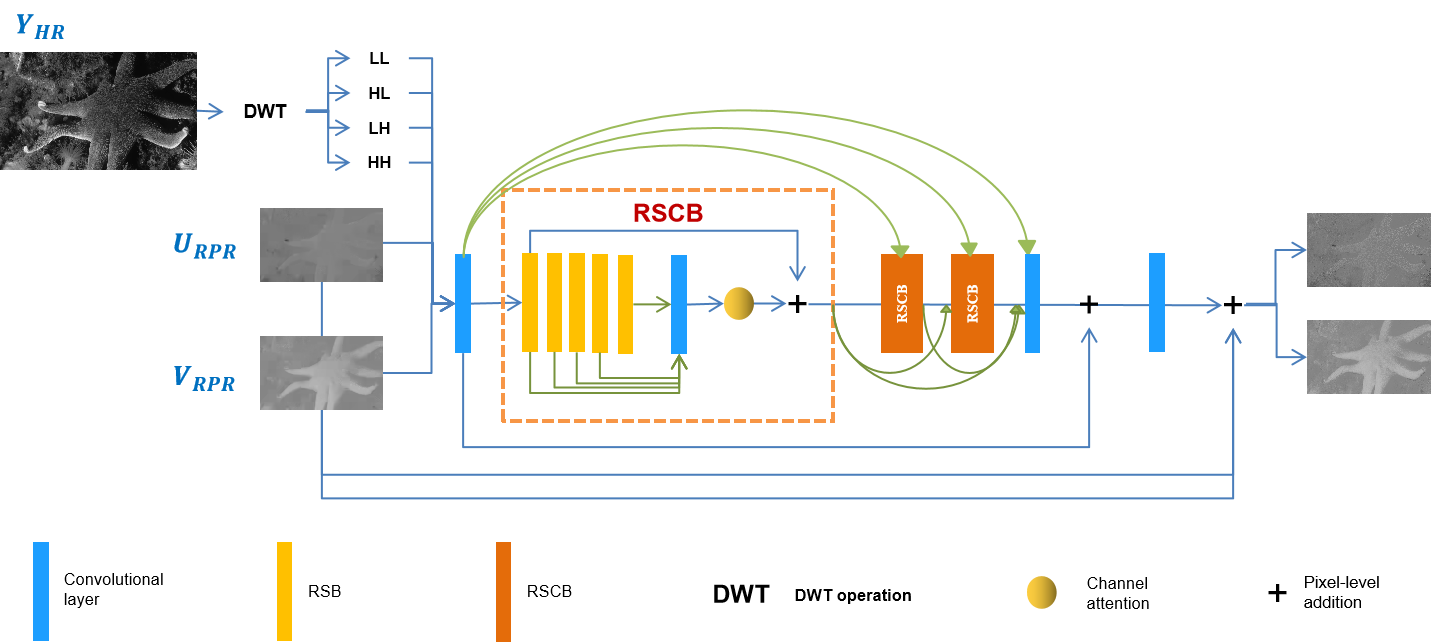
[JVET-AC0081](https://jvet-experts.org/doc_end_user/current_document.php?id=12283) EE1-related: A CNN Filter for RPR-Based Super-Resolution Using Wavelet Decomposition Combined with GOP Level Adaptive Resolution [H. Lan, C. Jung (Xidian Univ.), Y. Liu, M. Liu (OPPO), J. Nam, S. Yoo, J. Lim, S. H. Kim (LGE)]

This contribution proposes a CNN filter for RPR-based super-resolution using wavelet decomposition combined with GOP level adaptive resolution. The proposed CNN filter takes the LR reconstructed frame (), LR prediction frame () and RPR upsampled frame () as the input for RPR-based SR. Wavelet decomposition is adopted to make the same size as and and obtain the relationship between high frequency and low frequency components. The proposed CNN filter is combined with GOP level adaptive resolution in JVET-Z0065 to adaptively select a scale factor (×1.0 and ×2.0) at the GOP level. Experimental results show that the proposed CNN filter in Y channel achieves -4.68% and -0.30% BD-rate reductions in AI and RA configurations over VTM-11.0\_NNVC-2.0 anchor, respectively.

The proposed CNN filter takes three inputs of , and for RPR-based SR. Thus, this filter can learn the mapping relationship between LR and HR frames as well as recover textures while reducing blocking artifacts.



**Fig. 2** Illustration of the proposed network architecture for luma with wavelet decomposition. DWT: Discrete wavelet transform.



**Fig. 3** Illustration of the proposed network architecture for chroma with wavelet decomposition. DWT: Discrete wavelet transform.

Table 1. Network information of the proposed CNN filter testing in the training stage.

|  |  |  |
| --- | --- | --- |
| **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: |  |
|  |  |

Table 2: Network information for the proposed CNN filter testing in inference stage

|  |  |  |
| --- | --- | --- |
| **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 convolutions |
| 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: |  |
|  |  |

Model size 2.38 M parameters

>2000 kMAC/pixel

Gains reported in abstract are onöy for class A sequences. For the overall data set, gain would be 1.75%. This is less than the best performing method in EE1, which had 2.1% but only around 500 kMAC/pixel. No good tradeoff complexity/performance.

[JVET-AC0126](https://jvet-experts.org/doc_end_user/current_document.php?id=12328) EE1-related: Reduced complexity through channel redistribution in NN head [P. Wennersten, J. Ström, D. Liu (Ericsson)]

This contribution proposes to modify the number of channels in the head of the luma NN loopfilter to reduce the complexity. Comparing to the models in JVET-AB-EE1-1.5.3, it is reported that the proposed the luma model reduces the worst-case complexity by 9% (from 673 kMAC/pix to 615 kMAC/pix in block-basis), while the loss is 0.01% in RA. It is further reported that, also compared to the models in JVET-AB-EE1-1.5.3, the number of parameters goes down 2% (from 3.10 M to 3.03 M) and the training time is decreased by 12% (from 12 days to 10.5 days).

Result over NNVC-3.0, int16 Y: RA -9.60%, AI -7.31%, LDB -8.17%

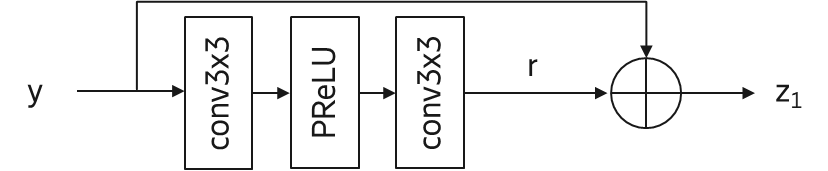
Result over NNVC-3.0 filter set#1, int16 Y: RA -0.13%, AI 0.23%, LDB 0.37%

Result over JVET-AC0089 (JVET-AB-EE1-1.5.3), int16 Y: RA 0.01%, AI 0.08%, LDB 0.12%

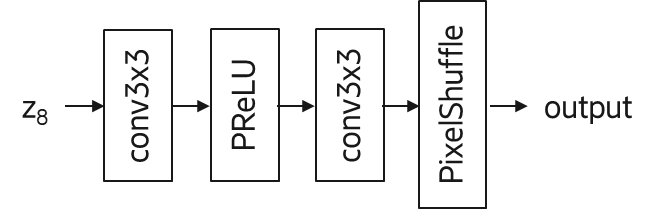
Figure 2 shows the k-th residual block. In this contribution, there are eight residual blocks. Figure 3 shows the last part of the network. These two are the same as in JVET-AB-EE1-1.5.3.



*Figure 1. Head of luma network. The inputs are combined to form the input y to the next part of the network.*



*Figure 2. The k-th residual block (k=0..7). The output y of the head is fed into a first residual block with input z0=y, which also takes the inputs rec, pred, part, bs, qp, and IPB. The output z1 is then fed into another such residual block.*



*Figure 3. The output of the last residual block is fed into this last part of the network.*

Input with 192 channels for reconstruction, 24 channels for prediction, 12 channels bs/qp/IPB

|  |  |  |
| --- | --- | --- |
| **Network Information in Training Stage** | | |
| Mandatory | GPU Type | GPU: NVIDIA A100-PCIE-40GB |
| Framework: | PyTorch v1.9 |
| Number of GPUs per Task | 1 |
|  |  |
| Epoch: | 340 |
| Batch size: | 32 |
| Loss function: | Weighted L1 and L2 |
| Training time: | 10.5 days |
| Training data information: | DIV2K, BVI-DVC |
| Training configurations for generating compressed training data (if different to VTM CTC): | VTM-11.0, JVET-W-EE1-1.6, qp {17, 22, 27, 32, 37, 42} |
| Optional |  |  |
| Number of iterations |  |
| Patch size | 256x256 |
| Learning rate: | 1e-4 and 1e-5 |
| Optimizer: | ADAM |
| Preprocessing: |  |
| Mini-batch selection process: |  |
| Other information: |  |
|  |  |

|  |  |  |
| --- | --- | --- |
| **Network Information in Inference Stage** | | |
| Mandatory | HW environment: | |
| GPU Type | N/A |
| Framework: | SADL [5] |
| Number of GPUs per Task | 0 |
|  |  |
| Total Parameter Number | See Table 1 |
| Parameter Precision (Bits) | Int16 |
| Memory Parameter (MB) | See Table 1 |
| Multiplay Accumulate (MAC) | See Table 1 |
| Optional |  |  |
| Total Conv. Layers |  |
| Total FC Layers |  |
| Total Memory (MB) |  |
| Batch size: | 1 |
| Patch size | 128x128, 256x256 |
| Changes to network configuration or weights required to generate rate points |  |
| Peak Memory Usage (Total) |  |
| Peak Memory Usage (per Model) |  |
| Border handling |  |
| Other information: |  |
|  |  |

It was commented that the approach to give different priority to the different inputs of the network (by varying number of channels) is an interesting approach to reduce complexity and increase performance, and that a similar approach could be used for filter set#0.

Support was expressed to investigate in EE. In this context, also the impact of the selection of the number of channels could be relevant. For example, it was pointed out that in SIMD implementation channel numbers which are powers of 2 are most effective to reduce run time.

[JVET-AC0312](https://jvet-experts.org/doc_end_user/current_document.php?id=12517) Cross-check of JVET-AC0126: EE1-related : reduced complexity through channel redistribution in NN head [T. Dumas (InterDigital)] [late]

[JVET-AC0156](https://jvet-experts.org/doc_end_user/current_document.php?id=12360) [EE1-related] RTNN: An In-loop Filter Based on Resblock and Transformer [H. Zhang, C. Jung (Xidian Univ.), Y. Liu, M. Li (OPPO)]

This contribution presents a convolutional neural network (CNN)-based in-loop filter. The proposed CNN filter explores the combination of Resblock and Transformer block (TB) to build a new backbone network with efficiency. Moreover, a new attention module is proposed to better refine features by introducing auxiliary information. In addition, the training strategy in JVET-AB0090 [1] is used to train the proposed CNN filter. Compared with EE1-anchor-2.0 (VTM-11.0\_NNVC-2.0), the proposed CNN filter achieves average {8.49%, 22.98%, 24.07%} BD-rate reductions for {Y, Cb, Cr} under AI configuration and average {9.23%, 22.47%, 22.31%} BD-rate reductions for {Y, Cb, Cr} under RA configuration.

Fig. 1 shows the feature extraction and reconstruction of the luminance component, while Fig. 2 shows the feature extraction and reconstruction of the chrominance component. First, the number of feature maps is allocated according to the richness of the information contained in the input to reduce the complexity while ensuring performance. Secondly, for the chroma component, according to different categories of input information, a gradual fusion architecture is used to obtain more accurate fusion features. Note that the luma model does not use partition information for B Slice.

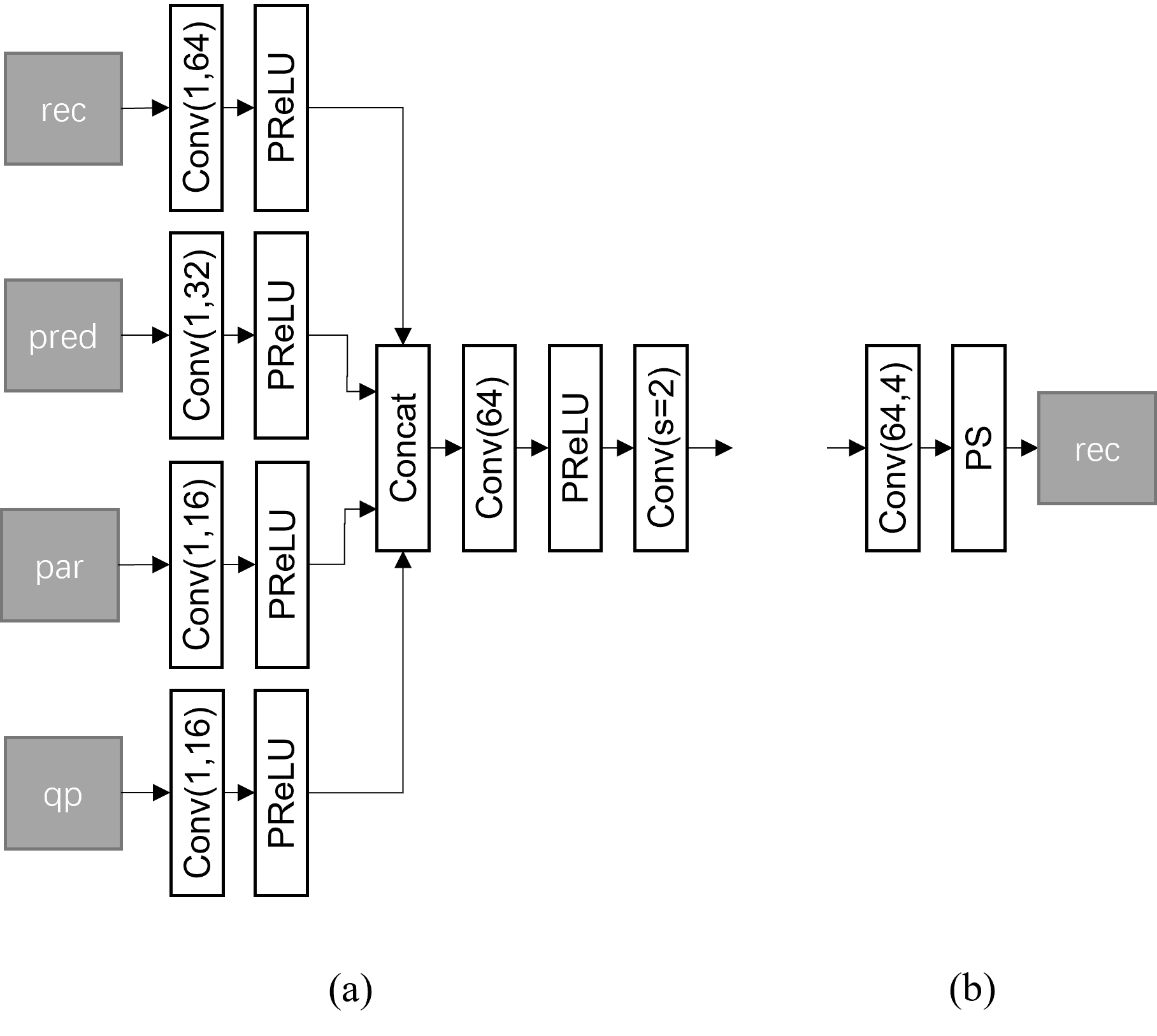


Fig. 1 Network architecture of feature extraction and reconstruction of luma component. (a) Feature extraction part. (b) Reconstruction part.

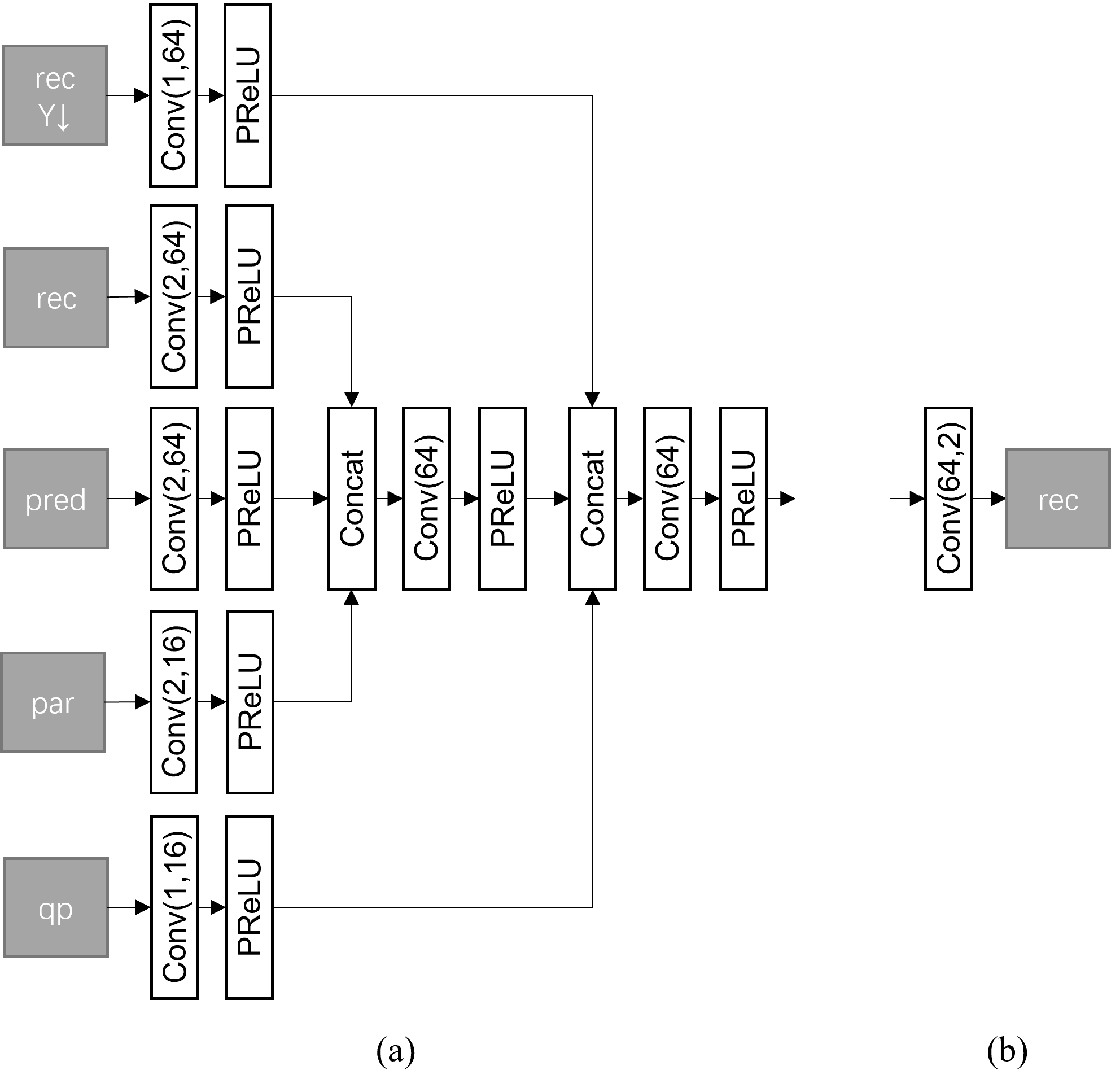


Fig. 2 Network architecture of feature extraction and reconstruction of chroma component. (a) Feature extraction part. (b) Reconstruction part.

Network architectures of the backbone are shown in Fig. 3. In addition, for the luma model, the backbone network includes 3 RABs and 6 TBs, while for the chroma model, the backbone network includes 1 RAB and 3 TBs.



Fig. 3 Network architectures of the proposed backbone. (a) Residual block. (b) Residual attention block (RAB). (c) Backbone. TB: Transformer block.

|  |  |  |
| --- | --- | --- |
| **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: | ~70h/luma model, ~30h/chroma 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: | L1 and L2 |
| Optional |  |  |
| Number of iterations |  |
| Patch size | 144x144 |
| Learning rate: | 1e-4 |
| Optimizer: | ADAM |
| Preprocessing: |  |
| 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) | 1.4M for luma, 0.47M for chroma |
| Total Number of Parameters (All Models) | 3.74M (4 models) |
| Parameter Precision (Bits) | 32 |
| Memory Parameter (MB) | 14.96M |
| Multiply Accumulate (MAC)/pixel | 353K for luma, 116K for chroma |
| Optional |  |  |
| Total Conv. Layers | 54 Depthwise convolution layers  174 Convolution layers  (4 models total) |
| Total FC Layers |  |
| 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: |  |
|  |  |

It was reported that the interface is similar to filter set#0.

Lower in complexity than the filters in NNVC3.0.

Interesting to investigate alternative architectures. Would implementation in SADL be possible? Likely, but would need to be checked in detail.

How much benefit due to attention blocks and transformer blocks? Was not studied in detail.

Investigate in EE.

[JVET-AC0178](https://jvet-experts.org/doc_end_user/current_document.php?id=12382) EE1-related: 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 built based on basic residual blocks with wide activation and large receptive field. The proposed filter is obtained by revising the one presented in JVET-AB0179 with two aspects of changes. First, a unified model is used to deal with different types of slices. Second, the chroma model is simplified. It is reported that kMAC/pixel is reduced by 17% and the model storage is reduced by 66% compared with filter set #1. Compared with NNVC-3.0 filter set #1 and NNVC-3.0 anchor, the proposed filter shows BD-rate changes reported as below:

Compared with NNVC-3.0 filter set #1:

RA: {-1.72%, -3.05%, -2.41%}, LB: {-2.72%, -8.64%, -10.17%}, AI: {-0.93%, 1.91%, 2.26%}

Compared with NNVC-3.0 anchor:

RA: {-11.04%, -22.99%, -22.26%}, LB: {-11.00%, -23.30%, -23.66%}, AI: {-8.37%, -18.59%, -18.84%}

Fig. 2.gives the architecture of the proposed CNN filter, 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. The proposed model takes reconstruction, prediction, slice type, and qp as input, therefore is equal to 4. 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 19 for. In the end, there is a TailBlock mapping the embedded features from backbone to the final output.



*Fig. 2*. *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 | 2.05M |
| Parameter Precision (Bits) | 16 (I) |
| Memory Parameter (MB) | 4.17 MB |
| Multiply Accumulate (kMAC/pixel) | 446K/pixel (frame-level input)  564K/pixel (block-level input) |
| Optional |  |  |
| Total Conv. Layers | 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: | 140h |
| 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: |  |

Slice type is input once per frame.

2 models (1 for luma and chroma each, chroma model is smaller). The luma model is actually having larger number of resblocks (though simpler) than filter set#1, where however the latter has 4 models

It was commented that implementation-wise it is advantageous to have a number of channels being power-of-two (it was commented by another expert that this may however be specific to SADL, PyTorch, TensorFlow, whereas in hardware this might not be the case).

Investigate in EE.

[JVET-AC0197](https://jvet-experts.org/doc_end_user/current_document.php?id=12401) EE1-1.1-related: More refinements on NN based in-loop filter [R. Chang, L. Wang, X. Xu, S. Liu (Tencent)]

This contribution reports more refinements on NN based in-loop filter for single model and dual models. The potential performance for single model based on real-iterative training is reported in this contribution. Besides, a dual-model filter based on new training scripts is proposed to achieve better performance. Based on NNVC-3.0, the test results are shown in order of RA and AI configurations as follows.

Test 1 (single model with real-iterative training similar to EE1-1.1.3):

RA: -10.11% -21.26% -21.14% EncT: 153% DecT: 55477%

AI : -7.70% -16.71% -18.00% EncT: 148% DecT: 31664%

Test 2 (dual models trained based on new training scripts):

RA: -10.26% -18.50% -18.32% EncT: 141% DecT: 45393%

AI : -7.95% -18.74% -19.60% EncT: 141% DecT: 26222%

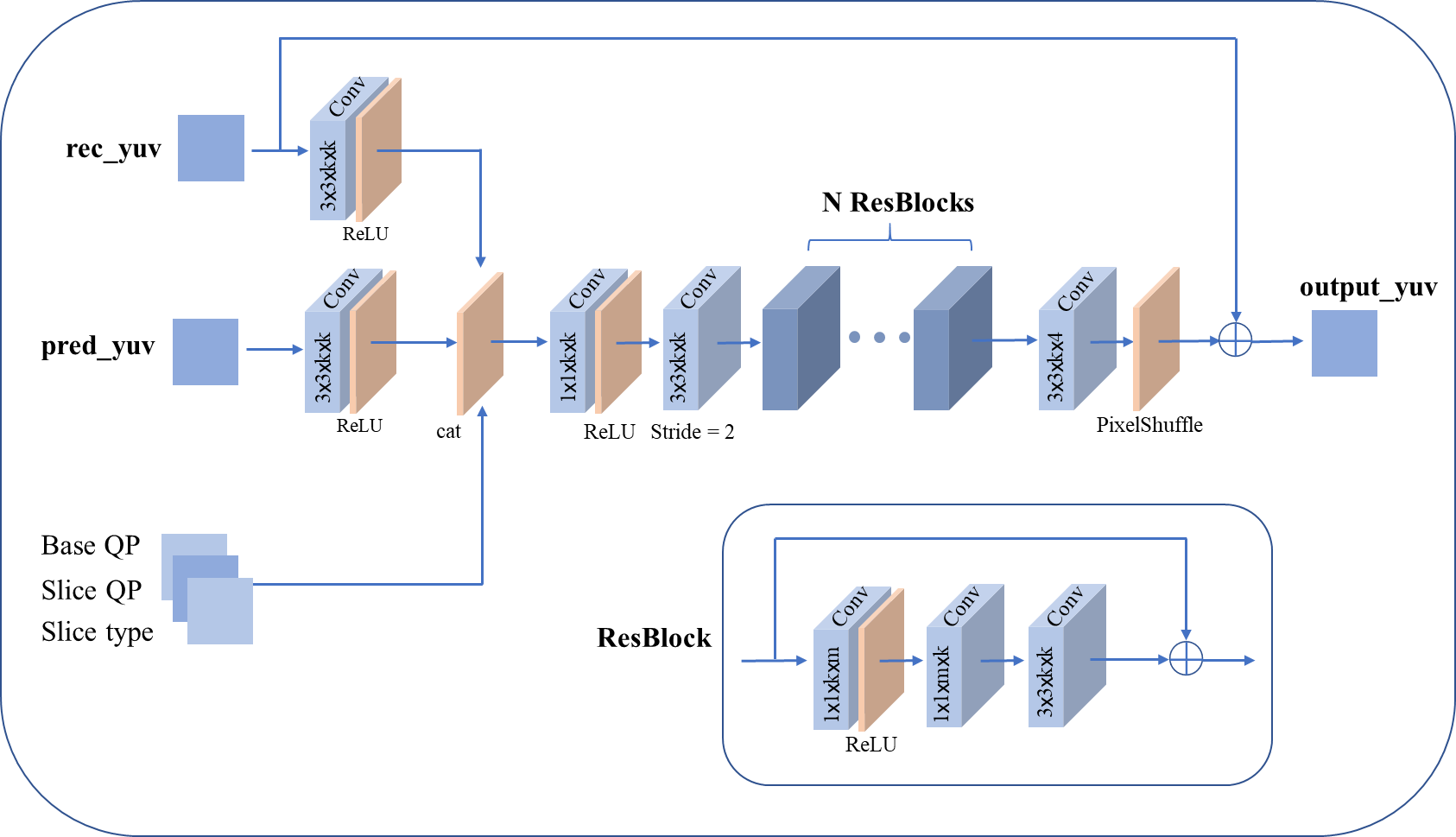


Fig. 1: Network architecture

Architccture identical with filter set#0 (but only 1 model independent of slice type and luma/chroma in test 1; test 2 has still 2 models for luma/chroma but higher number of luma res blocks, lower number for chroma.

|  |  |  |
| --- | --- | --- |
| **Network Information in Training Stage** | | |
| Mandatory | GPU Type | Tesla A100 40GB |
| Framework: | Pytorch v1.9.0 |
| Number of GPUs per Task | 1 |
|  |  |
| Epoch: | ~100 |
| Batch size: | 64 |
| Loss function: | Weighted L1 and L2 |
| Training time (for 1 model): | ~10 days |
| Training data information: | DIV2K, TVD, BVI-DVC |
| Training configurations for generating compressed training data (if different to VTM CTC): |  |
| Optional |  |  |
| Number of iterations | 1200 |
| Patch size | 144x144 |
| Learning rate: | 1e-4 |
| Learning rate update strategy |  |
| Optimizer: | ADAM |
| Preprocessing: | flipped, rotated |
| Mini-batch selection process: | random cropped |
| Training data update strategy: |  |
| Other information: |  |
|  |  |

|  |  |  |  |
| --- | --- | --- | --- |
| **Network Information in Inference Stage** | | | |
|  | Filters | Test 1 | Test 2 |
| Mandatory | HW environment: | | |
| GPU Type | CPU only | |
| Framework: | SADL | |
| Number of GPUs per Task | 0 | |
|  |  | |
| Number of Parameters (Each Model) | 1.9M | 1.66M for luma, 0.96M for chroma |
| Total Number of Parameters (All Models) | 1.9M | 2.62M |
| Parameter Precision (Bits) | 16 | |
| Memory Parameter (MB) | 3.8MB/model, 1 model in all | 3.3MB/model for luma, 1.9MB/model for chroma, 2 models in all |
| Multiply Accumulate (kMAC/pixel) | 485K (w/o block extension)  615K (w/ block extension) | 488K (w/o block extension)  618K (w/ block extension) |
| Calculation Method | block basis | |
| Optional |  |  | |
| Total Conv. Layers | 101 | 89 |
| Total FC Layers | 0 | |
| 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: |  | |
|  |  | |

Investigate in EE.

[JVET-AC0309](https://jvet-experts.org/doc_end_user/current_document.php?id=12514) Crosscheck of JVET-AC0197 test 1 (EE1-1.1-related: More refinements on NN based in-loop filter) [Z. Xie (OPPO)] [late]

[JVET-AC0310](https://jvet-experts.org/doc_end_user/current_document.php?id=12515) EE1-related: Combination of EE1-1.5 and EE1-1.7 [Y. Li, K. Zhang, L. Zhang (Bytedance), D. Liu, J. Ström, M. Damghanian, P. Wennersten, K. Andersson (Ericsson)] [late]

This contribution presents the results of a combination test of EE1-1.5.3 in JVET-AC0089 and EE1-1.7.1 in JVET-AC0177. EE1-1.5.3 proposes to reduce model memory by using a shared model for the in-loop filtering of different types of slices. EE1-1.7.1 proposes to improve the in-loop filtering performance by introducing an additional CNN model. The combined test is to achieve reduced model memory and improved performance compared with NNVC-3.0 filter set #1.

BD-rate changes of {Y, Cb, Cr} of the combined test anchored on EE1-1.5.3 is {-0.73%, 0.06%, 0.06%} for RA.

BD-rate changes of {Y, Cb, Cr} of the combined test anchored on NNVC-3.0 filter set #1 is {-0.86%, -0.18%, -1.35%} for RA.

BD-rate changes of {Y, Cb, Cr} of the combined test anchored on NNVC-3.0 anchor is {-10.26%, -20.88%, -21.43%} for RA.

BD-rate changes of {Y, Cb, Cr} of EE1-1.7.1 anchored on NNVC-3.0 filter set #1 is {-0.67%, 0.03%, 0.05%} for RA.

From results reported above, it is asserted that the method proposed in EE1-1.7.1 can bring similar gains on top of NNVC-3.0 filter set #1 and EE1-1.5.3.

The combination is made by applying 1.7.1 to the three highest temporal layers, 1.5.3. to the remaining layers. This means that different models are used for different temporal layers. In total, 3 models (1 for the higher layers, 2 for the lower ones)

It was commented that increasing the PSNR quality of higher temporal might not be as much give subjective improvement.

It was commented by different experts that the introduction of an additional model is undesirable.

E. Alshina commented that the interesting aspect of 1.7.1 is the introduction of filtering taking additional input from reference frames. To allow experts investigating that technology, it was suggested to include in NNVC software (in the version of combination with 1.5.3), but turn it off by default for comparison with other proposals not using such an approach in the EE.

Decision: Adopt JVET-AC0310, but make 1.5.3 the default configuration, and allow disabling 1.7.1 component. Further, the method shall be changed to use the actual reference frame (not the unfiltered decoded frame)

[JVET-AC0348](https://jvet-experts.org/doc_end_user/current_document.php?id=12553) Crosscheck of JVET-AC0310 (EE1-related: Combination of EE1-1.5 and EE1-1.7) [C. Zhou (vivo)] [late]

[JVET-AC0328](https://jvet-experts.org/doc_end_user/current_document.php?id=12533) EE1-related: Combination of EE1-1.1 and EE1-1.2 [R. Chang, L. Wang, X. Xu, S. Liu (Tencent)] [late]

This contribution reports the combination test results of EE1-1.1.2 in JVET-AC0194 and EE1-1.2.1 in JVET-AC0195. EE1-1.1.2 optimizes the performance by retraining the models base on new training scripts. EE1-1.2.1 proposes a NN filter which is involved in the partitioning decision process. Based on the NNVC-3.0, the combination test results are shown in order of RA and AI configurations as follows.

RA: -9.99% -21.05% -20.50% EncT: 160% DecT: 55750%

AI : -7.95% -16.80% -17.52% EncT: 172% DecT: 31594%

This contribution shows that the RDO concept using the RDO decision network that was optimized fo another EE part is also beneficial in combination with EE1-1.1.2 (which had already been adopted)

Decision: Adopt JVET-AC0328, but disable the RDO network in the encoder by default (unless such an approach is generally enabled for both filter sets #0 and #1)

[JVET-AC0340](https://jvet-experts.org/doc_end_user/current_document.php?id=12545) Crosscheck of JVET-AC0328 (EE1-related: Combination of EE1-1.1 and EE1-1.2) [D. Liu (Ericsson)] [late]

[JVET-AC0342](https://jvet-experts.org/doc_end_user/current_document.php?id=12547) Crosscheck of JVET-AC0328 (EE1-related: Combination of EE1-1.1 and EE1-1.2) [Z. Xie (OPPO)] [late]

### Improvements of NNVC software beyond EE1 (1)

Contributions in this area were discussed in session 19 at 2215–2225 UTC on Tuesday 17 Jan. 2023 (chaired by JRO).

[JVET-AC0065](https://jvet-experts.org/doc_end_user/current_document.php?id=12267) Non-EE1: On flipping of input and output of model in NNVC filter set 0 [Z. Xie, Y. Yu, H. Yu, D. Wang (OPPO)]

In this contribution, flipping operations applied to the inputs and output of the NNVC filter set 0[1] neural network are proposed to further improve the performance. The input of filter set 0 includes reconstructed samples, prediction samples, BaseQP, SliceQP and SliceType. When a flipping operation is applied, only the reconstructed samples and prediction samples are flipped and fed into the network with the other inputs to perform the inference. Then the output of network is flipped back to restore the order of the filtered samples. On top of NNVC-3.0 filter set 0, the performance of the proposed method is reported as below:

Compared with NNVC-3.0 filter set 0:

RA: Y -0.18%, U -0.37%, V -0.44%, EncT 113%, DecT 100%;

LDB: Y -0.57%, U -1.46%, V -2.03%, EncT 114%, DecT 102%.

Flipping is only applied in B slices. Signalled by slice-level flags, depending on which all CTUs in the given slice are flipped horizontally, vertically, or not.

Encoding runtime is increased significantly, and decrease should be investigated. It would be interesting to investigate if there are frequent changes of the selected flipping mode over the sequence, to avoid filtering twice at the encoder to find the best mode.

Invstigate in EE. Should also be tested in combination with RDO decision.

[JVET-AC0303](https://jvet-experts.org/doc_end_user/current_document.php?id=12508) Crosscheck of JVET-AC0065 (Non-EE1: On flipping of input and output of model in NNVC filter set 0) [R. Chang (Tencent)] [late]

### Other aspects of neural network-based video coding (4)

Contributions in this area were discussed in session 19 at 2230–2310 UTC on Tuesday 17 Jan. 2023 (chaired by JRO).

[JVET-AC0090](https://jvet-experts.org/doc_end_user/current_document.php?id=12292) AhG11: Neural Network-based Reference CU Quality Enhancement for Motion Compensation Prediction [Y. Chu, Z. Wang, W. Zhang, S. Li (Hisense)]

This contribution proposes to enhance the Coding Unit (CU) before motion compensation with neural network. The proposed method is mandatorily performed on CU with minimum size of 16×16. Experimental results demonstrate that, compared with VTM-11.0\_nnvc-2.0, the proposed method achieves 0.86% BD-rate reductions for Y component on Class D under the LDP configuration.

Fig. 1 (a) is the process of integrating neural network into conventional motion compensation while Fig. 1 (b) is the process of integrating the neural network into affine motion compensation. In the conventional motion compensation, the VVC standard method finds the reference CU from the current CU according to the MV. Since the fractional interpolation filter will reduce the image size, it is necessary to expand the boundary of the reference CU by 4 pixels and then perform DCTIF to ensure that the CU size is consistent before and after filtering. The residual CU is obtained by subtracting the filtered prediction CU from the original CU, then it is transformed and quantized. Our mothed is designed to input the padded image patch into neural network, and the output image patch is then interpolated and filtered. In the affine motion compensation, the VVC standard method is to split the current CU into several 4 × 4 sub-CUs, and calculate the MV of each sub-CU according to the 4-parameter or 6-parameter affine models. The reference sub-CUs will then be found through MV. By performing padding, fractional interpolation on the reference sub-CUs, the final predicted sub-CUs can be obtained. Finally, the predicted sub-CUs will be combined to gets the predicted CU. In our method, we first get the max region containing all the reference sub-CUs, and then pad the region and input it to the neural network. Then, based on the corresponding position of the output image patch, the padding sub-CUs can be filtered by DCTIF to get the predicted sub-CUs.

**We trained three networks, one for the CU block with size 64×64, one for the CU block with size 64×32, 32×64, 32×32, and the last one for the CU block with size 16×64, 64×16, 32×16, 16×32, 16×16.**



(a) Conventional motion compensation

 (b) Affine motion compensation

Fig. 1 motion compensation in VVC



Fig. 2 The structure of the proposed neural network.

Table 1. Training information.

|  |  |  |
| --- | --- | --- |
| **Training information** | | |
| Mandatory | GPU Type | NVIDIA Geforce RTX 3090 |
| CPU Type | Intel(R) Core(TM) i9-10900X CPU @ 3.70GHz |
| Framework: | PyTorch 1.8.0 |
| Number of GPUs per Task: | 1 |
| Epoch: | 240 |
| Batch size: | 64 |
|  | Total Parameter Number: | 738.43k/model (3 models in total) |
|  | Parameter Precision (Bits): | 32 (F) |
|  | Training data information: | BVI-DVC |
|  | Training configurations for generating compressed training data (if different to VTM CTC): | QP =22 |
| Optional | Patch size | 64x64, 32x32, 16x16 |
| Learning rate: | 1e-4 |
| Optimizer: | ADAM |
| Loss function: | MSE |

|  |  |  |
| --- | --- | --- |
| **Inference information** | | |
| Mandatory | GPU Type | N/A |
| CPU Type | Intel(R) Core(TM) i9-10900X CPU @ 3.70GHz |
| Framework: | PyTorch 1.8.0 |
| Number of GPUs per Task | 0 |
| Total Parameter Number | 738.43k/model (3 models in total) |
| Parameter Precision (Bits) | 32 (F) |
|  | Memory Parameter (KB) | 2611K/model (3 models in total) |
| Optional | Total Conv. Layers | 22 |
| Total FC Layers | 0 |
| Batch size: | 1 block at a time |
| Patch size | Depending on the block (64x64, 32x32, 16x16) |

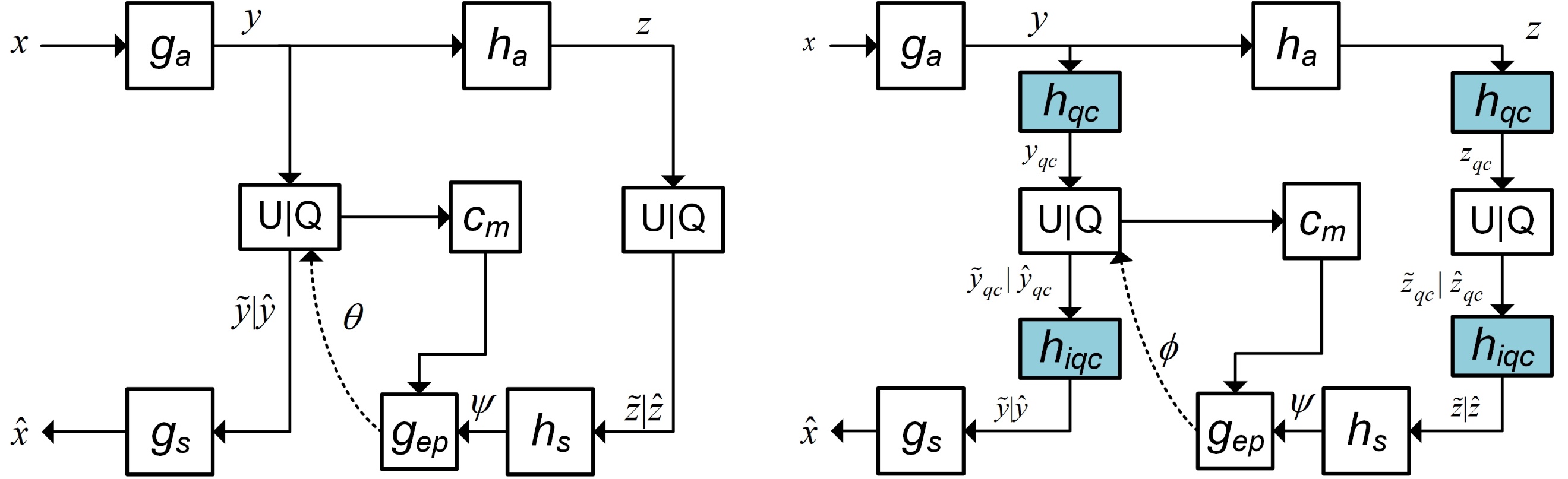
Goal is enhancement of prediction

Further study recommended

* Complexity reduction should be possible
* Would it still have gain when a NN-based loop filter is used?
* Performance on larger resolutions?
* Where is the gain coming from? Correction of affine subblocks?

[JVET-AC0091](https://jvet-experts.org/doc_end_user/current_document.php?id=12293) AhG11: Fourier Series and Laplacian Noise-based Quantization Error Compensation for End-to-End Learning-based Image Compression [S. Jiang, Z. Wang, W. Zhang, S. Li (Hisense)]

Quantization is a core operation in lossy image compression. In the end-to-end learning-based image compression framework, quantization is conducted by a rounding operation during test, while it is replaced by additive uniform noise during training, leading to a mismatched problem between train and test. To address this problem, this contribution presents a quantization error compensation method for the end-to-end learning-based image compression framework. The method uses Fourier series to approximate the periodic changes of the quantization error, and adds Laplacian noise to the quantized latent during test. The proposed method can be flexibly combined with different end-to-end learning-based image compression methods. Experimental results report that higher coding efficiency can be achieved by adding the proposed method with the state-of-the-art methods.

\

(a) Baseline model (b) Proposed model

Fig. 1. Operational diagrams of learning-based compression models. (a) Baseline model [1], (b) Proposed model with quantization error compensation.

Further study recommended: Would it still have an effect when RDOQ was applied? That might violate the assumption about the distribution of the quantization error. This could potentially be investigated with the E2E module in SADL.

[JVET-AC0114](https://jvet-experts.org/doc_end_user/current_document.php?id=12316) AHG11: Deep Reference Frame Generation for Inter Prediction Enhancement [J. Jia, Y. Zhang, H. Zhu, Z. Chen (Wuhan Univ.), Z. Liu, 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-AB0114, this contribution optimizes the network architecture and modifies the TMVP strategy for better performance and less complexity. It is reported that this method can achieve {LDB: -3.18%/-11.89%/-9.21%, RA: -4.32%/-10.71%/-10.07%} bitrate savings for Y/U/V components respectively based on VTM-11.0\_nnvc-2.0.

In the v2 version, we detailed the description of TMVP modification.

This method has been presented in JVET-X0060, JVET-Y0096, JVET-AA0082, and JVET-AB0114 at previous meetings. Different from the previous work, there are several modifications in this contribution:

a) Optimize the network architecture with more efficiency and fewer parameters.

b) Modify the TMVP strategy while using DRF.



Fig. 1: The frame work of the proposed DFR method

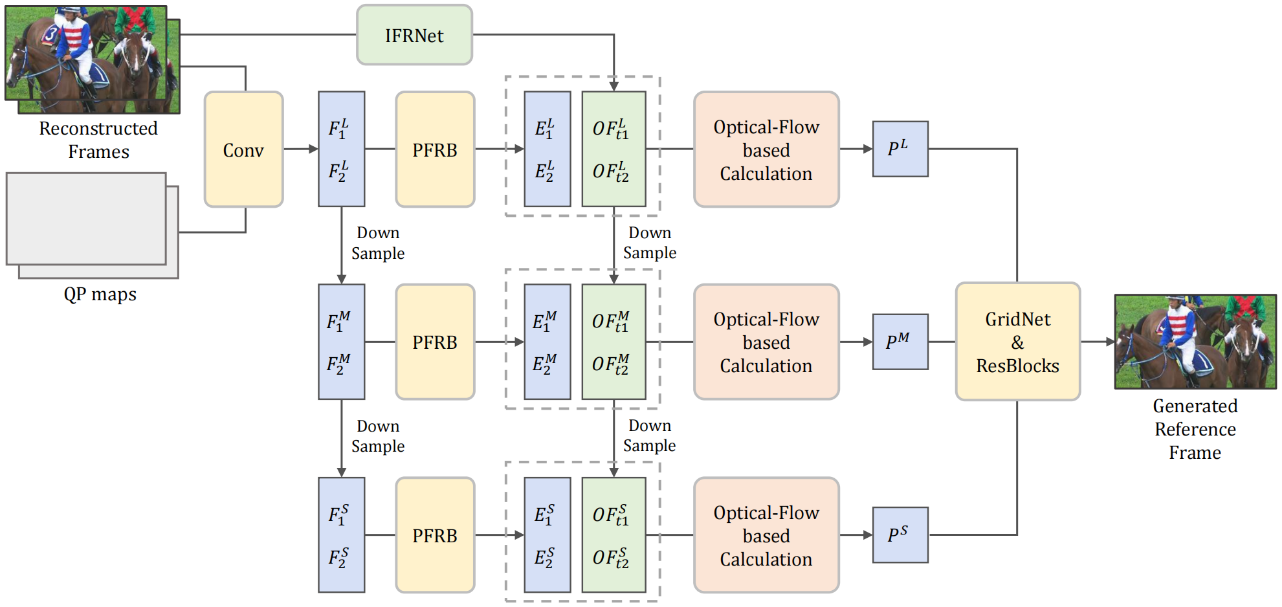


Fig. 2: The architecture of the DRF network

|  |  |  |
| --- | --- | --- |
| **Network Information in Training Stage** | | |
| Mandatory | GPU Type | GTX 1080ti x 8 x 12GB |
| Framework | PyTorch v1.9.0 |
| Number of GPUs per Task | 8 |
|  |  |
| Epoch | 200 |
| Batch size | 3Kx24 (RA) 4Kx18 (LDB) |
| Loss function | L1 |
| Training time (for 1 model) | 168h |
| 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) | 7797K (RA) 11512K (LDB) |
| Total Number of Parameters (All Models) | 19309K |
| Parameter Precision (Bits) | 32 |
| Memory Parameter (MB) | 29.8 (RA) 44.0 (LDB) |
| Multiply Accumulate (kMAC/pixel) | 329 (RA) 398(LDB) |
|  |  |
| Optional | Total Conv. Layers |  |
| Total FC Layers |  |
| Total Memory (MB) |  |
| Batch size |  |
| Patch size | 256x256 |
| Changes to network configuration or weights required to generate rate points |  |
| Peak Memory Usage (Total) |  |
| Peak Memory Usage (per Model) |  |
| Border handling |  |
| Other information |  |
|  |  |

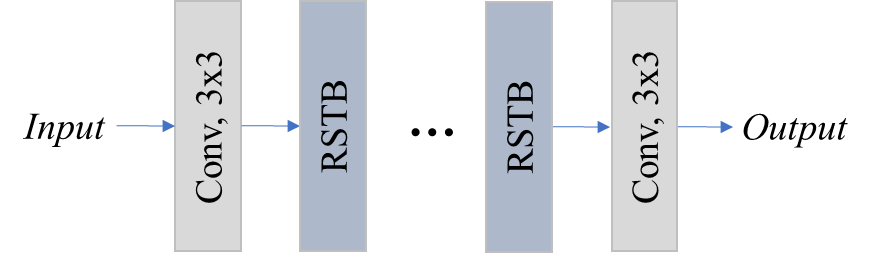
Interesting method, targeting improved inter prediction.

Investigate in EE (new category: NN based inter coding)

[JVET-AC0179](https://jvet-experts.org/doc_end_user/current_document.php?id=12383) AHG11: Swin-Transformer based In-Loop Filter for Natural and Screen Contents [J. Li, K. Zhang, L. Zhang, M. Wang (Bytedance)]

This proposal presents swin-transformer based in-loop filters which are trained for natural contents and screen contents, respectively. Adaptive selection for multiple models is introduced to adapt various contents. Compared with VTM-11.0\_NNVC2, the proposed method reportedly shows on average {7.85%, 20.15%, 20.47%} BD-rate reductions for {Y, Cb, Cr} components, under AI configuration.

The architecture of the proposed swin-transformer based in-loop filters is illustrated in Fig.1, which includes head, backbone, and tail modules. The head module is composed of convolutional layers which are utilized to extract features from inputs. The backbone module involves N sequential-connected Residual Swin Transformer Block (RSTB). The tail module recovers the final reconstruction by using convolutional layers.



*Fig. 1*. *Architecture of the swin-transformer based filter*

|  |  |  |
| --- | --- | --- |
| **Network Information in Inference Stage** | | |
| Mandatory | HW environment: | |
| GPU Type | N/A |
| Framework: | PyTorch v1.6 |
| Number of GPUs per Task | 0 |
|  |  |
| Total Parameter Number | 0.85M/model (4 models) |
| Parameter Precision (Bits) | 32 (F) |
| Memory Parameter (MB) | 8.1 MB/model (4 models) |
| MAC (Giga) | 1052 K/pixel  1188 K/pixel |
| Optional |  |  |
| Total Conv. Layers |  |
| Total FC Layers |  |
| Total Memory (MB) |  |
| Batch size: | 1 |
| Patch size | 128×128 |
| 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: | 50 |
| Batch size: | 64 |
| Training time: | 100h/model |
| Training data information: | DIV2K |
| 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: |  |

Compared to another proposal based on transformer networks (JVET-AC0156), more complex, and 1% less coding gain (in AI); this current proposal does not provide RA.

Proponents indicate that the method is not mature enough for EE.

## AHG12: Enhanced compression beyond VVC capability (92)

### Summary and BoG reports

Contributions in this area were discussed in session 4 at 2355–0130+1 UTC on Wednesday 11 Jan. 2023 (chaired by JRO), in session 8 at 2320–0130+1 UTC on Thursday 12 Jan. 2023 (chaired by JRO), andin sessions 9 and 10 at 1300–1500 UTC and 1520-1640 on Friday 13 Jan. 2023 (chaired by JRO.

[JVET-AC0024](https://jvet-experts.org/doc_end_user/current_document.php?id=12259) EE2: Summary report of exploration experiments on enhanced compression beyond VVC capability [V. Seregin, J. Chen, G. Li, K. Naser, J. Ström, F. Wang, M. Winken, X. Xiu, K. Zhang (EE coordinators)]

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

The software basis for this EE is ECM-7.0, released at <https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tags/ECM-7.0>. ECM-7.0 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-ab-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-ab-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 | Directional planar prediction | LGE  S. Yoo  [JVET-AC0082](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0082-v1.zip) | InterDigital  K. Naser  [JVET-AC0261](https://jvet-experts.org/doc_end_user/current_document.php?id=12466) |
| 1.2 | Improvements on planar horizontal and planar vertical mode | WILUS  K. Kim  [JVET-AC0105](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0105-v1.zip) | InterDigital  K. Naser  [JVET-AC0262](https://jvet-experts.org/doc_end_user/current_document.php?id=12467) |
| 1.3 | Horizontal and vertical planar modes | Alibaba  X. Li  [JVET-AC0084](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0084-v1.zip) | OPPO  F. Wang  [JVET-AC0243](https://jvet-experts.org/doc_end_user/current_document.php?id=12447) |
| 1.4 | Combination of directional planar prediction methods (Test 1.1 + Test 1.2 + Test 1.3) | LGE  S. Yoo  WILUS  K. Kim  Alibaba  X. Li  [JVET-AC0083](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0083-v1.zip) | ETRI  W. Lim |
| 1.6 | Chroma fusion | vivo  C. Zhou  [JVET-AC0119](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0119-v1.zip) | Alibaba  J. Chen |
| 1.7 | Adaptive reference region DIMD | Sharp  Z. Fan  [JVET-AC0045](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0045-v1.zip) | Xiaomi  P. Andrivon |
| 1.8 | Location-dependent DIMD | Nokia  S. Blasi  [JVET-AC0098](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0098-v1.zip) | Xiaomi  P. Andrivon |
| 1.9 | TIMD with directional blending | Nokia  S. Blasi  [JVET-AC0099](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0099-v1.zip) | Xiaomi  P. Andrivon |
| 1.10 | Optimizing the use of available decoded reference samples | InterDigital  T. Dumas  [JVET-AC0094](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0094-v2.zip) | Kwai  H.-J. Jhu  [JVET-AC0168](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0168-v1.zip) |
| 1.11a | Test 1.8 + Test 1.9 | Nokia  S. Blasi  [JVET-AC0100](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0100-v1.zip) | Xiaomi  P. Andrivon |
| 1.11b | Test 1.8 + Test 1.9 + Test 1.10 | Nokia  S. Blasi  InterDigital  T. Dumas  [JVET-AC0101](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0101-v1.zip) | Xiaomi  P. Andrivon |
| 1.11c | Test 1.7 + Test 1.8 + Test 1.9 + Test 1.10 | Sharp  Z. Fan  Nokia  S. Blasi  InterDigital  T. Dumas  [JVET-AC0046](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0046-v1.zip) | Xiaomi  P. Andrivon |
| 1.12a | Gradient and location based CCCM | Nokia  R. G. Youvalari  [JVET-AC0054](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0054-v1.zip) | Qualcomm  Y.-J. Chang  [JVET-AC0281](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0281-v1.zip) |
| 1.12b | Gradient based CCCM | Nokia  R. G. Youvalari  [JVET-AC0054](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0054-v1.zip) | Qualcomm  Y.-J. Chang  [JVET-AC0281](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0281-v1.zip) |
| 1.13 | CCCM using non-downsampled luma samples | Kwai  H.-J. Jhu  [JVET-AC0147](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0147-v1.zip) | Alibaba  X. Li  [JVET-AC0247](https://jvet-experts.org/doc_end_user/current_document.php?id=12452) |
| 1.14 | No luma subsampling for CCCM | Qualcomm  V. Seregin  [JVET-AC0147](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0147-v1.zip) | Alibaba  X. Li  [JVET-AC0247](https://jvet-experts.org/doc_end_user/current_document.php?id=12452) |
| 1.15 | Test 1.13 + Test 1.14 | Kwai  H.-J. Jhu  Qualcomm  V. Seregin | withdrawn |



In the “planar” category, 1.2 provides best tradeoff. Run time is confirmed by cross-checkers, who also report that no additional RDO checks are done; the gain seems to be due to introduction of additional mode.

Decision: Adopt JVET-AC0105 (test 1.2)

1.6 improves the chroma fusion method from ECM (approx. doubling the gain) by introducing switchable fusion models.

Decision: Adopt JVET-AC0119 (test 1.6)

1.7-1.10 relate to DIMD/TIMD, 1.11 are combinations. 1.8 and 1.10 have the best tradeoff in terms of encoder runtime. 1.8 does not increase decoder runtime, while 1.10 does (probably due to additional usage of wide-angular modes). The decoder run time is not critical in the exploration at this point. A combination of 1.8+1.10 was not tested, but from the combinations 1.11a vs. 1.11b it can be concluded that the gains would be almost additive.

Decision: Adopt JVET-AC0098 (test 1.8)

Decision: Adopt JVET-AC0094 (test 1.10)

1.12 extends CCCM by an additional mode, where 1.12a uses the position of a sample, wheeas 1.12b does not. 1.12a clearly performs better and gives good gain (somewhat better than 1.6 which also targets chroma improvement). A combination with 1.6 was no tested, but it was argued by the proponents that conceptually 1.6 and 1.12a do not target the same source of gain (1.6 uses collocated luma samples, whereas 1.12a uses surrounding chroma samples). The contribution JVET-AC0054 also reports results that the gains of 1.12a are additive with 1.13/1.14.

Decision: Adopt JVET-AC0054 (test 1.12a)

1.13 and 1.14 disable downsampling in CCCM to improve its performance in particular for screen content. It is notd that CCCM as tool gave 0.8%/1.1% for class F, and 2.7%/1.5% for TGM (in AI/RA). The feature can be switched at block level. It was further reported that the switching is necessary, otherwise there would be loss in particular for camera captured content. The version with encoder optimization is straightforward and gives additional run time reduction and gain.

Decision: Adopt JVET-AC0147 (test 1.13)

|  |  |  |  |
| --- | --- | --- | --- |
| **2 Inter prediction (add left column with numbers)** | | | |
| Decoder-side affine model refinement | Alibaba  J. Chen  [JVET-AC0144](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0144-v1.zip) | Kwai  X. Xiu |
| Sub-block processing for affine DMVR without subblock downsampling and without linear regression | Qualcomm  H. Huang  Alibaba  J. Chen | withdrawn |
| Sub-block processing for affine DMVR with subblock downsampling and without linear regression | Qualcomm  H. Huang  Alibaba  J. Chen | withdrawn |
| Sub-block processing for affine DMVR | Qualcomm  H. Huang  Alibaba  J. Chen  [JVET-AC0144](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0144-v1.zip) | Xiaomi  M. Blestel |
| Sub-block processing for affine DMVR with subblock downsampling and with linear regression | Qualcomm  H. Huang  Alibaba  J. Chen | withdrawn |
| Control-point motion vector refinement for affine DMVR step 1 | Qualcomm  H. Huang  [JVET-AC0144](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0144-v1.zip) |  |
| Control-point motion vector refinement for affine DMVR all steps | Qualcomm  H. Huang  [JVET-AC0144](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0144-v1.zip) |  |
| Extension of affine DMVR to other modes | Qualcomm  H. Huang | withdrawn |
| Test 2.1 + Test 2.2 | Qualcomm  H. Huang  Alibaba  J. Chen  [JVET-AC0144](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0144-v1.zip) | Xiaomi  M. Blestel  Bytedance  L. Zhao |
| Test 2.1 + Test 2.2 + Test 2.3 step 1 | Qualcomm  H. Huang  Alibaba  J. Chen  [JVET-AC0144](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0144-v1.zip) | Xiaomi  M. Blestel  Bytedance  L. Zhao |
| Test 2.1 + Test 2.2 + Test 2.3 | Alibaba  J. Chen  Qualcomm  H. Huang  [JVET-AC0144](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0144-v1.zip) |  |
| Pixel based affine motion compensation  (minimum subblock size is conditionally set to be 1×1 when OBMC flag is false, otherwise 4×4) | Qualcomm  Z. Zhang  [JVET-AC0158](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0158-v1.zip) | Kwai  X. Xiu |
| Pixel based affine motion compensation  (minimum subblock size is always set to be 1×1) | Qualcomm  Z. Zhang  [JVET-AC0158](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0158-v1.zip) | Kwai  X. Xiu |
| ARMC merge candidate list reordering for AMVP-merge mode for low-delay pictures | Qualcomm  K. Cui  LGE  H. Jang  [JVET-AC0150](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0150-v1.zip) | Xiaomi  F. Le Léannec |
| Adaptive merge candidate list reordering for AMVP-Merge mode at CU level | Qualcomm  K. Cui  [JVET-AC0150](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0150-v1.zip) | Xiaomi  F. Le Léannec |
| Enhanced temporal motion information derivation | Bytedance  L. Zhao  [JVET-AC0185](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0185-v1.zip) | Alibaba  J. Chen |
| Enhanced temporal motion information derivation with SbTMVP template matching using adjacent subblock MVs | Bytedance  L. Zhao  Tencent  L.-F. Chen  [JVET-AC0185](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0185-v1.zip) | Google  X. Li |



Test 2.1-2.4:

Test 2.2 gives straightforward gain (though small, 0.1%) without increased encoding/decoding time.

According to proponents, the standalone benefit of test 2.3a *without*MMVD and adapt. DMVR (as used in combination 2.4b) would be 0.08%, of 0.02% are retained in combination with 2.4a (which is 2.2 and modified 2.1).

According to proponents, the standalone benefit of test 2.1 with search reduction would be 0.07%, of which 0.03% are retained in combination with 2.2 in 2.4a.

2.3b does not provide a reasonable tradeoff.

The gains of the modified 2.1/2.3a proposals are not additive with 2.2.

Test 2.4b requires two refinement steps (one for 2.1 and one for 2.3a).

Test 2.4c did not have complete results, but results were provided later and were showing that gains are somewhat addtive. It was decided die include 2.1 and 2.3 (also in combination) again in next EE

Decision: Adopt JVET-AC0144 (test 2.2)

Test 2.5a: For camera captured content, OBMC is not disabled globally. Pixel based affine is not used when OBMC is used locally. Main intent of this is reduction of run time, but it also gives a small compression benefit. For screen content, OBMC is disabled globally. This is the reason that 2.5a and 2.5b give identical results for screen content.

It is pointed out that in class F two sequences have loss

The pixel based affine mode is an additional affine mode that is signalled. The “normal” subblock affine mode, PROF etc. still can be used.

There are also changes in the OBMC flag signalling both for 2.5a and 2.5b. The motivation behind is that OBMC is always enabled in merge mode. It was suggested by several experts that a solution which does just introduce pixel-based affine mode would be much cleaner. According to proponents, the gain without the modification in OBMC would be 0.16%. It was requested to add such results to the contribution.

There were some issues reported with 2.5a on screen content (software crash). The usage of non-CTC setting of disabling OBMC for screen content was debated.

No changes to deblocking.

Decision: Adopt JVET-AC0158 (test 2.5a) with the following modification: It was agreed to adopt a solution with pixel-based affine as in test 2.5a (not using pixel based affine whenever OBMC is used in the given block), but without the modification of OBMC flag signalling.

The aspect of disabling OBMC globally for screen content was deferred.

Further study (EE) is suggested on the aspect of modifying OBMC flag signalling for the benefit of more frequent usage of pixel based affine.

Test 2.6: The expected benefit (also in compression) was not confirmed in the EE. According to proponents, there is some overlap with modifications that were adopted to ECM7. No action.

Test 2.7: Both variants of improving the temporal MV derivation show gain, slightly higher and slightly less runtime increase for 2.7a.

Decision: Adopt JVET-AC0185 test 2.7a.

|  |  |  |  |
| --- | --- | --- | --- |
| **3 Screen content coding** | | | |
| Direct block vector mode for chroma prediction (using CU level flag in MODE\_INTRA) | Xidian Univ.  J.-Y. Huo  X. Hao  OPPO  M. Li  [JVET-AC0071](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0071-v1.zip) | Alibaba  X. Li  [JVET-AC0245](https://jvet-experts.org/doc_end_user/current_document.php?id=12450) |
| Direct block vector mode for chroma prediction (without signalling) | Xidian Univ.  J.-Y. Huo  X. Hao  OPPO  M. Li  [JVET-AC0071](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0071-v1.zip) | Alibaba  X. Li  [JVET-AC0245](https://jvet-experts.org/doc_end_user/current_document.php?id=12450) |
| Block vector difference sign prediction for IBC blocks | Xidian Univ.  J.-Y. Huo  X. Hao  OPPO  M. Li  [JVET-AC0072](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0072-v1.zip) | Alibaba  X. Li  [JVET-AC0246](https://jvet-experts.org/doc_end_user/current_document.php?id=12451) |
| Block Vector Difference Prediction for IBC blocks: 6 bins (2 sign bins and 4 suffix bins) | Ofinno  A. Filippov  [JVET-AC0104](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0104-v1.zip) | Alibaba  R.-L. Liao  [JVET-AC0210](https://jvet-experts.org/doc_end_user/current_document.php?id=12414)  InterDigital  K. Naser  [JVET-AC0265](https://jvet-experts.org/doc_end_user/current_document.php?id=12470) |
| Block Vector Difference Prediction for IBC blocks: 8 bins (2 sign bins and 6 suffix bins) | Ofinno  A. Filippov  [JVET-AC0104](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0104-v1.zip) | Alibaba  R.-L. Liao  [JVET-AC0210](https://jvet-experts.org/doc_end_user/current_document.php?id=12414) |
| Block Vector Difference Prediction for IBC blocks: 10 bins (2 sign bins and 8 suffix bins) | Ofinno  A. Filippov  [JVET-AC0104](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0104-v1.zip) | Alibaba  R.-L. Liao  [JVET-AC0210](https://jvet-experts.org/doc_end_user/current_document.php?id=12414) |
| Block Vector Difference Prediction for IBC blocks: 12 bins (2 sign bins and 10 suffix bins) | Ofinno  A. Filippov  [JVET-AC0104](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0104-v1.zip) | Alibaba  R.-L. Liao  [JVET-AC0210](https://jvet-experts.org/doc_end_user/current_document.php?id=12414) |
| Test 3.3a + Test 3.1a | Ofinno  A. Filippov  [JVET-AC0104](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0104-v1.zip) | Xiaomi  M. Radosavljevic  [JVET-AC0268](https://jvet-experts.org/doc_end_user/current_document.php?id=12473) |
| Test 3.3a + Test 3.1b | Ofinno  A. Filippov  [JVET-AC0104](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0104-v1.zip) | Xiaomi  M. Radosavljevic  [JVET-AC0268](https://jvet-experts.org/doc_end_user/current_document.php?id=12473) |
| BVP candidates clustering and BVD sign derivation for reconstruction-reordered IBC mode | Ofinno  D. Ruiz  [JVET-AC0060](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0060-v1.zip) | Tencent  H. Zhang  [JVET-AC0233](https://jvet-experts.org/doc_end_user/current_document.php?id=12437) |
| Tests 3.1 + Test 3.2 + Test 3.3 + Test 3.4 | Xidian Univ.  J.-Y. Huo  X. Hao  OPPO  M. Li  Ofinno  A. Filippov  Ofinno  D. Ruiz  [JVET-AC0113](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0113-v1.zip) | Alibaba  R.-L. Liao  [JVET-AC0211](https://jvet-experts.org/doc_end_user/current_document.php?id=12415) |
| Combined IBC and intra prediction | Kwai  C. Ma  Bytedance  Y. Wang  [JVET-AC0112](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0112-v2.zip) | OPPO  L. Zhang  [JVET-AC0244](https://jvet-experts.org/doc_end_user/current_document.php?id=12448) |
| Combined IBC and intra prediction (IBC-CIIP) | Bytedance  Y. Wang  Kwai  C. Ma  [JVET-AC0112](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0112-v2.zip) | OPPO  L. Zhang  [JVET-AC0244](https://jvet-experts.org/doc_end_user/current_document.php?id=12448) |
| 3.6a + IBC with geometry partitioning (IBC-GPM) + IBC with local illumination compensation (IBC-LIC) | Kwai  C. Ma  Bytedance  Y. Wang | Duplicated of 3.6f |
| 3.6b + IBC with geometry partitioning (IBC-GPM) + IBC with local illumination compensation (IBC-LIC) | Bytedance  Y. Wang  Kwai  C. Ma  [JVET-AC0112](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0112-v2.zip) | OPPO  L. Zhang  [JVET-AC0244](https://jvet-experts.org/doc_end_user/current_document.php?id=12448) |
| Combination of 3.6a and 3.6b | Kwai  C. Ma  Bytedance  Y. Wang  [JVET-AC0112](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0112-v2.zip) | Google  X. Li  [JVET-AC0202](https://jvet-experts.org/doc_end_user/current_document.php?id=12406) |
| 3.6e + IBC with geometry partitioning (IBC-GPM) + IBC with local illumination compensation (IBC-LIC) | Kwai  C. Ma  Bytedance  Y. Wang  [JVET-AC0112](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0112-v2.zip) | Google  X. Li  [JVET-AC0202](https://jvet-experts.org/doc_end_user/current_document.php?id=12406) |
| Chroma IBC method as in VTM-5.0 | Tencent  Y. Wang  [JVET-AC0080](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0080-v1.zip) | Google  X. Li  [JVET-AC0133](https://jvet-experts.org/doc_end_user/current_document.php?id=12335) |
| Test 3.1 + Test 3.7 | Tencent  Y. Wang  Xidian Univ.  J.-Y. Huo  X. Hao  OPPO  M. Li  [JVET-AC0125](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0125-v1.zip) | Google  X. Li  [JVET-AC0206](https://jvet-experts.org/doc_end_user/current_document.php?id=12410) |



It was commented that 3.3a is preferable, as the additional complexity of 3.3b..d is not justified by the additional gain.

Comparing 3.7 and 3.8.x with 3.1x indicates that 3.1 has still advantages over the old method from 3.7, even if it takes up some elements from 3.1 (as in 3.8). 3.1 uses template matching, and 3.8b uses it as well. It is concluded that 3.1x has advantages over 3.7 and 3.8.

Comparing 3.1a and 3.1b, the encoder of 3.1a has slightly higher complexity due to an additional RD check. On the other hand, the implicit derivation of 3.1b makes the decoder slightly more complex. 3.1a has slightly better performance (0.05% for class F), which is also retained in the combination tests. From these considerations, 3.1a is asserted to have slight preference over 3.1b.

It was commented that 3.2a and 3.3x could be in competition, where 3.3x tries predicting additional bins beyond the sign. However, 3.3 is not using exactly the same method for sign prediction as 3.2. It was clarified during the discussion that in the combination tests 3.5, actually the method from 3.3 was used. Actually, after getting this information, it can be concluded that the gains from 3.1a, 3.3a, and 3.4 are practically adding up in 3.5a.

Decision: Adopt JVET-AC0113 (version 3.5a)

Tests 3.6x

3.6a and 3.6b are competing methods for IBC/intra combined prediction, a) using adaptive weighting, and b) using fixed weighting for blending. 3.6e is a combination which selects one of the two methods and signal at SPS level which one is used. This is determined based on the hash values of the first picture of the sequence, which are generated for performing fast BV estimation for IBC. This gives an advantage of class TGM (where the standalone methods give both roughly 0.1%, the combination gives 0.44%), whereas in class F the combination is not better than 3.6b. However, during the discussion it was clarified that the method 3.6a in 3.6e is different from the original, as it uses the combined prediction also with AMVP (like 3.6b), whereas the results for 3.6a are reported using it only with merge.

It is generally agreed that switching between two methods of weight generation for blending is undesirable.

Standalone results for 3.6a (also using it for AMVP) are not available, therefore it is impossible to conclude from the results what the benefit of adaptive weights over fixed weights is.

It is further commented that obviously LIC gives good gain, but it was only tested on top of the combination of combined intra/IBC and GPM. Therefore, it is impossible to conclude how much of GPM’s gain is retained in the overall combination (e.g. 3.6d). Comparing 3.6b and 3.6g, the gain of GPM appears to be around 0.25% for class F and 0.7% in AI.

Results for 3.6a also used in AMVP would be necessary to make a decision about the benefit of adaptive weights, otherwise 3.6b would be a candidate for adoption.

Results of LIC on top of 3.6b would be necessary to conclude about the benefit of GPM in the overall combination.

LIC was definitely assessed to be a candidate for adoption.

The requested results were reported in session 27 in an update of JVET-AC0112. Accordingly,

* IBC-GPM provides 0.3% for class F and 0.7% for class TGM on top of IBC-LIC. This shows that the gain of IBC-GPM is additive with LIC.
* Standalone 3.6a (adap. Weights) with AMVP versus 3.6b (fixed weights) with AMVP is 0.4% better for for TGM, and 0.1% worse for class F. In the combination with IBC-LIC and IBC-GPM, the gain for 3.6a in class TGM is only 0.25% .

It was commented that it would be undesirable to implement a high-level switching between two weighting methods. 3.6b is clearly less complex with fixed weighting. It performs better for class F, where typically gains are more difficult to achieve than for class TGM.

Decision: Adopt JVET-AC0112 test 3.6d

Further investigation on adaptive weighting in next EE.

|  |  |  |  |
| --- | --- | --- | --- |
| **4 Transform** | | | |
| 4.1a | Modifications of MTS and LFNST for IntraTMP coded block | WILUS  D. Kim  [JVET-AC0115](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0115-v2.zip) | OPPO  F. Wang  [JVET-AC0240](https://jvet-experts.org/doc_end_user/current_document.php?id=12444) |
| 4.1b | Generation of MPM list using the derived intra prediction mode | InterDigital  K. Naser  [JVET-AC0115](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0115-v2.zip) | OPPO  F. Wang  [JVET-AC0240](https://jvet-experts.org/doc_end_user/current_document.php?id=12444) |
| 4.1c | Test 4.1a + Test 4.1b | WILUS  D. Kim  InterDigital  K. Naser  [JVET-AC0115](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0115-v2.zip) | OPPO  F. Wang  [JVET-AC0240](https://jvet-experts.org/doc_end_user/current_document.php?id=12444) |
| 4.2 | Non-separable primary transform for intra coding | Qualcomm  P. Garus  LGE S. Kim  [JVET-AC0130](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0130-v1.zip) | InterDigital K. Naser  [JVET-AC0264](https://jvet-experts.org/doc_end_user/current_document.php?id=12469) |



Several experts supported 4.1a as a straightforward simplification with slight compression gain and runtime reduction over the original method (which had been adopted by the last meeting). 4.1b does not have benefit according to the combination 4.1c (note there is a typo 4.2b->4.1b in the results table above)

Decision: Adopt JVET-AC0115 test 4.1a.

In Test 4.2a, DCT-II plus LFNST transform combination is replaced with NSPT for the block shapes 4x4, 4x8, 8x4, and 8x8.

In Test 4.2b, DCT-II plus LFNST transform combination is replaced with NSPT for the block shapes 4x4, 4x8, 8x4, 8x8, 4x16, 16x4, 8x16, and 16x8.

It was confirmed that the transforms in 4.2b are a superset of those in 4.2a.

It was asked how much gain comes from the transforms for block sizes 8x16 and 16x8. It was not exactly reported, but the proponent confirms that they contribute to the additional gain, whereas another transform for block size 16x16 that was investigated did not provide significant additional benefit.

(the motivation behind this question is that the necessary storage size and amount of multiplications increases significantly with larger transform kernels, which is not asserted to be a problem in the context of this exploration.)

One expert mentions that the gain improved relative to the last meeting, and also from the first to the second software version that was provided. It was confirmed by proponents that re-training was done.

Run times are based on a SIMD implementation. This makes the decoder faster than before according to the cross-checker. He also reports that the code is very clean.

It was mentioned that the gain is larger for lower-resolution classes such as C and D, whereas it is smaller for class A. The proponents however confirmed that the training included material with high resolution.

Decision: Adopt JVET-AC0130 (test 4.2b)

|  |  |  |  |
| --- | --- | --- | --- |
| **5 In-loop filtering** | | | |
| 5.1a | Using prediction samples for adaptive loop filter | Kwai  C. Ma  [JVET-AC0162](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0162-v1.zip) | Bytedance  W. Yin  [JVET-AC0188](https://jvet-experts.org/doc_end_user/current_document.php?id=12392) |
| 5.1b | Using residual samples for adaptive loop filter | Kwai  C. Ma  [JVET-AC0162](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0162-v1.zip) | Bytedance  W. Yin  [JVET-AC0188](https://jvet-experts.org/doc_end_user/current_document.php?id=12392)  Alibaba  J. Chen |
| 5.2 | ALF with diversified extended taps | Bytedance  W. Yin  [JVET-AC0183](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0183-v1.zip) | Kwai  C. Ma  [JVET-AC0288](https://jvet-experts.org/doc_end_user/current_document.php?id=12392) |
| 5.3a | Test 5.1a + Test 5.2 | Kwai  C. Ma  Bytedance  W. Yin  [JVET-AC0184](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0184-v2.zip) | Google  X. Li  [JVET-AC0214](https://jvet-experts.org/doc_end_user/current_document.php?id=12418) |
| 5.3b | Test 5.1b + Test 5.2 | Kwai  C. Ma  Bytedance  W. Yin  [JVET-AC0184](https://jvet-experts.org/doc_end_user/documents/29_Teleconference/wg11/JVET-AC0184-v2.zip) | Google  X. Li  [JVET-AC0214](https://jvet-experts.org/doc_end_user/current_document.php?id=12418) |



It was reported verbally that 5.2 standalone in AI would give a gain of 0.03%. It was not tested in EE, as currently only a different software implementation is available for intra which would give this gain.

It is understood that decoding time increases due to additional computations, but why is there such a difference between 5.3a/b in RA? It should be irrelevant if residual samples or prediction samples are used.

It was commented that it is difficult to understand what the combination is about, and a precise description about how this is implemented is missing.

5.2 has gain only in RA, no benefit in LB. Combined results 5.3 are not available for LB.

Decision: Adopt JVET-AC0162 (5.1b)

### EE2 contributions: Enhanced compression beyond VVC capability (31)

Contributions in this area were discussed in the context of the EE summary report JVET-AC0024 unless otherwise noted.

For actions decided to be taken, see section 5.3.1, unless otherwise noted.

[JVET-AC0045](https://jvet-experts.org/doc_end_user/current_document.php?id=12230) EE2-1.7: Adaptive Reference Region DIMD [Z. Fan, Y. Yasugi, T. Ikai (Sharp)]

[JVET-AC0226](https://jvet-experts.org/doc_end_user/current_document.php?id=12430) Cross-check of JVET-AC0045 (EE2-1.7: Adaptive Reference Region DIMD) [P. Andrivon (Xiaomi)] [late]

[JVET-AC0046](https://jvet-experts.org/doc_end_user/current_document.php?id=12231) EE2-1.11c: Combination of EE2-1.7, EE2-1.8, EE2-1.9, and EE2-1.10 [Z. Fan, Y. Yasugi, T. Ikai (Sharp), S. Blasi, J. Lainema (Nokia), T. Dumas, K. Reuzé (InterDigital)]

[JVET-AC0054](https://jvet-experts.org/doc_end_user/current_document.php?id=12239) EE2-1.12: Gradient and location based convolutional cross-component model (GL-CCCM) for intra prediction [R. G. Youvalari, P. Astola, J. Lainema (Nokia)]

[JVET-AC0281](https://jvet-experts.org/doc_end_user/current_document.php?id=12486) Crosscheck of JVET-AC0054 (EE2-1.12: Gradient and location based convolutional cross-component model (GL-CCCM) for intra prediction) [Y.-J. Chang (Qualcomm)] [late]

[JVET-AC0060](https://jvet-experts.org/doc_end_user/current_document.php?id=12262) EE2-3.4: BVP candidates clustering and BVD sign derivation for Reconstruction-Reordered IBC mode [D. Ruiz Coll, J.-K. Lee, V. Warudkar (Ofinno)]

[JVET-AC0233](https://jvet-experts.org/doc_end_user/current_document.php?id=12437) Crosscheck of JVET-AC0060 (EE2-3.4: BVP candidates clustering and BVD sign derivation for Reconstruction-Reordered IBC mode) [H. Zhang (Tencent)] [late]

[JVET-AC0071](https://jvet-experts.org/doc_end_user/current_document.php?id=12273) EE2-3.1: Direct block vector mode for chroma prediction [J.-Y. Huo, X. Hao, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.), M. Li, L. Zhang, J. Ren (OPPO)]

[JVET-AC0245](https://jvet-experts.org/doc_end_user/current_document.php?id=12450) Crosscheck of JVET-AC0071 (EE2-3.1: Direct block vector mode for chroma prediction) [X. Li (Alibaba)] [late]

[JVET-AC0072](https://jvet-experts.org/doc_end_user/current_document.php?id=12274) EE2-3.2: Block vector difference sign prediction for IBC blocks [J.-Y. Huo, Z.-Y. Zhang, X. Hao, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.), M. Li, L. Zhang, J. Ren (OPPO)]

[JVET-AC0246](https://jvet-experts.org/doc_end_user/current_document.php?id=12451) Crosscheck of JVET-AC0072 (EE2-3.2: Block vector difference sign prediction for IBC blocks) [X. Li (Alibaba)] [late]

[JVET-AC0080](https://jvet-experts.org/doc_end_user/current_document.php?id=12282) EE2-3.7: Chroma IBC method as in VTM-5.0 [Y. Wang, X. Xu, X. Zhao, R. Chernyak, S. Liu (Tencent)]

[JVET-AC0133](https://jvet-experts.org/doc_end_user/current_document.php?id=12335) Cross-check of JVET-AC0080 EE2-3.7 on Chroma IBC method as in VTM-5.0 [X. Li (Google)]

[JVET-AC0082](https://jvet-experts.org/doc_end_user/current_document.php?id=12284) EE2-1.1: Directional planar prediction [S. Yoo, J. Choi, J. Nam, M. Hong, J. Lim, S. Kim (LGE)]

[JVET-AC0261](https://jvet-experts.org/doc_end_user/current_document.php?id=12466) crosscheck of JVET-AC0082 (EE2-1.1 : Directional planar prediction) [K. Naser (InterDigital)] [late]

[JVET-AC0083](https://jvet-experts.org/doc_end_user/current_document.php?id=12285) EE2-1.4: Combination of directional planar prediction methods [S. Yoo, J. Choi, J. Nam, M. Hong, J. Lim, S. Kim (LGE), K. Kim, D. Kim, J.-H. Son, J. S. Kwak (Wilus), X. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba)]

[JVET-AC0301](https://jvet-experts.org/doc_end_user/current_document.php?id=12506) Crosscheck of JVET-AC0083 (EE2-1.4: Combination of directional planar prediction methods) [W. Lim, S.-C. Lim (ETRI)] [late]

[JVET-AC0084](https://jvet-experts.org/doc_end_user/current_document.php?id=12286) EE2-1.3: Horizontal and vertical planar modes [X. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba)]

[JVET-AC0243](https://jvet-experts.org/doc_end_user/current_document.php?id=12447) Crosscheck of JVET-AC0084 (EE2-1.3: Horizontal and vertical planar modes) [F. Wang (OPPO)] [late]

[JVET-AC0094](https://jvet-experts.org/doc_end_user/current_document.php?id=12296) EE2-1.10: Optimizing the use of reference samples [K. Reuzé, T. Dumas, K. Naser, Y. Chen (InterDigital)]

[JVET-AC0168](https://jvet-experts.org/doc_end_user/current_document.php?id=12372) Crosscheck of JVET-AC0094 (EE2-1.10: Optimizing the use of reference samples) [H.-J. Jhu (Kwai)]

[JVET-AC0098](https://jvet-experts.org/doc_end_user/current_document.php?id=12300) EE2-1.8: Location-dependent DIMD [S. Blasi, J. Lainema (Nokia)]

[JVET-AC0227](https://jvet-experts.org/doc_end_user/current_document.php?id=12431) Cross-check of JVET-AC0098 (EE2-1.8: Location-dependent DIMD) [P. Andrivon (Xiaomi)] [late]

[JVET-AC0099](https://jvet-experts.org/doc_end_user/current_document.php?id=12301) EE2-1.9: TIMD with directional blending [S. Blasi, J. Lainema (Nokia)]

[JVET-AC0228](https://jvet-experts.org/doc_end_user/current_document.php?id=12432) Cross-check of JVET-AC0099 (EE2-1.9: TIMD with directional blending) [P. Andrivon (Xiaomi)] [late]

[JVET-AC0100](https://jvet-experts.org/doc_end_user/current_document.php?id=12302) EE2-1.11a: Combination of Test 1.8 and Test 1.9 [S. Blasi, J. Lainema (Nokia)]

[JVET-AC0229](https://jvet-experts.org/doc_end_user/current_document.php?id=12433) Cross-check of JVET-AC0100 (EE2-1.11a: Combination of Test 1.8 and Test 1.9) [P. Andrivon (Xiaomi)] [late]

[JVET-AC0101](https://jvet-experts.org/doc_end_user/current_document.php?id=12303) EE2-1.11b: Combination of Test 1.8, Test 1.9 and Test 1.10 [S. Blasi, J. Lainema (Nokia), T. Dumas, K. Reuzé (InterDigital)]

[JVET-AC0230](https://jvet-experts.org/doc_end_user/current_document.php?id=12434) Cross-check of JVET-AC0101 (EE2-1.11b: Combination of Test 1.8, Test 1.9 and Test 1.10) [P. Andrivon (Xiaomi)] [late]

[JVET-AC0231](https://jvet-experts.org/doc_end_user/current_document.php?id=12435) Cross-check of JVET-AC0102 (EE2-1.11c: Combination of Test 1.7, Test 1.8, Test 1.9 and Test 1.10) [P. Andrivon (Xiaomi)] [late]

Cross-check of a withdrawn document?

[JVET-AC0104](https://jvet-experts.org/doc_end_user/current_document.php?id=12306) EE2-3.3: Block Vector Difference Prediction for IBC blocks [A. Filippov, V. Rufitskiy (Ofinno)]

[JVET-AC0210](https://jvet-experts.org/doc_end_user/current_document.php?id=12414) Crosscheck of JVET-AC0104 (EE2-3.3: Block Vector Difference Prediction for IBC blocks) [R.-L. Liao (Alibaba)] [late]

[JVET-AC0265](https://jvet-experts.org/doc_end_user/current_document.php?id=12470) crosscheck of JVET-AC0104 (EE2-3.3: Block Vector Difference Prediction for IBC blocks) [K. Naser (InterDigital)] [late]

[JVET-AC0268](https://jvet-experts.org/doc_end_user/current_document.php?id=12473) Crosscheck of JVET-AC0104 (EE2-3.3: Block Vector Difference Prediction for IBC blocks) [M. Radosavljević (Xiaomi)] [late]

[JVET-AC0105](https://jvet-experts.org/doc_end_user/current_document.php?id=12307) EE2-1.2: Improvements on planar horizontal and planar vertical mode [K. Kim, D. Kim, J.-H. Son, J.-S. Kwak (WILUS)]

[JVET-AC0262](https://jvet-experts.org/doc_end_user/current_document.php?id=12467) crosscheck of JVET-AC0105 (EE2-1.2: Improvements on planar horizontal and planar vertical mode) [K. Naser (InterDigital)] [late]

[JVET-AC0112](https://jvet-experts.org/doc_end_user/current_document.php?id=12314) EE2-3.6: IBC-CIIP, IBC-GPM, and IBC-LIC [Y. Wang, K. Zhang, L. Zhang, N. Zhang (Bytedance), C. Ma, X. Xiu, W. Chen, H.-J. Jhu, C.-W. Kuo, N. Yan, X. Wang (Kwai)]

[JVET-AC0202](https://jvet-experts.org/doc_end_user/current_document.php?id=12406) Crosscheck EE2-3.6-ef: IBC-CIIP, IBC-GPM, and IBC-LIC [X. Li (Google)] [late]

[JVET-AC0244](https://jvet-experts.org/doc_end_user/current_document.php?id=12448) Crosscheck of JVET-AC0112 (EE2-3.6 a, b & d: IBC-CIIP, IBC-GPM, and IBC-LIC) [L. Zhang (OPPO)] [late]

[JVET-AC0113](https://jvet-experts.org/doc_end_user/current_document.php?id=12315) EE2-3.5: Combination of test EE2-3.1, test EE2-3.2, test EE2-3.3, and test EE2-3.4 [A. Filippov, V. Rufitskiy, D. Ruiz Coll, J.-K. Lee, V Warudkar (Ofinno), J.-Y. Huo, X. Hao, Z.-Y. Zhang, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.), M. Li, L. Zhang, J. Ren (OPPO)]

[JVET-AC0211](https://jvet-experts.org/doc_end_user/current_document.php?id=12415) Crosscheck of JVET-AC0113 (EE2-3.5: Combination of test EE2-3.1, test EE2-3.2, test EE2-3.3, and test EE2-3.4) [R.-L. Liao (Alibaba)] [late]

[JVET-AC0115](https://jvet-experts.org/doc_end_user/current_document.php?id=12317) EE2-4.1a: modifications of MTS and LFNST for Intra TMP coded blocks [D. Kim, K. Kim, J.-H. Son, J. Kwak (WILUS), K. Naser, T. Poirier, F. Galpin, A. Robert (InterDigital)]

[JVET-AC0240](https://jvet-experts.org/doc_end_user/current_document.php?id=12444) Crosscheck of JVET-AC0115 (EE2-4.1: Experimental results of EE2-4.1a, EE2-4.1b, and EE2-4.1c) [F. Wang (OPPO)] [late]

[JVET-AC0119](https://jvet-experts.org/doc_end_user/current_document.php?id=12321) EE2-1.6: On Chroma Fusion improvement [C. Zhou, Z. Lv, J. Zhang (vivo)]

[JVET-AC0291](https://jvet-experts.org/doc_end_user/current_document.php?id=12496) Crosscheck of JVET-AC0119 (EE2-1.6: On Chroma Fusion improvement) [J. Chen (Alibaba)] [late]

[JVET-AC0125](https://jvet-experts.org/doc_end_user/current_document.php?id=12327) EE2-3.8: Combination of chroma IBC tests [Y. Wang, X. Xu, X. Zhao, R. Chernyak, S. Liu (Tencent), J. Huo, X. Hao, Y. Ma, F. Yang (Xidian Univ.), J. Ren, M. Li (OPPO)]

[JVET-AC0206](https://jvet-experts.org/doc_end_user/current_document.php?id=12410) Crosscheck JVET-AC0125 EE2-3.8 on Combination of chroma IBC tests [X. Li (Google)] [late]

[JVET-AC0130](https://jvet-experts.org/doc_end_user/current_document.php?id=12332) EE2-4.2: Non-Separable Primary Transform for Intra Coding [P. Garus, M. Coban, B. Ray, V. Seregin, M. Karczewicz (Qualcomm), M. Koo, J. Zhao, J. Lim, S. Kim (LGE)]

[JVET-AC0264](https://jvet-experts.org/doc_end_user/current_document.php?id=12469) crosscheck of JVET-AC0130 (EE2-4.2: Non-Separable Primary Transform for Intra Coding) [K. Naser (InterDigital)] [late]

[JVET-AC0144](https://jvet-experts.org/doc_end_user/current_document.php?id=12346) EE2 Test 2.1, 2.2, 2.3 and 2.4: 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)]

[JVET-AC0277](https://jvet-experts.org/doc_end_user/current_document.php?id=12482) Cross-check of JVET-AC0144 tests 2.2 and 2.4b [M. Blestel (Xiaomi)] [late]

[JVET-AC0285](https://jvet-experts.org/doc_end_user/current_document.php?id=12490) Crosscheck of JVET-AC0144 Test 2.4a/b [L. Zhao (Bytedance)] [late]

[JVET-AC0287](https://jvet-experts.org/doc_end_user/current_document.php?id=12492) Cross-check of JVET-AC0144 (EE2-2.1: affine DMVR) [W. Chen (kwai)] [late]

[JVET-AC0147](https://jvet-experts.org/doc_end_user/current_document.php?id=12349) EE2-1.13 and 1.14: 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)]

[JVET-AC0117](https://jvet-experts.org/doc_end_user/current_document.php?id=12319) Cross-check of EE2-1.13: CCCM using non-downsampled luma samples [T. Dumas (InterDigital)]

[JVET-AC0247](https://jvet-experts.org/doc_end_user/current_document.php?id=12452) Crosscheck of JVET-AC0147 (EE2-1.13 and 1.14: CCCM using non-downsampled luma samples) [X. Li (Alibaba)] [late]

[JVET-AC0150](https://jvet-experts.org/doc_end_user/current_document.php?id=12352) EE2-2.6: ARMC merge candidate list reordering for AMVP-merge mode [K. Cui, C. S. Coban, Z. Zhang, H. Huang, V. Seregin, M. Karczewicz (Qualcomm), H. Jang (LGE)]

[JVET-AC0297](https://jvet-experts.org/doc_end_user/current_document.php?id=12502) Cross-check of JVET-AC0150 "EE2-2.6: ARMC merge candidate list reordering for AMVP-merge mode" [F. Le Léannec (Xiaomi)] [late]

[JVET-AC0158](https://jvet-experts.org/doc_end_user/current_document.php?id=12362) EE2-2.5: Pixel based affine motion compensation [Z. Zhang, H. Huang, Y. Zhang, P. Garus, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-AC0280](https://jvet-experts.org/doc_end_user/current_document.php?id=12485) Crosscheck of JVET-AC0158 (EE2-2.5: Pixel based affine motion compensation) [X. Xiu (Kwai)] [late]

[JVET-AC0162](https://jvet-experts.org/doc_end_user/current_document.php?id=12366) EE2-5.1: 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-AC0188](https://jvet-experts.org/doc_end_user/current_document.php?id=12392) Crosscheck of JVET-AC0162 (EE2-5.1: Using prediction or residual samples for adaptive loop filter) [W. Yin (Bytedance)] [late]

[JVET-AC0294](https://jvet-experts.org/doc_end_user/current_document.php?id=12499) Crosscheck of JVET-AC0162 (EE2-5.1: Using prediction samples or residual samples for adaptive loop filter) [J. Chen (Alibaba)] [late]

[JVET-AC0183](https://jvet-experts.org/doc_end_user/current_document.php?id=12387) EE2-5.2: ALF with Diversified Extended Taps [W. Yin, K. Zhang, L. Zhang (Bytedance)]

[JVET-AC0288](https://jvet-experts.org/doc_end_user/current_document.php?id=12493) Crosscheck of JVET-AC0183 (EE2-5.2: ALF with Diversified Extended Taps) [C. Ma (Kwai)] [late]

[JVET-AC0184](https://jvet-experts.org/doc_end_user/current_document.php?id=12388) EE2-5.3: Combination Tests of 5.1 and 5.2 [W. Yin, K. Zhang, L. Zhang (Bytedance), C. Ma, X. Xiu, C.-W. Kuo, W. Chen, H.-J. Jhu, N. Yan, X. Wang (Kwai)] [late]

Initial version rejected as “placeholder”.

[JVET-AC0214](https://jvet-experts.org/doc_end_user/current_document.php?id=12418) Crosscheck of JVET-AB0184 on EE2-5.3: Combination Tests of 5.1 and 5.2 [X. Li (Google)] [miss] [late]

[JVET-AC0224](https://jvet-experts.org/doc_end_user/current_document.php?id=12428) Crosscheck of JVET-AC0184 on Test 5.3b: Combination Tests of 5.1b and 5.2 [L.-F. Chen (Tencent)] [miss] [late]

[JVET-AC0185](https://jvet-experts.org/doc_end_user/current_document.php?id=12389) EE2-2.7: Enhanced temporal motion information derivation [L. Zhao, K. Zhang, L. Zhang (Bytedance), L.-F. Chen, R. Chernyak, X. Zhao, X. Xu, S. Liu (Tencent)]

[JVET-AC0292](https://jvet-experts.org/doc_end_user/current_document.php?id=12497) Crosscheck of JVET-AC0185 (EE2-2.7: Enhanced temporal motion information derivation) [J. Chen (Alibaba)] [late]

[JVET-AC0216](https://jvet-experts.org/doc_end_user/current_document.php?id=12420) Crosscheck of EE2-2.7b: Enhanced temporal motion information derivation [X. Li (Google)] [late]

### EE2 related contributions (9)

Contributions in this area were discussed in session 10 at 1645–1730 UTC on Friday 13 Jan. 2023 (chaired by JRO).

[JVET-AC0093](https://jvet-experts.org/doc_end_user/current_document.php?id=12295) EE2-related: test 2.5a and PROF improvement [F. Galpin, A. Robert, K. Naser (InterDigital)] [late]

In JVET-AB0168, it has been proposed to compute the motion vector of affine block on a per-pixel basis in some conditions. In EE2 test 2.5, this change as well as some modifications related to OBMC conditions are studied. During VVC design, the PROF algorithm was simplified for complexity reason. It is proposed to study the impact of removing some of these simplifications and the interaction with the affine 1x1 proposed in EE2 test 2.5. By removing the PROF simplifications and using the same OBMC changes as in test 2.5, simulation results reportedly achieve Y, U, V BD-rate change with encoder and decoder run time with ECM-7.0 as anchor as follows:

Combination of PROF improvement and EE2 test 2.5a without affine 1x1:

RA -0.04% -0.04% -0.02% for 101% and 101% encT and decT.

LDB: -0.01% -0.11% -0.21% for 101% and 101% encT and decT.

Combination of PROF improvement and EE2 test 2.5a

RA -0.19% -0.45% -0.54% for 103% and 102% encT and decT.

LDB: -0.04% -0.41% -0.34% for 104% and 102% encT and decT.

Initial results are shown with test 2.5a without OBMC changes and without 1x1, which brings the RA gain to 0.02%.

It is noted that the additional gain in the last 2 lines of the abstract is coming from 1x1 affine.

No significant gain – no action.

[JVET-AC0305](https://jvet-experts.org/doc_end_user/current_document.php?id=12510) Crosscheck of JVET-AC0093 (EE2-related: test 2.5a and PROF improvement (Test 2)) [Z. Zhang (Qualcomm)] [late]

[JVET-AC0121](https://jvet-experts.org/doc_end_user/current_document.php?id=12323) EE2-related: on GL-CCCM improvement (test 1.12a) [P. Bordes, K. Naser, F. Galpin, E. François (InterDigital), R. G. Youvalari, P. Astola, J. Lainema (Nokia)]

In EE2 test 1.12a, a gradient and location based convolutional cross-component model (GL-CCCM) for chroma prediction is explored to improve compression efficiency of ECM, where three types of templates for deriving the CC-models are defined: full template, top-only template, and left-only template. It is proposed to extend the template selection to six. Reported PSNR-Y,U,V BD-rate variations of this proposal compared to ECM7.0 anchor are: -0.06%, -0.99%, -1.06% for AI configuration, -0.02%, -0.53%, -0.68% for RA configuration.

Tradeoff of encoder runtime vs. benefit seems slightly worse than for the adopted EE 1.12a.

Investigate in EE.

[JVET-AC0232](https://jvet-experts.org/doc_end_user/current_document.php?id=12436) Cross-check of JVET-AC0121 (EE2-related: on GL-CCCM improvement (test 1.12a)) [P. Andrivon (Xiaomi)] [late]

[JVET-AC0140](https://jvet-experts.org/doc_end_user/current_document.php?id=12342) EE2-related: Additional results for test EE2-3.3 [A. Filippov, V. Rufitskiy (Ofinno)] [late]

(abstract)

Was discussed in context of EE2 summary.

[JVET-AC0269](https://jvet-experts.org/doc_end_user/current_document.php?id=12474) Crosscheck of JVET-AC0140 (EE2-related: Additional results for test EE2-3.3) [M. Radosavljević (Xiaomi)] [late]

[JVET-AC0307](https://jvet-experts.org/doc_end_user/current_document.php?id=12512) Crosscheck of JVET-AC0140 (EE2-related: Additional results for test EE2-3.3 (Test EE2-3.3a combined with Test EE2-3.4)) [Z. Zhang (Qualcomm)] [miss] [late]

[JVET-AC0165](https://jvet-experts.org/doc_end_user/current_document.php?id=12369) EE2-related: Additional results to EE2-3.4 [D. Ruiz Coll, J.-K. Lee, V Warudkar (Ofinno)]

(abstract)

Was discussed in context of EE2 summary.

[JVET-AC0186](https://jvet-experts.org/doc_end_user/current_document.php?id=12390) EE2-related: EE2-2.7 with further encoder optimizations [L. Zhao, K. Zhang, L. Zhang (Bytedance)]

(abstract)

Was discussed in context of EE2 summary.

[JVET-AC0293](https://jvet-experts.org/doc_end_user/current_document.php?id=12498) Crosscheck of JVET-AC0186 (EE2-related: EE2-2.7 with further encoder optimizations) [J. Chen (Alibaba)] [late]

[JVET-AC0205](https://jvet-experts.org/doc_end_user/current_document.php?id=12409) EE2-related: Modifications of MTS and LFNST for IntraTMP [R. G. Youvalari, D. Bugdayci Sansli, J. Lainema (Nokia)] [late]

In ECM-7.0, an IntraTMP coded block uses block size dependent implicit MTS kernels as primary transform and LFNST transform set 0 as the secondary transform. Intra TMP coded blocks are assigned planar mode as the intra prediction mode. In JVET- AB0115 (EE2-4.1), it is proposed to derive the intra prediction mode for IntraTMP block using DIMD method.

This contribution proposes to inherit the intra prediction mode of an IntraTMP block from its reference block or from the neighborhood of its reference block. The proposed method has a simpler approach for selecting the intra prediction mode and it is reported to provide slightly higher coding gain than the EE tests.

The impacts on coding efficiency and runtimes over ECM-7.0 are reportedly {for Y, U, V, EncT, DecT}:

Test A:

AI { -0.05%, -0.01%, -0.07%, 101%, 101%}, RA { -0.02%, -0.08%, -0.06%, 98%, 95%}

Test B (in addition to Test A, the derived intra prediction mode is used in MPM list of neighboring blocks):

AI { -0.06%, -0.03%, -0.01%, 101%, 101%}, RA { -0.xx%, -0.xx%, -0.xx%, 1xx%, 1xx%}

No obvious advantage over the adopted EE2 4.1a.

No action.

[JVET-AC0276](https://jvet-experts.org/doc_end_user/current_document.php?id=12481) EE2-related: Complexity reduction for Decoder Side Control Point Motion Vector Refinement (test 2.3b) [M. Bestel, F. Le Léannec, P. Andrivon, M. Radosavljević (Xiaomi), H. Huang, Y. Zhang, Z. Zhang, C.-C. Chen, V. Seregin, M. Karczewicz (Qualcomm)] [late]

This contribution attempts to optimize the affine DMVR algorithm of test EE2-2.3 that was initially proposed in JVET-AB0178, in terms of trade-off between coding efficiency increase and complexity. It brings several aspects on top of test 2.3b, which are:

* Distortion based early termination
* 6-parameters affine model early termination
* Affine DMVR usage according to CU size and shape
* Adaptation of affine DMVR merge candidate list construction

The proposed modification of top of test EE2-2.3 reportedly leads to the following BD-rates and runtimes relative to ECM-7.0 anchor.

The proposed method reportedly leads to the following BD-rate performances over ECM-7.0 anchors.

RA : XXX% / -XXX% / -XXX% (Y/Cb/Cr). Runtime : XXX% (Enc), XXX% (Dec)

From the partial results shown, about 2/3 of the gain from 2.3b is retained, while the increase in encoder/decoder run time is reduced to about 1/3. From the current results, this would not be a reasonable tradeoff yet.

Further improvements are foreseen by proponents.

Investigate in EE.

It was noted that another proposal that could be further investigated in next EE might be 2.4c for which the results were not complete by the time of current EE review.

[JVET-AC0300](https://jvet-experts.org/doc_end_user/current_document.php?id=12505) Crosscheck of JVET-AC0276 EE2-related: Complexity reduction for Decoder Side Control Point Motion Vector Refinement (test 2.3b) [V. Rufitskiy, A. Filippov (Ofinno)] [late]

[JVET-AC0323](https://jvet-experts.org/doc_end_user/current_document.php?id=12528) EE2-related: Additional Fixed Filter for ALF [W. Yin, K. Zhang, L. Zhang (Bytedance)] [late]

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, additional fixed-filter-output based taps are introduced to provide additional texture information for ALF.

On top of ECM-7.0, simulation results of the proposed are reported as below:

AI: -0.05%, 0.01%, 0.01%, 101% 100%.

RA: -0.05%, 0.02%, -0.01%, 100%, 100%.

LB: -0.05%, -0.17%, -0.13%, 100%, 100%.

Was presented in session 17 (chaired by JRO).

Additional results show that gain is still retained on top of the adopted method EE2-5.1

Several experts expressed interest to investigate in EE.

[JVET-AC0343](https://jvet-experts.org/doc_end_user/current_document.php?id=12548) Crosscheck of JVET-AC0323 (EE2-related: Additional Fixed Filter for ALF) [N. Hu (Qualcomm)] [late]

[JVET-AC0332](https://jvet-experts.org/doc_end_user/current_document.php?id=12537) On worst case decoder complexity of JVET-AC0130 [J. Gan, Y. Yu (OPPO)] [late]

This contribution proposes reducing the number of basis vectors used in the NSPT of JVET-AC0130 for 8x16 and 16x8 block sizes from 40 to 32, which is asserted to maintain the worst-case decoder complexity at the inverse transform stage to 32 multiplications per sample. The reported results are AI -0.43% -0.57% -0.53% enc 103% dec 100%.

The results confirm that the 8x16 and 16x8 transforms are beneficial in terms of gain. It is further shown that the reduction of the number of basis vectors practically has no loss.

It was pointed out that the DCT 256x256, and MST in general, may be more costly, implementation-wise. So, the contribution would not reduce the worst case.

No need to take action.

+crosscheck JVET-AC0347

[JVET-AC0347](https://jvet-experts.org/doc_end_user/current_document.php?id=12552) Crosscheck of JVET-AC0332 on worst case decoder complexity of JVET-AC0130 [X. Li (Google)] [late]

### ECM modifications beyond EE2 (52)

Contributions in this area were discussed in session 11 at 2100–2300 UTC on Friday 13 Jan. 2023 (chaired by JRO), in sessions 15 and 16 at 2100–2300 UTC and at 2320-0130+1 UTC on Monday 16 Jan. 2023 (chaired by JRO), in session 17 at 1435-1500 UTC on Tuesday 17 Jan. 2023 (chaired by JRO), in session 18 at 1525-1730 UTC on Tuesday 17 Jan. 2023 (chaired by Y. Ye), in session 20 at 2330-0145+1 UTC on Tuesday 17 Jan. 2023 (chaired by Y. Ye), and in session 24 at 2320-0125+1 UTC (chaired by JRO).

[JVET-AC0048](https://jvet-experts.org/doc_end_user/current_document.php?id=12233) Non-EE: Modification of TIMD [Y. Yasugi, T. Ikai (Sharp)] [late]

Initial version rejected as “placeholder”.

This contribution proposes a modification of template-based intra derivation (TIMD) mode on top of ECM-7.0 software. This modification allows explicit selection of template from three regions; TL, T and L. Experimental results reportedly show BD-rate -0.09% and -0.03% for AI and RA conditions for test a (no block restrictions), respectively.

Gain in AI is 0.07% with block size restriction, no change in RA.

Encoding time increase 19% in AI, 4% in RA.

With this tradeoff, the proposal is not attractive. Only slight increase in encoding time would be acceptable. Investigate in EE, but would only be considered when encoding is largly decreased.

[JVET-AC0248](https://jvet-experts.org/doc_end_user/current_document.php?id=12453) Crosscheck of JVET-AC0048 (Non-EE: Modification of TIMD) [X. Li (Alibaba)] [late]

[JVET-AC0053](https://jvet-experts.org/doc_end_user/current_document.php?id=12238) AHG12: Simplified linear model solver [J. Lainema, P. Astola, A. Aminlou, R. G. Youvalari (Nokia)]

Some ECM tools, such as CCCM (Convolutional Cross-Component Model) and GLM (Gradient Linear Model), need to solve sets of linear equations when calculating filter coefficients. ECM-7 currently solves those using LDL decomposition. This contribution proposes to replace that with a more straight-forward approach based on Gaussian elimination. It is asserted the proposed approach is beneficial in multiple ways:

* Using row-based matrix operations (enabling SIMD/parallel implementation)
* Smaller amount of arithmetic operations
* Smaller amount of numbers multiplied together (leading to reduced dynamic range)
* Smaller amount of back-substitutions (leading to reduced latency)

In addition to the asserted implementation benefits, the CTC simulations reportedly indicate some tendency towards improved coding efficiency:

AI { 0.00%, -0.15%, 0.03%, 99%, 100%}, RA { -0.01%, -0.10%, -0.02%, 99%, 99%}

LB { 0.00%, -0.38%, -0.42%, 98%, 100%}, LP { 0.00%, -0.07%, -0.34%, 98%, 100%}

The deviation may be due to usage of integer arithmetics, where the reduced number of iterations reduces deviation from optimum.

Changes made in CCCM and GLM.

Not a new tool, rather optimization of implementation of a math function

Code is clean and well understandable according to cross-checker.

Decision: Adopt JVET-AC0053

[JVET-AC0282](https://jvet-experts.org/doc_end_user/current_document.php?id=12487) Crosscheck of JVET-AC0053 (AHG12: Simplified linear model solver) [Y.-J. Chang (Qualcomm) [late]

[JVET-AC0059](https://jvet-experts.org/doc_end_user/current_document.php?id=12261) AHG12: TMP using Reconstruction-Reordered for screen content coding (RR-TMP) [J.-K. Lee, D. Ruiz Coll, V. Warudkar (Ofinno)]

This contribution proposes TMP using Reconstruction-Reordered (RR-TMP) for screen content video coding. When RR-TMP is applied, the prediction signal is generated by matching the flipped L-shaped template causal neighbor of the current block with another reference block for intraTMP in a predefined search area. RR-TMP finds the best-matched flipped template position, flips the corresponding reference block according to the flip type, and uses it as a prediction block of the current block. RR-TMP can be further improved by allowing an adjusted RR-TMP block vector (BV) for intra block copy (IBC). Simulation results based on ECM-7.0 are reported below:

AI: Class F -0.19%, 0.05%, -0.17%, 100%, 101%; Class TGM -0.31%, -0.28%, -0.28%, 100%, 102%

RA: Class F -0.10%, -0.20%, -0.25%, 101%, 100%; Class TGM -0.11%, -0.22%, -0.10%, 101%, 100%

No gain was observed on camera captured content.

It was expressed by other experts that this is straightforward, gain has reasonable tradeoff with complexity and might not have much overlap with proposals adopted to next ECM. Cross-checker also supports it.

Investigate in EE.

[JVET-AC0242](https://jvet-experts.org/doc_end_user/current_document.php?id=12446) Crosscheck of JVET-AC0059 (AHG12: TMP using Reconstruction-Reordered for screen content coding (RR-TMP)) [C.-C. Chen (Qualcomm)] [late]

[JVET-AC0067](https://jvet-experts.org/doc_end_user/current_document.php?id=12269) Non-EE2: Modifications on template-based multiple reference line intra prediction [L. Xu, Y. Yu, H. Yu, D. Wang (OPPO)]

This contribution proposes three modifications on template-based multiple reference line intra prediction. The first modification is to use neighboring 9 position PUs’ intra modes to fill the intra candidate list. The second modification is to use partitioning angles to fill the intra candidate list when neighboring PUs are coded by SGPM or GPM. The third modification is to extend angular precision from 65 to 129 in TMRL mode. The proposed modifications are implemented on top of ECM-7.0 software [1], and reportedly provide Y, U, V-BD rate reduction with encoder and decoder runtime as follows:

AI: -0.08%/-0.10%/-0.10% EncT:98% DecT:100%

RA:

Not intended as simplification, targeting better compression, but reasonable tradeoff in run time (confirmed by cross-checker).

Was supported to be investigated in EE. The individual benefits of the different elements should be reported.

[JVET-AC0313](https://jvet-experts.org/doc_end_user/current_document.php?id=12518) Crosscheck of JVET-AC0067 (Non-EE2: Modifications on template-based multiple reference line intra prediction) [H. Wang (Qualcomm)] [late]

[JVET-AC0068](https://jvet-experts.org/doc_end_user/current_document.php?id=12270) Non-EE2: multi-candidate IntraTMP [F. Wang, L. Zhang, Y. Yu, H. Yu, D. Wang (OPPO)]

This contribution proposes a multi-candidate IntraTMP method. In the proposed method, a candidate list is constructed with the candidate BVs ranked in ascending order of their template matching costs, and the index of selected candidate is signaled in the bit-stream. On top of ECM-7.0, the performance of multi-candidate IntraTMP is reported as follow:

AI {-0.19% Y, -0.22% U, -0.22% V, 98% EncT, 101% DecT}

RA {-0.06% Y, -0.08% U, -0.12% V, 99% EncT, 100% DecT}

This contribution further proposes to enlarge the search range for smaller blocks. On top of ECM-7.0, the performance of multi-candidate IntraTMP with minimum search range 128 is reported as follow:

AI {-0.32% Y, -0.31% U, -0.40% V, 103% EncT, 106% DecT}

RA {-0.12% Y, -0.18% U, -0.21% V, 100% EncT, 100% DecT}

Gain is not uniform over different classes, larger for class E.

Additional results are presented about equivalent search range extension in ECM does only provide small gain.

Search method of IntraTMP is modified to enable the multiple candidates option. It was asked if that is really necessary. It was pointed out that some elements from the IntraTMP search are removed that were included by intent.

Investigate in EE along with JVET-AC0069 and JVET-AC0070. Also investigate options to invoke multiple candidates without significantly changing the current search procedure of IntraTMP. Further, investigate the impact of different number of candidates.

[JVET-AC0259](https://jvet-experts.org/doc_end_user/current_document.php?id=12464) Crosscheck of JVET-AC0068 (Non-EE2: multi-candidate IntraTMP) [Z. Lv (vivo)] [late]

[JVET-AC0069](https://jvet-experts.org/doc_end_user/current_document.php?id=12271) Non-EE2: Intra Template-Matching Prediction Fusion [L. Zhang, F. Wang, Y. Yu, H. Yu, D. Wang (OPPO)]

This contribution proposes an Intra Template-Matching Prediction fusion method that blends multiple matched blocks derived from intra template matching process. The proposed method is implemented on top of ECM-7.0 software, it shows coding performance gains as follows:

Test a (IntraTMP fusion with SAD-based weights):

AI { -0.15% Y, -0.16% U, -0.18% V, 100% EncT, 100% DecT }

RA {xx% Y, xx% U, xx% V, xx% EncT, xx% DecT }

Test b (IntraTMP fusion with fixed weights):

AI { -0.15 % Y, -0.18 % U, -0.29 % V, 100 % EncT, 101 % DecT }

RA {xx% Y, xx% U, xx% V, xx% EncT, xx% DecT }

Test c (IntraTMP fusion with SAD-based weights and enlarged search range):

AI {-0.28% Y, -0.33% U, -0.32% V, 104% EncT, 106% DecT }

RA {-0.10% Y, -0.10% U, -0.16% V, 101% EncT, 100% DecT }

Same modified search procedure as in JVET-AC0068 (which selects one candidate and signals by index). Here, no signalling is used and fusion is done with up to 3 candidates. It is determined by a threshold how many are used.

Investigate in EE – see notes under JVET-AC0068.

[JVET-AC0260](https://jvet-experts.org/doc_end_user/current_document.php?id=12465) Crosscheck of JVET-AC0069 (Non-EE2: Intra Template-Matching Prediction Fusion) [Z. Lv (vivo)] [late]

[JVET-AC0070](https://jvet-experts.org/doc_end_user/current_document.php?id=12272) Non-EE2: combination of JVET-AC0068 and JVET-AC0069 [F. Wang, L. Zhang, Y. Yu, H. Yu, D. Wang (OPPO)]

In JVET-AC0068, a multi-candidate IntraTMP method is proposed. In JVET-AC0069, an IntraTMP fusion method is proposed. This contribution proposes a combination method of multi-candidate IntraTMP and IntraTMP fusion. On top of ECM7.0, the simulation results are reported as follow:

AI {-0.41% Y, -0.45% U, -0.47% V, 104% EncT, 107% DecT}

RA {xx% Y, xx% U, xx% V, xx% EncT, xx% DecT}

It was reported that with enlarged search range gain is up to 0.6%

Investigate in EE – see notes under JVET-AC0068.

[JVET-AC0326](https://jvet-experts.org/doc_end_user/current_document.php?id=12531) Crosscheck of JVET-AC0070 (Non-EE2: combination of JVET-AC0068 and JVET-AC0069) [K. Naser (InterDigital)] [miss] [late]

[JVET-AC0087](https://jvet-experts.org/doc_end_user/current_document.php?id=12289) Non-EE2: Intra TMP with half-pel precision [X. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba)]

In this contribution, half-pel precision is supported in Intra TMP. Specifically, half-pel positions in 8 directions adjacent to the optimal integer-pel position obtained by template matching are further enabled to Intra TMP. The selected precision and directions are signaled in CU level. A 4-tap DCT-IF interpolation filter is used for half-pel interpolation. Two variants of the proposed method with different coding gain and complexity trade-offs are tested. It is reported that on top of ECM-7.0, the overall coding performance impact for {Y, U, V, EncT, DecT} are

Test 1: {-0.12%, -0.15 %, -0.20%, 103%, 100%} AI and {-0.05%, -0.05%, -0.09%, 101%, 100%} RA.

Test 2: {-0.11%, -0.13%, -0.19%, 101%, 100%} AI and {-x.xx%, -x.xx%, -x.xx%, xxx%, xxx%} RA.

Test 2 does not use additional RDO decisions, whereas test 1 does.

Investigate in EE. Performance/complexity tradeoff should be like test 2 or better.

[JVET-AC0318](https://jvet-experts.org/doc_end_user/current_document.php?id=12523) Crosscheck of JVET-AC0087 (Non-EE2: Intra TMP with half-pel precision) [W.-H. Qiao (Xidian Univ.)] [late]

[JVET-AC0095](https://jvet-experts.org/doc_end_user/current_document.php?id=12297) AHG12: Fix related to pixel copy in motion compensation [K. Andersson, R. Yu (Ericsson)]

It was discovered that blocks coded with RPR inter prediction had artifacts in some cases for ECM. The reason was asserted to be the clipping operation in the SIMD code for pixel copy in motion compensation when generating intermediate uni-prediction sample values for bi-predicted blocks. The clipping operation clips those intermediate uni-prediction sample values to be within the sample value range ( 0, ( 1  <<  bitDepth ) – 1) before combining those values into bi-prediction sample values. It is asserted that this clipping operation is unnecessary and incorrect.

The proposed fix is to remove such clipping operation in the SIMD code for pixel copy to keep the accuracy for the intermediate uni-prediction sample values and perform the clipping after combining the uni-prediction sample values into bi-prediction sample values. The proposed fix is asserted to follow exactly the procedure in VTM.

It is proposed to include the fix related to the pixel copy for both normal (i.e., non-RPR) and RPR motion compensation.

The objective impact on RPR CTC is reported to be:

2X (Y/Cb/Cr): -0,09%/-0,25%/-0,20% (PSNR2), -0,09%/-0,18%/-0,20% (PSNR1)

1.5x (Y/Cb/Cr): -0,18% /-0,11%/-0,06% (PSNR2), -0,17%/-0,13%/-0,06% (PSNR1)

The objective impact on normal (non-RPR) CTC is reported to be:

RA (Y/Cb/Cr): 0.00%/0.00%/0.00%

LDB (Y/Cb/Cr): 0.00%/0.00%/0.00%

In v2 a second fix is proposed to make RPR simulations results repeatable.

It is not fully understood why in RPR SIMD and non-SIMD code implementations of pixel copy do not match. A cleaner fix would be to fix the SIMD implementation, rather than replacing it by non-SIMD code in RPR.

Furher study – currently RPR is not used in any experiments, such that an immediate workaround avoiding SIMD is not needed.

A ticket on the problem shall be issued.

For the other problem (padding at picture boundary) the problem is well understood and the fix should be implemented.

Decision(SW/BF): Adopt JVET-AC0095 second aspect (padding at picture boundary)

[JVET-AC0306](https://jvet-experts.org/doc_end_user/current_document.php?id=12511) Crosscheck of JVET-AC0095 (AHG12: Fix related to pixel copy in motion compensation) [Z. Zhang (Qualcomm)] [miss] [late]

[JVET-AC0097](https://jvet-experts.org/doc_end_user/current_document.php?id=12299) Non-EE2: SGPM combined with IntraTMP [C. Fang, S. Peng, D. Jiang, J. Lin, X. Zhang (Dahua)] [late]

Initial version rejected as “placeholder”.

This contribution introduces IntraTMP to the SGPM design. In this method, the two geometric blocks can not only use intra prediction modes, but also use IntraTMP mode. The proposed method was implemented on top of ECM-7.0, it reportedly provides {Y, U, V} BD-rate reduction with encoder and decoder runtimes as follows:

AI: -0.04%, -0.05%, -0.05%, 102% (EncT), 115% (DecT)

Gain is highest in class E

Currently not enough compression benefit to justify inclusion in EE. Further study to prove that combination of TMP with SGPM is useful.

[JVET-AC0103](https://jvet-experts.org/doc_end_user/current_document.php?id=12305) Non-EE2: AMVP mode with subblock-based temporal motion vector prediction [R.-L. Liao, J. Chen, Y. Ye, X. Li (Alibaba)]

This contribution proposes to apply SbTMVP to AMVP mode. Given a CU coded in AMVP with SbTMVP mode, the CU is predicted using a similar way to that of SbTMVP in merge mode except that the motion shift is signaled in the bitstream instead of deriving from left neighboring block. The motion shift comprises a derived MVP and a signaled MVD, and the MVD only has 32 candidates which are composed of 8 motion directions and 4 motion magnitudes. It is reported that on top of ECM-7.0, the overall coding performance impact for {Y, U, V, EncT, DecT} is {-0.10%, -0.08%, -0.12%, 104%, 100%} for RA and {-0.28%, -0.49%, -0.53%, 104%, 101%} for LB.

Increase in encoding time due to additional motion search

Only one AMVP candidate is used.

Investigate in EE. Reduction of encoding time should be targeted.

JVET-AC0213 is also targeting SbTMVP

[JVET-AC0220](https://jvet-experts.org/doc_end_user/current_document.php?id=12424) Crosscheck of JVET-AC0103 (Non-EE2: AMVP mode with subblock-based temporal motion vector prediction) [X. Xiu (Kwai)] [late]

[JVET-AC0107](https://jvet-experts.org/doc_end_user/current_document.php?id=12309) AHG12: Fusion of Intra Template Matching [J. R. Arumugam, A. Natesan, V. Valvaiker, J. N. Shingala (Ittiam), T. Lu, P. Yin, F. Pu, T. Shao, A. Arora, S. McCarthy (Dolby)]

This contribution proposes adaptive intra template fusion to further improve the coding gain of intra template matching. The fusion techniques are based on fusion of TIMD: the two searched templates from intra TM are adaptively fused with weights decided by TM cost.

The method was implemented on top of ECM 7.0. The Y, U, V BDRate, along with encoder and decoder runtime, are reported as follows:

AI: {-0.07%, -0.06%, -0.13%, 100%,100%}

RA: {-0.02%, -0.09%, -0.11%, 10X%,10X%}

Relation with JVET-AC0068..70, but less modifications to current ECM method. Investigate in EE along with those other proposals. One important aspect of that EE would be study of different tradeoff complexity vs. performance in the context of adaptive fusion.

[JVET-AC0321](https://jvet-experts.org/doc_end_user/current_document.php?id=12526) Crosscheck of JVET-AC0107 (AHG12: Fusion of Intra Template Matching) [R.-L. Liao (Alibaba)] [late]

[JVET-AC0336](https://jvet-experts.org/doc_end_user/current_document.php?id=12541) Crosscheck of JVET-AC0107 (AHG12: Fusion of Intra Template Matching) [Y. Wang (Bytedance)] [late]

[JVET-AC0108](https://jvet-experts.org/doc_end_user/current_document.php?id=12310) Non-EE2: Extend search area in Intra Template Matching Prediction (IntraTMP) [J.-Y. Huo, H.-Q. Du, H.-L. Zhang, Z.-Y. Zhang, W.-H. Qiao, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.)]

In ECM7.0, IntraTMP uses the template to search the most similar template in predefined search area and uses its corresponding reference block as the prediction block. IntraTMP divides the search area into four rectangular regions, i.e. R1, R2 , R3, and R4. This contribution proposes to modify two of the four rectangular regions which lets the adjacent neighboring region of the coding block being checked. On top of ECM-7.0, simulation results are reported as follows:

AI(Y/U/V): -0.13%/-0.13%/-0.18%, EncT 100%, DecT 100%

RA(Y/U/V): -0.xx%/-0.xx%/-0.xx%, EncT xx%, DecT xx%

Number of comparisons in search is slightly increased by the modification.

It was asked why the encoding time is largely varying over classes. Proponents express that the measurement was unreliable.

This proposal is asserted to give gain that could be additive with other IntraTMP proposals.

Investigate in EE. JVET-AC0068 also enlarges the search range, should be investigated together.

[JVET-AC0249](https://jvet-experts.org/doc_end_user/current_document.php?id=12454) Crosscheck of JVET-AC0108 (Non-EE2: Extend search area in Intra Template Matching Prediction (IntraTMP)) [X. Li (Alibaba)] [late]

[JVET-AC0109](https://jvet-experts.org/doc_end_user/current_document.php?id=12311) Non-EE2: Intra template matching (Intra TMP) based on linear filter model [J.-Y. Huo, W.-H. Qiao, X. Hao, Z.-Y. Zhang, H.-Q. Du, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.)]

In ECM7.0, Intra Template Matching Prediction (IntraTMP) is applied to camera-captured contents. This contribution proposes to apply a linear filter model for the prediction of Intra TMP (Intra TMP-FLM). The 6-tap linear filter consists of 5 spatial luma samples in the reference block and a bias term. Filter coefficients are derived for each block using the regression based the minimized MSE on samples between the reference template and current template. The test results on top of ECM-7.0 are as follows:

AI(Y/U/V): -0.12%/-0.14%/-0.18%, EncT 101%, DecT 101%

RA(Y/U/V): -0.xx%/-0.xx%/-0.xx%, EncT xx%, DecT xx%

The usage of the mode is signalled by a flag. No RDO is used to determine when to invoke it. The effect may be comparable to the effect of LIC.

This proposal is asserted to give gain that could be additive with other IntraTMP proposals.

Investigate in EE

[JVET-AC0250](https://jvet-experts.org/doc_end_user/current_document.php?id=12455) Crosscheck of JVET-AC0109 (Non-EE2: Intra template matching (Intra TMP) based on linear filter model) [X. Li (Alibaba)] [late]

[JVET-AC0110](https://jvet-experts.org/doc_end_user/current_document.php?id=12312) Non-EE2: A Fusion method of Intra Template Matching Prediction (Intra TMP) [J.-Y. Huo, H.-Q. Du, H.-L. Zhang, W.-H. Qiao, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.)]

In ECM7.0, Intra Template Matching Prediction(IntraTMP) is applied to camera-captured contents. This contribution proposes a IntraTMP Fusion method that blends several reference blocks to derive the final prediction block. These reference blocks are obtained by Block Vectors(BVs) in template matching search process. The test results on top of ECM-7.0 are as follows:

AI(Y/U/V): -0.11%/-0.11%/-0.15%, EncT 94%, DecT 102%

RA(Y/U/V): -0.xx%/-0.xx%/-0.xx%, EncT xx%, DecT xx%

Encoder time is unreliable. Cross check reports around 101%.

The search is not modified

Five candidates are always fused, and weights are derived similar to CCCM. Usage of the fusion mode is signalled.

Investigate in EE together with other fusion proposals. See notes under JVET-AC0107.

[JVET-AC0251](https://jvet-experts.org/doc_end_user/current_document.php?id=12456) Crosscheck of JVET-AC0110 (Non-EE2: A Fusion method of Intra Template Matching Prediction (Intra TMP)) [X. Li (Alibaba)] [late]

[JVET-AC0111](https://jvet-experts.org/doc_end_user/current_document.php?id=12313) Non-EE2: Combination of JVET-AC0108, JVET-AC0109 and JVET-AC0110 for Intra TMP [J.-Y. Huo, W.-H. Qiao, H.-Q. Du, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.), H. Yuan (Shandong Univ.)]

This contribution reports the combination of two methods for Intra TMP.

Method 1 is a combination of the methods proposed in JVET-AC0109 and JVET-AC0110.

The experimental results on top of ECM-7.0 are as below:

AI(Y/U/V): -0.18%/-0.24%/-0.24%, EncT 102%, DecT 102%

RA(Y/U/V): -0.xx%/-0.xx%/-0.xx%, EncT xx%, DecT xx%

Method 2 is a combination of the methods proposed in JVET-AC0108, JVET-AC0109 and JVET-AC0110.

The experimental results on top of ECM-7.0 are as below:

AI(Y/U/V): -0.41%/-0.49%/-0.56%, EncT 99%, DecT 103%

RA(Y/U/V): -0.xx%/-0.xx%/-0.xx%, EncT xx%, DecT xx%

The results indicate that the gains of JVET-AC0108 and JVET-AC0109 are additive, and also additive with fusion adaptation, for which JVET-AC0110 is an example.

[JVET-AC0252](https://jvet-experts.org/doc_end_user/current_document.php?id=12457) Crosscheck of JVET-AC0111 (Non-EE2: Combination of JVET-AC0108, JVET-AC0109 and JVET-AC0110 for Intra TMP) [X. Li (Alibaba)] [late]

[JVET-AC0120](https://jvet-experts.org/doc_end_user/current_document.php?id=12322) Non-EE2: template based intra MPM list construction [C. Zhou, Z. Lv, J. Zhang (vivo)]

This contribution proposes an intra most probable modes optimization method. Firstly, the intra modes derived from the neighbouring blocks and DIMD are sorted with SAD cost using a template, and added to the general MPM list in ascending order of SAD cost after the Planar mode. Then the remaining entries in the general MPM list are the sorted directional modes with added offset. There are two tests, the experimental results are summarized as follows:

On top of ECM-7.0

Test1: This method is disabled when the area of the current CU is greater than 1024.

AI: -0.07% (Y), -0.07% (U), -0.12% (V), 104% (EncT), 103% (DecT)

RA : -0.04% (Y), -0.08% (U), -0.08% (V), 100% (EncT), 100% (DecT)

Test2: To reduce the complexity, this method is disabled when the area of the current CU is greater than 32.

AI: -0.07% (Y), -0.07% (U), -0.13% (V), 103% (EncT), 102% (DecT)

It was asked how large the benefit of changing the signalling for the secondary MPM list would be. It is reported to be around 0.03%.

Several experts supported this proposal to be investigated in EE (along with other proposals targeting MPM modifications)

[JVET-AC0284](https://jvet-experts.org/doc_end_user/current_document.php?id=12489) Crosscheck of JVET-AC0120 (Non-EE2: template based intra MPM list construction) [F. Wang (OPPO)] [late]

[JVET-AC0314](https://jvet-experts.org/doc_end_user/current_document.php?id=12519) Crosscheck of JVET-AC0120 (Non-EE2: template based intra MPM list construction) [H. Wang (Qualcomm)] [late]

[JVET-AC0123](https://jvet-experts.org/doc_end_user/current_document.php?id=12325) MTT maximum depth correction for class B sequences at QP 22 in ECM [G. Laroche, P. Onno (Canon)]

The purpose of this contribution is to correct an error introduced by one ECM-6.0 adoption related to the maximum tree depth of Multiple Type Tree (MTT) for Class A when QP=22 in Random Access (RA). Unfortunately, the software change related to this adoption also unexpectedly affected the Class B sequences by accidentally increasing (from 2 to 3) the maximum Multiple Type Tree (MTT) depth when the QP=22 in RA. The intention of this contribution is to correct this error and to return to the initial setting for Class B sequences at QP=22 for RA to save encoder runtime while maintaining the intended adoption for the ECM 6.0.

The maximum MTT depth was introduced in JVET-AA0098 with the intent of maximum runtime reduction for class A sequences. Its application to class B had not originally been intended.

It was pointed out that the change also would apply to the VTM11-ecm6 anchor, and would also lead to a run time reduction of its encoding. This means the anchor bitstreams at least for class B@QP22 will change.

Decision (BF/SW): Adopt JVET-AC0123

[JVET-AC0235](https://jvet-experts.org/doc_end_user/current_document.php?id=12439) Cross-check of JVET-AC0123 "MTT maximum depth correction for class B sequences at QP 22 in ECM" [F. Le Léannec (Xiaomi)] [late]

[JVET-AC0124](https://jvet-experts.org/doc_end_user/current_document.php?id=12326) Non-EE2: Bi-prediction with block-level only out-of-bound management [F. Le Léannec, M. Blestel, P. Andrivon, M. Radosavljević (Xiaomi)]

This contribution proposes to remove all sample-adaptive processes in the enhanced bi-prediction method of JVET-Z0136 used in ECM-7.0. Instead, a block-level management of out-of-bound (OOB) samples is proposed. No more sample-by-sample OOB criteria evaluation is used, and no more sample-adaptive bi-prediction is applied.

The proposed method reportedly leads to the following BD-rate performances over ECM-7.0 anchors.

RA: 0.00% / -0.01% / -0.08% (Y/Cb/Cr). Runtime: 101% (Enc), 102% (Dec)

LDB : -0.01% / 0.19% / 0.09% (Y/Cb/Cr). Runtime: 97% (Enc), 98% (Dec)

Partial results confirmed by cross-checkers. One cross-checker suggests that the second aspect of the change (not applying BDOF offset in OOB subblocks) is unnecessary, and the current method of OOB handling could be retained.

It was commented that some aspects might be implementation dependent, e.g. the copying of the non-OOB content from the other reference picture.

Replacing the sample-level check by block-level check appears appropriate, but it was questioned if this might enforce superfluous subblock processing in some cases.

Benefit not obvious. No action.

[JVET-AC0219](https://jvet-experts.org/doc_end_user/current_document.php?id=12423) Crosscheck of JVET-AC0124 (Non-EE2: Bi-prediction with block-level only out-of-bound management) [X. Xiu (Kwai)] [late]

[JVET-AC0263](https://jvet-experts.org/doc_end_user/current_document.php?id=12468) Crosscheck of JVET-AC0124 (Non-EE2: Bi-prediction with block-level only out-of-bound management) [F. Galpin (InterDigital)] [late]

[JVET-AC0142](https://jvet-experts.org/doc_end_user/current_document.php?id=12344) Non-EE2: Cross-Component Discrete Mapping Model for Screen Content Coding [B. Vishwanath, K. Zhang, L. Zhang (Bytedance)]

In this contribution, a method of cross-component discrete mapping model (CCDMM) prediction is proposed for screen content coding. Similar to CCLM or CCCM, CCDMM learns a discrete mapping model between luma and chroma from reconstructed neighbors and derive chroma prediction from the co-located luma block with the model. On top of ECM-7.0, simulation results are reported as:

Class F: AI: 0.01%, 0.01%, -0.01%, 101%, 101%; RA: 0.01%, 0.13%, -0.20%, 101%, 100%

Class TGM: AI: -1.22%, -1.44%, -1.80%, 100%, 101%; RA: -0.39%, -0.33%, -0.61%, 101%, 100%

Why no gain on class F? This is untypical for SCC tools seen so far. Could it be the case that there are losses on camera captured or mixed content? In that case, a mechanism should be introduced to enable/disable.

No support by other experts to investigate in EE. Further study recommended.

[JVET-AC0146](https://jvet-experts.org/doc_end_user/current_document.php?id=12348) AHG12: Filtered Template Matching based Intra Prediction (FTMP) [R. G. Youvalari, D. Bugdayci Sansli, P. Astola, J. Lainema (Nokia)]

This contribution presents a template-matching based intra prediction mode for luma blocks, which is called Filtered Template Matching based intra Prediction (FTMP). Prediction is generated by applying a 6-tap filter on the reference block that is obtained from template-matching based search. The filter input consists of 5 spatial luma samples and a bias term. Filter coefficients are derived via regression based MSE minimization on reconstructed samples over the template areas of the current and reference blocks. The mode is signaled with a flag conditioned on intra TMP flag. The impact on coding efficiency and runtimes over ECM-7.0 is reportedly {for Y, U, V, EncT, DecT}:

AI { -0.14%, -0.15%, -0.19%, 105%, 101%}, RA { -0.04%, -0.04%, -0.11%, 101%, 100%}

It was commented that the results in the abstract refer to using 8 candidates; in the presentation, also results with less candidates are shown, but no encoder run times for those – it is likely that the run times are lower. Higher number of candidates also increases decoder complexity.

It was further commented that using more candidates is having a similar effect as fusion methods.

Is the method always applied? Signalled by flag, RDO check at encoder.

A similar approach is in the proposal JVET-AC0109, but different method of deriving the filtering, and different encoder decision. Main difference seems to be number of candidates. The proposal AC0109 seems to have a better tradeoff.

It was requested to communicate offline with the proponents of JVET-AC0109 to identify what the new aspects of this proposal are, and if that would be relevant to investigate in the EE. If possible, that should use the same codebase.

[JVET-AC0148](https://jvet-experts.org/doc_end_user/current_document.php?id=12350) Non-EE2: CCCM using multiple downsampling filters [Y.-J. Chang, V. Seregin, M. Karczewicz (Qualcomm)]

In this contribution, it is proposed to apply multiple downsampling filters to reconstructed luma samples corresponding to the prediction shape of a convolutional cross-component model for chroma intra prediction. The linear combination of these variously downsampled samples separately multiplied by derived coefficients composes a chroma predictor. The experimental results over ECM-7.0 are summarized as follows:

**AI:**  -0.02 % (Y), -1.05 % (U), -0.89 % (V), 104 % (EncT), 100 % (DecT)

**AI-F:** -0.19 % (Y), -1.07 % (U), -0.98 % (V), 103 % (EncT), 101 % (DecT)

**AI-TGM:** -0.75 % (Y), -1.14 % (U), -1.37 % (V), 101 % (EncT), 100 % (DecT)

**RA:** 0.01 % (Y), -0.69 % (U), -0.57 % (V)

**RA-F :** -0.28 % (Y), -1.15 % (U), -0.95 % (V)

**RA-TGM :** -0.2 % (Y), -0.44 % (U), -0.39 % (V)

It is expected (also by cross-checker) that the gains would still be retained in context with new adoptions to ECM. It was suggested to investigate in EE. Reduction of encoder run time to improve the tradeoff with the compression gain.

[JVET-AC0302](https://jvet-experts.org/doc_end_user/current_document.php?id=12507) Crosscheck of JVET-AC0148 (Non-EE2: CCCM using multiple downsampling filters) [J. Lainema (Nokia)] [late]

[JVET-AC0159](https://jvet-experts.org/doc_end_user/current_document.php?id=12363) Non-EE2: IBC MBVD list derivation [Z. Zhang, P. Nikitin, H. Huang, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes to derive the IBC MBVD list from the sample positions along the MBVD directions. The MBVD candidates search starts with checking template SAD costs of offsets added to BVP along each direction with the interval of M-pel. The search process repeats around the K candidates having the lowest TM cost with the half interval. The MBVD list is derived when the interval reaches 1-pel, the K candidates with the lowest TM cost are included into the final list.

The proposed method was implemented on top of ECM-7.0 and the test results are as follows:

Max BVD offset=128-pel, direction=4, start search step=8-pel

AI: -0.30% (Y), -0.17% (U), -0.26% (V), 100% (EncT), 99% (DecT) for class F

-0.64% (Y), -0.58% (U), -0.59% (V), 100% (EncT), 100% (DecT) for class TGM

RA: -0.28% (Y), -0.29% (U), -0.21% (V), 100% (EncT), 100% (DecT) for class F

-0.41% (Y), -0.35% (U), -0.24% (V), 99% (EncT), 100% (DecT) for class TGM

Max BVD offset=256-pel, direction=4, start search step=4-pel

AI: -0.73% (Y), -0.62% (U), -0.72% (V), 105% (EncT), 102% (DecT) for class F

-1.10% (Y), -0.98% (U), -1.03% (V), 105% (EncT), 110% (DecT) for class TGM

RA: -0.45% (Y), -0.47% (U), -0.64% (V), 102% (EncT), 99% (DecT) for class F

-0.84% (Y), -0.79% (U), -0.69% (V), 99% (EncT), 99% (DecT) for class TGM

In case of BVD offset 128, the number of searches is almost same as in current ECM.

It was suggested to investigate in EE. Reduction of encoder run time should be targeted.

[JVET-AC0296](https://jvet-experts.org/doc_end_user/current_document.php?id=12501) Cross-check of JVET-AC0159 (Non-EE2: IBC MBVD list derivation) [K. Andersson (Ericsson)] [late]

[JVET-AC0160](https://jvet-experts.org/doc_end_user/current_document.php?id=12364) Non-EE2: On the condition of OBMC [X. Li (Google)]

In this contribution a method to adaptively disable OBMC is proposed based on prediction samples. 1+% and 2+% average luma coding gains are reported for TGM class in RA and LDB with similar or even less encoding time. When CU level OBMC flag is removed (method2), the encoding complexity is reportedly further reduced. The efficiency impact on camera content is reportedly negligible: average luma BD-rate differences of 0.00%, 0.00%, 0.02% by method1 and 0.00%, 0.00%, 0.04% by method2 are reported for RA, LDB, and LDP, respectively.

Method 1 uses CU-level OBMC flag for explicitly disabling OBMC in case of screen content

Method 2 implicitly determines when to disable OBMC by a boundary condition

Both methods also require decoder changes.

JVET-AC0190 proposes a similar approach.

Study in EE – see further notes under JVET-AC0190.

[JVET-AC0218](https://jvet-experts.org/doc_end_user/current_document.php?id=12422) Crosscheck of JVET-AC0160 (Non-EE2: On the condition of OBMC) [H.-J. Jhu, X. Xiu (Kwai)] [late]

[JVET-AC0161](https://jvet-experts.org/doc_end_user/current_document.php?id=12365) Non-EE2: IBC adaptation for coding of natural content [B. Ray, H. Wang, C.-C. Chen, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution investigates adaptation of IBC tool for the coding of camera captured content.

The experimental results on top of ECM-7.0 are as below:

AI(Y/U/V): -0.90%/-0.91%/-0.99%, EncT 128%, DecT 100%

RA(Y/U/V): -0.34%/-0.35%/-0.40%, EncT 106%, DecT 102%

Further, RR-IBC and TM-IBC is disabled and encoding optimization is done. The results on top of ECM-7.0 are as follows:

AI(Y/U/V): -0.75%/-0.88%/-0.90%, EncT 110%, DecT 101%

Additionally, fractional pel motion compensation for IBC is also investigated on top of previous result. The results on top of ECM-7.0 are as follows:

AI(Y/U/V): -0.92%/-1.08%/-1.14%, EncT 113%, DecT 101%

Aspect 1 is only encoder change (replacing hash-based search) and configuration change

Aspect 2 intoduces fractional displacement which would be normative change

For RA, only used in I slices.

Investigate in EE

[JVET-AC0163](https://jvet-experts.org/doc_end_user/current_document.php?id=12367) Non-EE2: On non-separable primary transform for intra blocks [N. Yan, X. Xiu, W. Chen, H.-J. Jhu, C.-W. Kuo, C. Ma, X. Wang (Kwai)]

In this contribution, non-separable primary transform kernels are proposed for intra coded blocks. Specifically, besides the existing DCT-II + LFNST and adaptive MTS, additional non-separable transforms are applied to intra coded blocks with sizes ranging from 4x4 to 16x16. Furthermore, two simplification methods, namely transform kernel sharing and transform coefficient zeroing-out, are applied to reduce the computational complexity and memory usage of the proposed non-separable transform kernels.

Compared to ECM7.0 anchors, the performance of the proposed scheme is provided as follows.

AI: -0.86%% (Y), -0.68% (U), -0.71 % (V), 103% (EncT), 101% (DecT).

For the simplifications, only partial results are available which however indicate significant loss of performance.

Transforms are different from the ones adopted EE2-4.2 transform. Further, a transform is defined for 16x16.

It was commented that in comparison with the adopted proposal, there is higher gain for class E, while for other classes the gain is lower. The proponents claim that they have additional benefit on top of the adopted proposal.

It was commented that some of the simplifications might not be beneficial.

Benefit over adopted proposal not obvious.

[JVET-AC0253](https://jvet-experts.org/doc_end_user/current_document.php?id=12458) Crosscheck of JVET-AC0163 on Non-EE2: On non-separable primary transform for intra blocks [X. Li (Google)] [late]

[JVET-AC0164](https://jvet-experts.org/doc_end_user/current_document.php?id=12368) Non-EE2: Improvements on local illumination compensation in ECM7.0 [X. Xiu, N. Yan, H.-J. Jhu, W. Chen, C.-W. Kuo, C. Ma, X. Wang (Kwai)]

In this contribution, two aspects are proposed to further improve the coding efficiency of local illumination compensation (LIC) tool in ECM7.0. Firstly, it is proposed to extend the existing LIC design from uni-predicted blocks to bi-predicted blocks. Specifically, by such method, two sets of LIC parameters (i.e., scale and offset) are derived and applied to L0 and L1 prediction blocks separately, which are then combined to form the final bi-prediction of one bi-prediction block. Secondly, it is proposed to enable overlapped block motion compensation (OBMC) to the inter blocks that apply the LIC. Additionally, when one neighboring block of the current block is coded with the LIC, it is proposed to consider the LIC parameters of the neighboring block (besides its MVs) when generating the corresponding prediction samples for the OBMC of the current block. Compared to ECM7.0 anchors, simulation results reportedly show that the proposed modifications provide BD-rate savings of 0.40% and XX for RA and LDB configurations respectively.

Encoding run time increases by 7-8%

Investigate in EE, analyze benefit of the two elements separately, target reduction of encoding time.

[JVET-AC0225](https://jvet-experts.org/doc_end_user/current_document.php?id=12429) Crosscheck of JVET-AC0164 (Non-EE2: on Improvements on local illumination compensation in ECM7.0) [L.-F. Chen (Tencent)] [miss] [late]

[JVET-AC0167](https://jvet-experts.org/doc_end_user/current_document.php?id=12371) AHG12: Using block vector derived from IntraTMP as an IBC candidate for the current block [J.-K. Lee, D. Ruiz Coll, V. Warudkar (Ofinno)]

This contribution proposes to use a block vector (BV) derived from intra template matching prediction (IntraTMP) for the current block as a block vector predictor (BVP) for the AMVP IBC candidate list. After IntraTMP achieves the best reference block, and the TMP-BV is stored in the AMVP IBC list as a new BVP candidate for the current block. Simulation results based on ECM-7.0 are reported below:

AI: Class F: -0.07%, 0.10%, -0.12%, 100%, 97%; Class TGM: -0.04%, -0.00%, -0.02%, 98%, 97%;

RA: Class F: -0.10%, -0.04%, -0.13%, 104%, 102%; Class TGM: 0.03%, -0.10%, 0.01%, 103%, 103%

The proposal replaces the second AMVP BVP predictor with the proposed IntraTMP based BVP. When this candidate is selected, the decoder would need to perform IntraTMP to derive this predictor.

Simulation results are confirmed by cross checker.

It was commented that the gains for the SCC classes are rather low, some loss in class TGM in RA configuration.

It was commented that JVET-AB0061 which was adopted at the last meeting could be related. That contribution uses IntraTMP block vector from the neighboring blocks, whereas this contribution uses IntraTMP block vector from the current block. JVET-AB0061 achieves 0.04%/0.11% for class F/TGM in AI configuration.

JVET-AB0061 had low gains but was considered simple extension as it does not introduce any additional operations. Further study encouraged, especially in the context of whether the proposed method is beneficial for camera captured content as there is an EE2 test on IBC for camera captured content. Could be candidate for EE at next meeting if benefits are convincing.

[JVET-AC0283](https://jvet-experts.org/doc_end_user/current_document.php?id=12488) Crosscheck of JVET-AC0167 (AHG12: Using block vector derived from IntraTMP as an IBC candidate for the current block) [K. Kim (WILUS)] [late]

[JVET-AC0169](https://jvet-experts.org/doc_end_user/current_document.php?id=12373) Non-EE2: Template Matching for RR-IBC [C.-C. Chen, H. Huang, Z. Zhang, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes to extend template matching (TM) to RR-IBC (reconstruction-reordering intra block copy). In ECM-7.0, template matching is applied to TM AMVP and TM MRG modes of IBC only when flipping is disabled for IBC blocks. This contribution enables TM, that ECM-7.0 uses currently, for IBC CU’s with horizontal or vertical flipping type. Firstly, the template blocks are flipped horizontally or vertically before template matching cost is computed, and then the search process of template matching can be reused completely for BV refinement. Experimental results are summarized below.

* AI Class F : (Y) -0.05%, (U) 0.13%, (V) 0.11%, (EncT) 100%, (DecT) 100%
* AI Class TGM : (Y) -0.03%, (U) 0.01%, (V) -0.02%, (EncT) 100%, (DecT) 100%
* RA Class F : (Y) -0.07%, (U) -0.34%, (V) -0.04%, (EncT) 100%, (DecT) 99%
* RA Class TGM: (Y) 0.02%, (U) -0.10%, (V) 0.06%, (EncT) 101%, (DecT) 100%
* LB Class F : (Y) -0.25%, (U) -0.38%, (V) -1.21%, (EncT) 100%, (DecT) 100%
* LB Class TGM: (Y)-0.04%, (U) 0.11%, (V) 0.13%, (EncT) 100%, (DecT) 101%

It was asked if by removing the conditional checks, more template matching would have to be performed. It was commented by the proponent that the proposed method does not increase the amount of template matching that the decoder performs. The asserted inconsistency was introduced when RR-IBC was adopted.

It was commented that the gains for the SCC classes are rather low, some loss in class TGM in RA configuration.

No action at this time.

[JVET-AC0207](https://jvet-experts.org/doc_end_user/current_document.php?id=12411) Crosscheck of JVET-AC0169 (Non-EE2: Template Matching for RR-IBC) [D. Ruiz Coll, J.-K. Lee, V. Warudkar (Ofinno)] [late]

[JVET-AC0170](https://jvet-experts.org/doc_end_user/current_document.php?id=12374) Non-EE2: Fuse intra template matching prediction with intra prediction [Y. Wang, K. Zhang, L. Zhang (Bytedance)]

This contribution presents to fuse intra template matching prediction (Intra TMP) with intra prediction, in which the prediction block is the weighted sum of two prediction signals generated by Intra TMP and intra prediction. On top of ECM-7.0, simulation results of the proposed method are reported as below:

AI: {-0.07%, -0.08%, -0.09%; 102%, 101%};

RA: {-0.01%, 0.00%, -0.03%; 101%, 100%}.

Currently the proposed method uses fixed weights for fusion.

This method is used as an additional mode, signalled conditionally when intraTMP is on. The intra prediction mode is derived using TIMD.

There is a related contribution JVET-AC0201, which modifies CIIP by replacing the inter prediction in CIIP with intraTMP, and using the modified CIIP as an additional mode.

The proponent mentioned that some planned EE2 tests enlarge the search area, which bring additional benefits to this proposed method.

Consider combination with other intraTMP related EE tests. See further notes under JVET-AC0170.

[JVET-AC0333](https://jvet-experts.org/doc_end_user/current_document.php?id=12538) Crosscheck of JVET-AC0170 (Non-EE2: Fuse intra template matching prediction with intra prediction) [F. Pu (Dolby)] [late]

[JVET-AC0171](https://jvet-experts.org/doc_end_user/current_document.php?id=12375) AHG12: ECM-6 Tool Off Tests [X. Li (Google)]

This contribution reports tool off test results of ECM-6. Due to limited computing resources and some issues with the ECM reference software, only a subset of ECM tools were tested and reported. It is proposed to establish an Ad Hoc for the activity of ECM tool off test and analysis.

The testing methodology is described as follows

* Anchor: ECM-6 2b0d8ede
* Configurations: RA and LDP
* Tool off tests with cfg options
  + Some tools are completely off as no separate controls for the improvements in ECM
* VMAF and VMAF NEG (No Enhancement Gain) performance are also reported as supplementary information
* Decoding time was measured with YUV output in all tests

Revised contribution was requested to be uploaded

LDB configuration was not tested due to the long simulation time it requires. How are the VMAF and VMAF NEG values calculated? This was based on a tool provided by Netflix as open source software. Software licensing situation to be clarified.

Comparison against VTM19 anchor was also provided. The proponent reported that VTM19 shows small performance benefit over VTM11-ECM6 in RA configuration but small performance loss in LDB and LDP configurations.

This contribution proposes to

1. Establish an Ad Hoc to perform ECM tool off tests and analysis
2. Fix software issues which block tool off tests
3. Show other quality metric such as VMAF in tool off test results but do NOT consider new metrics when making decisions
4. Discuss whether OBMC should be allowed in low delay P
5. Use EnableTMTools to control the template matching based part of new adoptions for reliable TM test results

Several experts expressed support for establishing such an AHG. The work in this AHG could also be used to fix software issues, which was considered by several experts to be helpful. It was agreed to establish an AHG on tool-off tests (chaired by X. Li, vice chairs: L.-F. Chen, Z. Deng, J. Gan, E. François, H.-J. Jhu, R.-L. Liao, H. Wang).

Offline discussion (organized by X. Li) were requested to be conducted to come up with recommendations on granularity of testing (e.g. grouping of tools), suitable testing configurations and conditions, suitable anchors, etc.

Regarding OBMC in LDP, it was discussed and no action would be taken at this time to disable OBMC in LDP.

Regarding using EnableTMTools to control the template matching based tools, in ECM7, all TM-based inter tools can be turned on/off using EnableTMTools, IntraTMP can be turned on/off, but TM-based IBC cannot be turned on/off separately. To have this TM-based IBC control, an SPS flag needs to exist first. One AHG mandate should be to investigate missing tool controls in ECM, and suggest appropriate fixes to software.

[JVET-AC0354](https://jvet-experts.org/doc_end_user/current_document.php?id=12559) Report on the methodology discussions of ECM tool off tests [X. Li (coordinator)]

This document contains the report of offline discussions on the methodology of ECM tool off tests. The following three topics were discussed

* Tool grouping
* Suitable testing configurations and anchors
* Identify necessary control for TM based tools

In session 24, the concepts contained in this document were presented at 2330 UTC (chaired by JRO). It was proposed to group tools in 4 categories with different levels of asserted “implementation difficulty”.

This was generally asserted being a viable approach, whereas exact criteria about the allocation of a certain tool to a group would need to be worked out. It was suggested by some experts that “group 3” might better be split into two groups, one requiring decoder-side search and another which does not require. Also some differentiation between inter ans intra might be appropriate. For implementability, the degree of option to parallelize operation should also be considered as a criterion.

It was agreed that this is a reasonable starting point, and details to be further worked out in conference call(s).

It was also suggested that for the conf call, it would be good to provide an analysis about how many flags would currently be needed to enable/disable individual tools (or individual features of tools) in ECM.

For enabling/disabling, certain properties might also be controlled via software macro, or by encoder configuration.

[JVET-AC0172](https://jvet-experts.org/doc_end_user/current_document.php?id=12376) Non-EE2: IBC with fractional block vectors [W. Chen, X. Xiu, C. Ma, H.-J. Jhu, C.-W. Kuo, N. Yan, X. Wang (Kwai)] [late]

Initial version rejected as “placeholder”.

This contribution investigates the performance of the IBC mode with fractional BVs. Specifically, besides the existing integer BVs, two fractional BV precisions, including 1/2-pel and 1/4-pel, are introduced for the IBC model. For IBC AMVP, following the existing IBC design, additional adaptive block vector resolution (ABVR) signaling is introduced for the proposed fractional BV precisions. When any fractional BV precision is enabled, the corresponding BVPs and BVDs are in the unit of the selected precision. For IBC merge, fractional BVs are inherited through spatial neighbors.

Compared to the IBC mode in ECM-7.0, the average BD-rate performance of the fractional IBC for camera captured content is summarized as:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | |
|  | **Over ECM7.0 with IBC** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -0.27% | -0.30% | -0.30% | 108% | 99% |
| Class C | -0.17% | -0.15% | -0.17% | 109% | 100% |
| Class E | -0.99% | -0.88% | -1.02% | 108% | 100% |
| **Overall** |  |  |  |  |  |
| Class D | -0.21% | -0.33% | -0.35% | 109% | 100% |
| Class F |  |  |  |  |  |
| Class TGM |  |  |  |  |  |

Proposed method is asserted to be very similar to JVET-AC0161 aspect 2. The gains from JVET-AC0161 aspect 2 is approximately 0.17% for AI on top of JVET-AC0161 aspect 1 (with a much faster encoder that relies on reduced encoder search). The gains in this contribution are estimated to be 0.34% for AI (with a few high rate points filled in with estimated numbers) but the complexity in this contribution is also much higher.

It was commented that fractional BV may also be beneficial for mixed content such as class F. This could be tested in EE.

It was asked if this fractional BV can be applied to TM-based IBC, this is possible, but currently not part of the proposal.

[JVET-AC0173](https://jvet-experts.org/doc_end_user/current_document.php?id=12377) Non-EE2: ALF classification based on residual data [I. Jumakulyyev, N. Hu, Z. Zhang, V. Seregin, M. Karczewicz, H. Huang (Qualcomm)]

In adaptive loop filter (ALF) process of ECM-7.0, a luma coding tree block (CTB) can reference a filter set in an adaptation parameter set. A Laplacian classifier or a band-based classifier is applied to each 2x2 reconstructed luma block in the CTB. In this proposal, a new classifier that applied to the residual samples is proposed. In another aspect, the filter shape of an online derived ALF is modified. On top of ECM-7.0 with common test condition, simulation results of the proposed methods are reported as follows.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | All intra | | | Random access | | | Low delay B | | |
|  | BD-rate Y | EncT | DecT | BD-rate Y | EncT | DecT | BD-rate Y | EncT | DecT |
| Classifier | 0.00% | 100% | 101% | -0.13% | 100% | 101% | -0.30% | 101% | 101% |
| Filter shape | -0.01% | 101% | 102% | -0.06% | 100% | 101% | -0.03% | 102% | 102% |
| Combined | -0.01% | 101% | 100% | -0.18% | 100% | 100% |  |  |  |

It was commented that residual samples are already used in ALF filtering as a result of the adoption of JVET-AC0162 (EE2 test5.1b), so using residual samples in ALF classifier seems to be interesting.

It was commented that the gains from classifier are larger than the gains from filter shape.

[JVET-AC0175](https://jvet-experts.org/doc_end_user/current_document.php?id=12379) Non-EE2: Local-Boosting Cross-Component Prediction [K. Zhang, L. Zhang, Z. Deng (Bytedance)]

This contribution presents methods of local-boosting cross-component prediction (LB-CCP), comprising two aspects:

Aspect #1: Prediction samples of MM-CCLM/MM-CCCM can be filtered with neighbouring samples.

Aspect #2: Neighbouring template costs are calculated to determine the training sample range, as well as the cross-component prediction method used in the chroma fusion mode.

Besides, the binarization of cross-component prediction modes is modified.

On top of ECM-7.0, simulation results of the proposed method are reported as below:

AI: {-0.10%, -0.08%, -0.01%, 101%, 101%}; RA:{ -0.02%, -0.08%, -0.12%, 100%, 99%}

The aspect #1 of this contribution also includes a signaling change for MM-CCLM-L/T/LT and MM-CCCM-L/T/LT modes, which is orthogonal to the main purpose of aspect #1, i.e., filtering of prediction samples. The proponent reported that this part contributed to a minor coding gain (~0.01%). It was suggested to consider this a separate aspect and test it separately. It was asked if aspect #1 could be performed always rather than using a flag to turn on/off. The proponent reported that would result in some coding loss.

Investigate in EE. Separate tests of individual aspects should be performed.

[JVET-AC0176](https://jvet-experts.org/doc_end_user/current_document.php?id=12380) Non-EE2: Non-Local Cross-Component Prediction [K. Zhang, L. Zhang, Z. Deng (Bytedance)]

This contribution presents methods of non-local cross-component prediction.

Method #1: Non-adjacent cross-component prediction (NA-CCP) is proposed to derive CCCM models with samples in regions non-adjacent to the current block.

Method #2: History-based cross-component prediction (H-CCP) is proposed to inherit CCCM models from stored models of CCP-coded blocks in history.

On top of ECM-7.0, simulation results of the proposed method are reported as below:

Method #1 : AI: { -0.02%, -0.90%, -0.90%, 101%, 95%}; RA:{ -0.02%, -0.65%, -0.80%, 98%, 96%}

Method #2 : AI: { -0.03%, -0.91%, -0.94%, 107%, 96%};

It was commented that the non-adjacent cross-component prediction method does not use neighboring samples beyond what is already being used by other coding tools (e.g. IBC) in ECM7.

The runtime reported by the proponent is not accurate, the proposed method is not expected to reduce complexity.

It was asked if the method is also applicable to CCLM, and the proponent suggested that it is applicable to CCLM as well, although that was not part of the proposal. This could be further studied.

The two methods may be combined, which may result in additional gain. The combination of these two methods is suggested to be tested in EE as well.

JVET-AC0315 proposes a method for cross-component prediction. See further notes under JVET-AC0315.

Investigate in EE, including encoding time reduction and possible combination with JVET-AC0315.

[JVET-AC0182](https://jvet-experts.org/doc_end_user/current_document.php?id=12386) Non-EE2: High-Precision MV Refinement for BDOF [M. Salehifar, Y. He, K. Zhang, L. Zhang (Bytedance)]

High precision MV refinement is proposed for BDOF, based on the accurate BDOF parameter calculation.

In test #1, BDOF MV refinement parameters are more accurately calculated, with additional weights.

In test #2, on top of the test #1, BDOF MV refinement subblock size is reduced from 8×8 to 4×4.

Simulation results on top of the ECM-7.0 are reported as:

Test #1: RA: -0.14%, -0.17%, -0.11%, 100%, 101%

Test #2: RA: -0.25%, -0.33%, -0.28%, 106%, 115%

It was asked why subblock size reduction (test #2) could result in such high encoding time increae. This may be because the current code is less optimized, and further code optimization could reduce the runtime.

Investigate in EE.

[JVET-AC0345](https://jvet-experts.org/doc_end_user/current_document.php?id=12550) Crosscheck of JVET-AC0182 (Non-EE2: High-Precision MV Refinement for BDOF) [J. Zhao (LGE)] [late]

[JVET-AC0187](https://jvet-experts.org/doc_end_user/current_document.php?id=12391) Non-EE2: Template matching-based subblock motion refinement [L. Zhao, K. Zhang, Z. Deng, L. Zhang (Bytedance)] [late]

Initial version rejected as “placeholder”.

This contribution presents methods to refine subblock motion based on template matching (TM). Firstly, TM-based control points refinement is proposed for uni-directional affine candidates. Secondly, the motion shift to locate SbTMVP is also refined by TM. It is reported that on top of ECM-7.0, simulation results of the proposed method are as below:

RA: -0.09 % (Y), -0.12 % (U), -0.14 % (V), 107% (EncT), 102% (DecT)

LB: -0.13 % (Y), -0.12 % (U), -0.33 % (V), 111% (EncT), 108% (DecT)

It was commented that the runtime should be reduced for the performance vs. complexity tradeoff to be attractive. Current runtime is too high, which is due to two factors: additional template searches, and code being unoptimized.

Investigate in EE, with runtime reduction expected.

[JVET-AC0330](https://jvet-experts.org/doc_end_user/current_document.php?id=12535) Crosscheck of JVET-AC0187 (Non-EE2: Template matching-based subblock motion refinement) [R.-L. Liao (Alibaba)] [late]

[JVET-AC0189](https://jvet-experts.org/doc_end_user/current_document.php?id=12393) Non-EE2: SGPM without blending [Z. Deng, L. Zhang, K. Zhang, L. Zhao (Bytedance)]

In ECM-7.0, spatial geometric partitioning mode (SGPM) uses a pre-defined blending width based on the block size, without encoder selection or signalling. The blending width could be 16 samples on each side of the partitioning line, no matter a block is of camera-captured content or screen-content. In this contribution, it is proposed to allow SGPM without blending for SCC.

Simulation results reportedly show BD-rate reduction for Y, U, V components as follows.

AI configuration:

Class F: -0.25%, -0.14%, -0.18%; 100% enc time, 98% dec time

Class TGM: -0.38%, -0.31%, -0.35%; 99% enc time, 100% dec time

RA configuration:

Class F: -0.24%, -0.34%, -0.34%; 100% enc time, 100% dec time

Class TGM: -0.12%, -0.26%, -0.19%; 100% enc time, 100% dec time

Updated presentation deck to be uploaded.

SGPM tool-off results in ECM7 for SCC classes:

Class F: 0.15%/0.05% for AI/RA

Class TGM: 0.15%/0.08% for AI/RA

For natural sequences, it is reported by the proponent that SGPM tool-off shows 0.24%/0.12% for AI/RA.

It was commented that some blending will still occur when the proposed flag is set to 1, if the SGPM block has partition angle that is not completely horizontal or vertical, but the blending area will be much narrower than what is currently used.

It was commented by several experts that excessive blending for SCC content could be harmful, and this proposal solves that problem. Several experts, including the original proponent of SGPM, supported adopting this change for SCC content.

It was commented that the proposal uses PPS flag rather than SPS flag because it could allow a finer granularity control of the blending process (e.g. in the case when some natural pictures are followed by screen content pictures).

The proponent further suggested to set this flag to 1 in the case of SCC CTC.

It was suggested that further encoder optimization could use this proposed method to bring gains to some natural content or mixed content by switching the flag picture by picture.

Decision: Adopt JVET-AC0189.

Decision (CTC): set the PPS flag to 1 for class F and class TGM.

[JVET-AC0295](https://jvet-experts.org/doc_end_user/current_document.php?id=12500) Crosscheck of JVET-AC0189(Non-EE2: SGPM without blending) [J. Chen (Alibaba)] [late]

[JVET-AC0190](https://jvet-experts.org/doc_end_user/current_document.php?id=12394) Non-EE2: On OBMC [Z. Deng, K. Zhang, L. Zhang (Bytedance)]

In ECM, OBMC is always used on inter merge coded blocks if the CU is not LIC coded and W×H >=32, without encoder selection. However, OBMC may not be efficient on SCC. In this contribution, three methods are proposed to harmonize OBMC and SCC.

1. Method #1 (normative change):

OBMC is adaptively enabled or disabled on inter merge coded CUs, based on the block characteristics.

1. Method #2 (configuration change):

OBMC is disabled at sequence level on screen contents sequences.

1. Method #3 (encoder change):

At encode, the inter AMVP mode without OBMC is preferred by shrinking its RD cost.

Simulation results based on ECM-7.0 are reported as below:

Method #1:

RA configuration: Class F 0.00%, 100%, 98%; Class TGM -1.29%, 100%, 102%

LB configuration: Class F -0.02%, 102%, 103%; Class TGM -2.09%, 100%, 102%

Method #2:

RA configuration: Class F 0.45%, 98%, 87%; Class TGM -1.32%, 97%, 97%

LB configuration: Class F 0.84%, 98%, 91%; Class TGM -2.11%, 97%, 96%

Method #3:

RA configuration: Class F -0.01%, 100%, 100%; Class TGM -0.40%, 100%, 99%

LB configuration: Class F -0.05%, 100%, 100%; Class TGM -0.70%, 100%, 99%

Method 1 is similar to JVET-AC0160, but here decision is made at CU level rather than boundary level.

Investigate in EE together with JVET-AC0160. In method 1 of 0190 and both methods of 0160, it should also be investigated to use reference blocks rather than prediction blocks for the OBMC on/off decision (this would reduce dependencies due to sequential processing and allow parallelization)

[JVET-AC0337](https://jvet-experts.org/doc_end_user/current_document.php?id=12542) Crosscheck of JVET-AC0190 Method 3 [Z. Lv (vivo)] [late]

[JVET-AC0191](https://jvet-experts.org/doc_end_user/current_document.php?id=12395) Non-EE2: Modification to MPM list derivation [H. Wang, Y.-J. Chang, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution investigates a unified method to derive MPM lists for the regular intra and template-based multiple reference line intra (TMRL) modes coding. In method 1, an MPM derivation process is developed using the existing logic in the MPM generation of regular intra modes and TMRL intra modes. In method 2, on top of method 1, an intra mode is derived for the MIP/intraTMP coded CUs for future MPM list generation, and a re-ordering process is further applied to the MPM list.

The experimental results on top of ECM-7.0 are as follows:

Method 1

AI(Y/U/V): -0.03%/0.00%/-0.04%, EncT 100%, DecT 100%

RA(Y/U/V): -0.01%/-0.07%/-0.15%, EncT 100%, DecT 100%

Method 2

AI(Y/U/V): -0.06%/-0.06%/-0.11%, EncT 102%, DecT 100%

It was suggested to also investigate other options of mode derivation in method 2.

It was pointed out that a similar approach of mode derivation had been used in EE2-4.1b, but no gain was found.

Several experts supported this proposal to be investigated in EE (along with other proposals targeting MPM modifications)

[JVET-AC0258](https://jvet-experts.org/doc_end_user/current_document.php?id=12463) Crosscheck of JVET-AC0191 (Non-EE2: Modification to MPM list derivation) [Z. Lv (vivo)] [late]

[JVET-AC0192](https://jvet-experts.org/doc_end_user/current_document.php?id=12396) Non-EE2: Temporal block vector prediction [N. Zhang, K. Zhang, L. Zhang (Bytedance)]

In ECM, temporal block vector (BV) prediction is not utilized for either IBC merge or IBC AMVP mode. To further improve the efficiency of BV prediction, temporal BV prediction is proposed to mimic TMVP.

On top of the ECM7.0, simulation results of the proposed method are reported as below:

RA: Class F -0.07%, 100%, 99%; Class TGM -0.46%, 100%, 100%

LB: Class F -0.23%, 99%, 99%; Class TGM -0.51%, 100%, 100%

The BVs are already stored in ECM7 for the current picture, this proposal would additionally require the BVs and an idenitification whether a block has BV or MV to be stored for the reference pictures.

No scaling is performed for temporal BVs.

This contribution is tested in combination with JVET-AC0193. See further notes under JVET-AC0193.

Investigate in EE.

[JVET-AC0349](https://jvet-experts.org/doc_end_user/current_document.php?id=12554) Crosscheck of JVET-AC0192 (Non-EE2 Temporal block vector prediction) [L. Zhang (OPPO)] [late]

[JVET-AC0193](https://jvet-experts.org/doc_end_user/current_document.php?id=12397) Non-EE2: Copy-Padding for IBC [N. Zhang, K. Zhang, L. Zhang (Bytedance)]

In IBC mode, the reference block cannot be overlapped with the current block. In this contribution, copy-padding can be applied on the overlapped area. With copy-padding, an unreconstructed sample in the overlapped region can be padded by copying its prediction sample.

On top of the ECM-7.0, simulation results of the proposed method are reported as below:

AI: Class F -0.10%, 101%, 100%; Class TGM -0.25%, 99%, 100%

RA: Class F -0.13% , 100%, 99%; Class TGM -0.27%, 100%, 100%

On top of the ECM-7.0, simulation results of the proposed method combined with the temporal block vector prediction (TBVP) in JVET-AC0192 are reported as below:

AI: Class F -0.10%, 100%, 99%; Class TGM -0.25%, 101%, 100%

RA: Class F -0.17%, 101%, 99%; Class TGM -0.68%, 101%, 99%

It was commented that the proposed padding is performed sample-by-sample, which could result in some complexity increase. It would be desirable to investigate whether adding some constraints could help reduce this burden.

Investigate in EE. Also investigate whether the current sample-by-sample padding can be simplified.

[JVET-AC0350](https://jvet-experts.org/doc_end_user/current_document.php?id=12555) Crosscheck of JVET-AC0193 (Non-EE2: Copy-Padding for IBC) [L. Zhang (OPPO)] [late]

[JVET-AC0198](https://jvet-experts.org/doc_end_user/current_document.php?id=12402) Non-EE2: IntraTMP with multiple modes [P.-H. Lin, J.-L- Lin, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes to introduce additional intra template matching (IntraTMP) modes to enhance the coding performance without increasing much complexity for both encoder and decoder. The experimental results show that the luma BDrate gain is -0.13% with 101%/101% encoding/decoding time in AI condition on top of ECM-7.0.

Three modes: left template, above template, and L-shaped

Is the encoder run time reliable? There is reduction in some classes, and increase (up to 8%) in particular for the classes A.

It was pointed out that there is a division in the determination of the fusion weights. It was answered that similar determination methods are already in other places as of now.

No change to chroma

What is the individual benefit of the two aspects (adaptation of templates, modification of the fusion)? Should also be studied separately.

Cost for the three template modes is partially e-used, as the left and above templates are subsets of the L-shape.

The modified fusion is only used for L-shape.

Study in EE.

[JVET-AC0200](https://jvet-experts.org/doc_end_user/current_document.php?id=12404) AhG12 Dynamic CABAC models [F. Lo Bianco, F. Galpin, E. François (InterDigital)] [late]

A dynamic CABAC model is presented where a set of CABAC parameters modifications can be optimally selected at encoder in a set of N choices. The method allows to better adapt the CABAC model of each context to the data content without additional complexity at the decoder. It also allows to decrease the dependencies between the CABAC performance and the fixed CABAC parameters. The following results are reported for Y, U V bdrate and encoding/decoding time:

Test 1:

AI -0.02% -0.02% -0.02%, 102% and 101%

RA: -0.01% 0.14% 0.27%, 101% and 101%

Test 2:

AI -0.02% -0.02% -0.02%, 102% and 101%

RA -0.04% 0.18% 0.26%, 101% and 100%

without additional encoding or decoding runtime.

An offset value is determined by RDO at the encoder (using a table) and transmitted to the decoder, where it is applied to CABAC initialization parameters (CABAC process unchanged)

Test 1: 4 values in table; Test 2: 8 values

2 bits are used to transmit the table index of the offset values

Is the offset signalled individually for each CABAC context? No, but each context has its own lookup table.

Improvement in coding efficiency is very small. Proponents hope to improve by better training the tables. It was also found that some chroma loss occurs in RA

Cross-checker did not find encoder run time increase.

Interesting approach, but it is recommended to further study the possibility of compression improvement, not mature for EE yet.

[JVET-AC0237](https://jvet-experts.org/doc_end_user/current_document.php?id=12441) Cross-check of JVET-AC0200 [F. Le Léannec (Xiaomi)] [late]

[JVET-AC0201](https://jvet-experts.org/doc_end_user/current_document.php?id=12405) AHG12: CIIP extension with Intra Template Matching [K. Naser, P. Bordes, F. Galpin, K. Reuzé, A. Robert (InterDigital)] [late]

This contribution proposes an extension to CIIP to replace the inter-prediction part by IntraTMP. This enables CIIP with intra prediction only, and thus allow CIIP in I-slices. Compared to ECM-7.0 anchor, the following results are obtained:

AI: -0.08%, -0.04%, -0.07% with 99% EncT and 100% DecT

RA: xx%, xx%, xx% with xx EncT and xx DecT

In this contribution, additional results are provided where the proposed method was tested on top of JVET-AC0070, and the gain of 0.08% in AI was shown to be additive with the gains from JVET-AC0070.

The proponent of JVET-AC0070 suggests that enlarged search range (part of JVET-AC0070) can bring benefits to this method.

Cross checker reports that partial simulation results are matched. It was commented that this extension would allow CIIP to be used in intra slices. The proposal also enables the method in inter slices.

The fusion in this proposal uses the existing CIIP weights in ECM-7.0.

Differences between this method and the method in JVET-AC0170 are: 1. How the mode is signalled, 2. The fusion process (JVET-AC0170 uses block-level fixed weights), 3. The template used in IntraTMP search in JVET-AC0170 is modified and this proposal uses the existing template.

Investigate in EE. Consider combination with other intraTMP related EE tests.

[JVET-AC0320](https://jvet-experts.org/doc_end_user/current_document.php?id=12525) Crosscheck of JVET-AC0201 (AHG12: CIIP extension with Intra Template Matching) [A. Filippov, V. Rufitskiy (Ofinno)] [late]

[JVET-AC0203](https://jvet-experts.org/doc_end_user/current_document.php?id=12407) AHG12: On LMCS luma mapping in template processing for IBC TM-based tools [A. Filippov, V. Rufitskiy (Ofinno)] [late]

This contribution presents a technique of harmonizing the luma mapping process with the mechanism of computing template matching (TM) costs for IBC TM-based tools that use templates of a current block and candidate reference blocks in different domains, i.e. before and after LMCS inverse luma mapping. This proposal aims at aligning domains where TM costs are calculated for templates of the current block and the candidate reference blocks in P- and B-slices. We investigate 2 methods of how to perform this alignment:

* Method 1: TM costs are computed in spatial domain (i.e. after LMCS inverse luma mapping);
* Method 2: TM costs are computed in LMCS (mapped) domain (i.e. before LMCS inverse luma mapping).

Both methods were implemented and tested on top of ECM-7.0. Reportedly, the following BD-rate gain results are obtained for both methods:

* Method 1: RA configuration: {-0.00%, 0.01%, 0.08%} for class F and {-0.02%, -0.05%, -0.02%} for class TGM; LDB configuration: {-0.03%, 0.03%, -0.25%} for class F and {-0.07%, 0.06%, 0.01%} for class TGM;
* Method 2: RA configuration: {0.01%, 0.05%, 0.06%} for class F and {-0.03%, -0.05%, -0.03%} for class TGM; LDB configuration: {0.07%, 0.06%, -0.19%} for class F and {0.00%, 0.22%, 0.22%} for class TGM.

It was commented that the mapping of LMCS for the case of screen content is completely different from the case of HDR camera captured content, close to identity. This could be the reason why the gain reported here is very small (even though it can be assumed that there is some inconsistency). It is recommended to study (not in EE) the method with HDR camera captured content and IBC enabled (see JVET-AC0161, non-CTC), to find if the gain is larger for that case, and which of the two methods is better.

[JVET-AC0325](https://jvet-experts.org/doc_end_user/current_document.php?id=12530) Crosscheck of JVET-AC0203 (AHG12: On LMCS luma mapping in template processing for IBC TM-based tools) [K. Naser (InterDigital)] [late]

[JVET-AC0327](https://jvet-experts.org/doc_end_user/current_document.php?id=12532) Crosscheck of JVET-AC0203 (AHG12: On LMCS luma mapping in template processing for IBC TM-based tools) [M. Blestel (Xiaomi)] [late]

[JVET-AC0212](https://jvet-experts.org/doc_end_user/current_document.php?id=12416) Non-EE2: Template matching for IBC BVD suffix derivation [P. Nikitin, Z. Zhang, H. Huang, V. Seregin, M. Karczewicz (Qualcomm)] [late]

This contribution proposes BVD suffix derived using template matching search, and a flag is signalled to indicate whether the suffix is transmitted or equal to zero.

AI: -0.29% (Y), -0.10% (U), -0.37% (V), 99% (EncT), 102% (DecT) for class F

-0.68% (Y), -0.63% (U), -0.63% (V), 103% (EncT), 102% (DecT) for class TGM

RA: -0.27% (Y), -0.17% (U), -0.25% (V), 94% (EncT), 96% (DecT) for class F

-0.39% (Y), -0.45% (U), -0.28% (V), 102% (EncT), 102% (DecT) for class TGM

RDO search at encoder? No.

How many template matching operations are performed? Depends on size of prefix, could become large when prefix is large (due to the case that BVD is typically rather small, at least for class F). This might be the reason why for some sequences runtime is significantly larger. Worst case complexity might be rather large.

How would that combine with new adoptions, in particular sign prediction? Might not be straightforward. One expert comments that according to his opinion gain might no longer be achievable on top of new ECM adoptions. This is however a case that always may happen, and in cases when the gain disappears ontop of new ECM, technology investigated in EE would not be adopted, anyway.

Investigate in EE. Beyond the best way of combining with new adoped proposals, also methods should be investigated to avoid excessively large search ranges.

[JVET-AC0322](https://jvet-experts.org/doc_end_user/current_document.php?id=12527) Crosscheck of JVET-AC0212 (Non-EE2: Template matching for IBC BVD suffix derivation) [M. Radosavljević (Xiaomi)] [late]

[JVET-AC0213](https://jvet-experts.org/doc_end_user/current_document.php?id=12417) Non-EE2: SbTMVP with MMVD [L.-F. Chen, R. Chernyak, X. Zhao, X. Xu, S. Liu (Tencent)] [late]

This contribution proposes the MMVD on SbTMVP candidate in subblock MMVD merge list. The syntax element, MMVD index, is signaled to indicate the MMVD offset for the SbTMVP merge candidate, and this MMVD offset is used to derive the offset value of the displacement vector (DV) for each SbTMVP-MMVD candidate. The final DV for each SbTMVP-MMVD candidate is calculated from the DV of SbTMVP merge candidate with the selected offset. By using the different DV offset, different subblock-based motion field could be obtained to form a SbTMVP-MMVD candidate list. In order to improve the coding performance, the subblock-based template-matching is also used to reorder all SbTMVP-MMVD candidates in the candidate list by using the template-matching (TM) cost in ascending order. Finally, only the first 16 candidates in the reordered candidate list could be signaled. The overall simulation results on top of ECM-7.0 reference software are reported as below.

* RA: -0.07%/-0.07%/-0.09% with 103% EncT and 103% DecT
* LB: 0.xx%/0.xx%/0.xx% with xxx% EncT and xxx% DecT

Decoding time is larger than in JVET-AC0103 due to TM-based reordering. Another aspect is that SbTMVP candidate derivation is modified in JVET-AC0103 to avoid unavailable case, whereas JVET-AC0213 leaves it unchanged.

Investigate in EE. Reduction of encoding and decoding time should be targeted. Identify if there is some aspect that a combination with JVET-AC0103 would provide benefit, or if the two methods are just competing.

[JVET-AC0217](https://jvet-experts.org/doc_end_user/current_document.php?id=12421) Crosscheck of JVET-AC0213 on SbTMVP with MMVD [X. Li (Google)] [late]

[JVET-AC0239](https://jvet-experts.org/doc_end_user/current_document.php?id=12443) Non-EE2: Prediction of MVD magnitude suffix bins [A. Filippov, V. Rufitskiy, K. Suverov (Ofinno)] [late]

This contribution presents a technique for predicting MVD magnitude suffix bins. In fact, we proposed a technique that is an extension of MVD sign prediction, which is a part of ECM-7.0. In this design, not just signs but also magnitude suffix bins of exponential Golomb code 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-7.0, the following BD-rate gain results are reportedly obtained for Test 1, where a maximum of 6 bins (including sign bins) are predicted for blocks with width and height larger than 4 (affine motion compensation is enabled for these block sizes), and maximum of 2 bins are predicted for blocks that have either width or height equal to 4 (affine motion compensation is disabled for these block sizes) {-0.05%, -0.00%, -0.01%} for RA configuration and {-0.08%, 0.02%, -0.03%} for LDB configuration. On top of ECM-7.0, the following BD-rate gain results are reportedly obtained for Test 2, where a maximum of 6 bins (including sign bins) are predicted for blocks with width and height larger than 4 (affine motion compensation is enabled for these block sizes), and maximum of 4 bins are predicted for blocks that have either width or height equal to 4 (affine motion compensation is disabled for these block sizes) {-0.08%, -0.02%, -0.04%} for RA configuration and {-0.11%, -0.06%, -0.06%} for LDB configuration.

Reasonable tradeoff, and unlikely that the gain will disappear due to new adoptions.

Study in EE.

[JVET-AC0315](https://jvet-experts.org/doc_end_user/current_document.php?id=12520) Non-EE2: Cross-component merge mode for chroma intra coding [C.-M. Tsai, H.-Y. Tseng, C.-Y. Chuang, C.-W. Hsu, C.-Y. Chen, T.-D. Chuang, O. Chubach, Y.-W. Chen, Y.-W. Huang, S.-M. Lei (MediaTek)] [late]

In this proposal, a cross-component merge (CCMerge) mode for chroma intra coding is introduced, comprising of cross-component model parameters inheritance for the current chroma block from its spatial adjacent and non-adjacent neighbors, or default models. The spatial adjacent and non-adjacent neighboring information is collected from the previously blocks coded by CCLM, MMLM, CCCM, GLM, chroma fusion, and CCMerge modes.

On top of ECM-7.0, simulation results of the proposed method are reported as below:  
AI: {Y BD-rate = -0.12%, Cb BD-rate = -1.95%, Cr BD-rate = -2.08%, EncT = 107%, DecT = 100%}

In this contribution, if the merge candidate represents 2-parameter linear model, then the offset is not inherited, and is instead derived.

This contribution applies the merge concept to all cross-component prediction methods, whereas JVET-AC0176 applies the history-based concept only to CCCM and CCLM, which could be a reason for this contribution having higher gains.

It would be desirable to reduce the encoding time – currently full RDO is performed, resulting in encoding time increase.

Investigate in EE, including encoding time reduction and possible combination with JVET-AC0176.

[JVET-AC0341](https://jvet-experts.org/doc_end_user/current_document.php?id=12546) Crosscheck of JVET-AC0315 (Non-EE2: Cross-component merge mode for chroma intra coding) [J. Chen (Alibaba)] [late]

[JVET-AC0335](https://jvet-experts.org/doc_end_user/current_document.php?id=12540) Non-EE2: Content adaptive OBMC enabling [K. Cui, Z. Zhang, V. Seregin, M. Karczewicz (Qualcomm)] [late]

This contribution proposes a method to adaptively control OBMC based on hash block hit percentage. It reports {-0.02%, -0.02%, -0.03%} YUV BD-rate gains for Class F and {-1.28%, -0.80%, -0.69%} YUV BD-rate gains for Class TGM under RA configuration.

Decision is done on basis of first frame for entire sequence. It is noted that JVET-0190 also includes one non-normative method of disabling OBMC for screen content by sequence-level flag, but not using an automatic setting of that flag as in this contribution.

This is an encoder-only method, for information only.

It was asked if the method reliably recognizes screen content. In class F, OBMC is not turned off for two sequences with mixed content. It was reported that turning OBMC off for entire class F would result in 0.45% loss for that class.

It is asserted that this is a straightforward approach to automatically disable OBMC without class-dependent settings. This is a non-normative encoder change. As an E is planned for disabling OBMC at block level, it is also interesting to use this method as an anchor.

Decision(SW): Adopt JVET-AC0335. Investigate during ECM8 software integration if it has any impact on camera-captured content. If not, it does not need to made configurable (but should be possible to be disabled by macro). If it would have negative impact on camera-captured content, it should be made configurable, and in CTC only be enabled for classes F and TGM.

# High-level syntax (HLS) and related proposals (32)

## AHG9: SEI messages on neural-network post filter (23)

Contributions in this area were initially discussed in session 5 at 1300–1515 UTC on Thursday 12 Jan. 2023 (chaired by JRO) – see notes under JVET-AC0208. Further review in BoG JVET-AC0324 (G. Sullivan)

[JVET-AC0324](https://jvet-experts.org/doc_end_user/current_document.php?id=12529) BoG report on neural-network based post-filtering contributions [G. J. Sullivan (BoG coordinator)]

Version 3 of the report was presented in session 21 at 1305-1320 on Wednesday 18 Jan. 2023 (chaired by JRO). All recommendations up to this version were decided to be approved. The BoG planned to meet again in parallel with session 24.

During the discussion, it was mentioned that an additional example where new pictures (or components of other modality) are generated at the output of an NNPF could be depth maps generated from video pictures.

It was reported during session 25 that the BoG had finished on all open issues. No items left that would need to be further revised. Version 4 of the BoG report which was shortly presented reflects the recommendations of the BoG. No request for further discussion – JVET approved all recommendations as decisions.

[JVET-AC0208](https://jvet-experts.org/doc_end_user/current_document.php?id=12412) AHG9: Summary of Proposals Related to Neural-Network Post-Filter SEI Messages [S. Deshpande (Sharp), Y.-K. Wang (Bytedance)]

This document aims to provide a summary of proposals related to neural-network post-filter SEI messages, i.e., the NNPFC SEI message and the NNPFA SEI message. It is suggested that this summary is used for the reviewing of these proposals, such that the discussions can be in a more structured and efficient manner. Contents of v3 version of this document is same as v2 version, with all change marks accepted. In v4 summary item and corresponding JVET proposal is updated based on the v2 of that JVET proposal. In v5 one new document summary is added and some other summaries are modified based on updated inputs.

It is intended to generate an output document “Improvements under consideration for neural network post filter SEI messages” (called IuC in the sequel).

* **Summary and List of Questions**

This section aims to list a summary of each proposal and a list of design questions related to those proposals.

1. NNPFC: Should the value range of StrengthControlVal be transferred from between 0 and 1 to the range of the input tensor. If yes, add the following? (JVET-AC0047 proposal 2)

When nnpfc\_inp\_format\_idc is equal to 1, the variable StrengthControlVal is modified as follows:

StrengthControlVal = Floor ( StrengthControlVal \* ( ( 1  <<  inpTensorBitDepth ) – 1 ) )

Agreed that it makes sense to define the StrengthControlVal as integer when the input data are integer. It had historically been used to give the network information about the QP of the coded video, for which 6 bits would be sufficient. In this proposal, the precision is aligned with the bit depth of the input, which seems to make sense as the strength control could also be used for other input rather than QP.

Decision: Include in IuC document.

2. NNPFA: Add a new flag to handle output order which may be different than decoding order persistence of post-filtering process. Thus signal a new decoding\_order\_flag if nnpfa\_persistence\_flag is equal to 1? (JVET-AC0047 proposal 3)

Proponents explain that an example where this would be needed is when filtering is not applied to all pictures, e.g. not filtering certain temporal hierarchy levels of B pictures. It is claimed that, if such a flag would not exist, the encoder might have a problem. Other experts believe that there is no problem if the activation is used properly. The concept of persistence does not mean that the activation must be enabled for every picture. No action.

3. NNPFC: Move the syntax element nnpfc\_purpose to earlier position in NNPFC and signal it unconditionally? (JVET-AC0061 proposal 1)

The intent of an earlier position is to avoid parsing of other syntax elements if a decoding devices would not want to execute a certain purpose. It is asserted that in the case when the syntax element is moved, the conditional signalling of that syntax element is also unnecessary. The proposal also suggests to send the purpose along with a partial update (which is also agreed), however there should be a constraint that the purpose of the update remains the same.

Decision: Include in IuC, with constraint that purpose of a partial update shall not change relative to the base filter.

4. NNPFC: Fix the bug in StoreOutputTensors equation 83’s for loop by changing numInputPics to numOutputPics (JVET-AC0061 proposal 2A, JVET-AC0127, JVET-AC0154 item 8)

Obvious bug fix. Decision: Include in IuC

5. NNPFC: Fix the bug that for a given case the value of numOuputPics is unspecified, by deriving numOuputPics to be equal to 1 when nnpfc\_purpose has value other than value 5 (JVET-AC0061 proposal 2B, JVET-AC0127, JVET-AC0154 item 8)

Obvious bug fix. Decision: Include in IuC

6. NNPFC: Handle the undefined bitdepth for chroma sample values in the input integer tensor? If yes choose from one of the options below (JVET-AC0061 proposal 3)

Option A: Define a syntax element for bit depth of chroma sample values in the input integer tensor

Option B: Modify semantics of existing syntax element for luma bitdepth to also apply to chroma bitdepth

It is agreed that an action is necessary. It is commented that the same problem occurs for the output tensor. More experts expressed preference for option A, as it allows more flexibility and adding the additional syntax element(s) is straightforward.

Decision: Include option A in IuC (both for input and output tensors)

7. NNPFC: Specify just the range of the input tensor instead of specifying the detailed input process, thus removing the detailed input process? (JVET-AC0154 item 6)

Several experts expressed the opinion that the detailed process should not be removed. For example, it is mentined that the data input to the network should use an equivalent conversion that was used in training. No agreement for this change. It is however emphasized that exact processing in an SEI message’s implementation is not mandatory.

8. NNPFC: Modify a constraint when nnpfc\_constant\_patch\_size\_flag is equal to 0? (JVET-AC0062 proposal A)

Modification: The value of inpPatchWidth + 2\*overlapSize shall be a positive integer multiple of nnpfc\_patch\_width\_minus1 + 1 and inpPatchWidth shall be less than or equal to CroppedWidth. The value of inpPatchHeight  + 2\*overlapSize shall be a positive integer multiple of nnpfc\_patch\_height\_minus1 + 1 and inpPatchHeight  shall be less than or equal to CroppedHeight.

The proponent reports that the problem was found when implementing a filter with overlap. It was initially thought that the problem would also exist for the case nnpfc\_constant\_patch\_size\_flag is equal to 1, but that was not the case, as in that case it is not required that the input data width/height is an integer multiple of patch width/height.

It is generally agreed that a problem exists. It is further mentioned that part of the problem may be that the meaning of the syntax elements for patch width/height is somehat different for the two possible values of nnpfc\_constant\_patch\_size\_flag. There could be other ways solving the problem.

[Note: Items 9 to 20 are on picture rate upsampling.]

Only item 10. below was initially discussed in session 5.

9. NNPFC: Add two proposed constraints related to picture rate upsampling? (JVET-AC0085 proposal 1)

1. 16× numInputPics >= numOutputPics

Constraint 1 aims to constrain the upper bound for picture rate upsampling in terms of the ratio between numOutputPics and numInputPics to avoid unlimited output number of the pictures (considering computational capability, the implementation cost, and assertedly inferior performance of picture rate upsampling if the output picture number is too much larger than the input picture number, especially for the high temporal complexity scenarios).

1. numInputPics <=16× numOutputPics

Constraint 2 aims to constrain the lower bound for picture rate upsampling. (Assertion: If the temporal complexity of the video sequence is high, more input pictures are required for better picture rate upsampling performance, as in proposed Constraint 1. On the contrary, if the temporal complexity is low, too many input picture for temporal rate upsampling is not necessary, which is proposed by Constraint 2.)

10. NNPFC: Replace the current syntax elements nnpfc\_num\_input\_pics\_minus2 and nnpfc\_interpolated\_pics[ i ] with two new syntax elements for picture rate upsampling? the first specifying the number of interpolated pictures between the current picture and the previous picture in output order minus 1, and the second specifying half of the number of input pictures minus 1, where the input pictures, including the current picture itself, are consecutive in output order and listed in increasing output order, and half of them precede the current picture in output order, and half minus 1 of them succeed the current picture in output order? (JVET-AC0085 proposal 2)

Was initially reviewed in session 5. It is the opinion of several experts that there is no problem with the current syntax, and the proposal would give up some flexibility (e.g., number of input pictures needs to be an even number). The main intent of the proposal is a syntax simplification. Was further discussed in BoG after reviewing other proposals.

Question is raised if there is a software implementation of any changes suggested?

Further review to be conducted in BoG (G. Sullivan).

* **List of Documents**

This section lists the documents (and abstracts) for which summary in section 1 is prepared.

[**JVET-AC0047**](https://jvet-experts.org/doc_end_user/current_document.php?id=12232) AHG9: On bugfix for NNPFC and NNPFA SEI messages, [T. Chujoh](mailto:chujoh.takeshi@sharp.co.jp), Y. Yasugi, T. Ikai (Sharp)

In this contribution, several problems of NNPFC (neural-network post-filter characteristics) SEI and NNPFA (neural-network post-filter activation) SEI messages are reported, and their solutions are proposed. The first problem is the syntax of NNPFC SEI is not correct when the value of nnpfc\_purpose is equal to 4, which is solved by a syntax change. The second problem is even if nnpfc\_inp\_format\_idc is equal to 1, a filtering strength control value, StrengthControlVal, is a real number in the range of 0 to 1, which is solved by adding an equation for the transform to an unsigned integer. The third problem is if the output order is different from the decoding order, the encoder may not be able to decide the persistence of post-filter processing, which is solved by adding a flag indicating the order when nnpfa\_persistence\_flag of NNPFA SEI is equal to 1.

[**JVET-AC0061**](https://jvet-experts.org/doc_end_user/current_document.php?id=12263) AHG9: Comments on Neural-network Post-filter Characteristics SEI Message, [S. Deshpande (Sharp)](mailto:sdeshpande@sharplabs.com)

Some modifications are proposed to the Neural-network post-filter characteristics SEI Message. Following is proposed:

* Proposal 1: It is proposed to move nnpfc\_purpose syntax element to signal it earlier in the NNPFC SEI message.
* Proposal 2: Two bug-fixes are proposed related to variables in NNPFC SEI message.
  + Bug-fix A: It is asserted that the outermost for loop in equation 84 which defines the process StoreOutputTensors( ), should use the variable numOutputPics instead of numInputPics.
  + Bug-fix B: It is asserted that when nnpfc\_purpose has a value other than value 5, the variable numOutputPics is not derived and it should be derived.
* Proposal 3: It is proposed to either define a new syntax element for specifying the bit depth of chroma sample values in the input integer tensor or to define the semantics of an existing syntax element to apply to chroma samples in addition to the luma samples.

**[JVET-AC0062](https://jvet-experts.org/doc_end_user/current_document.php?id=12264)** AHG9: On NNPFC SEI, [S. Deshpande](mailto:sdeshpande@sharplabs.com), A. Sidiya (Sharp)

Some modifications are proposed to the Neural-network post-filter characteristics (NNPFC) SEI Message. Following is proposed:

* Proposal A: It is proposed to modify a constraint when nnpfc\_constant\_patch\_size\_flag is equal to 1. It is asserted that without the proposed modification, the actual input size to the post-processing filter may not be a multiple of a specific value as may be required by some typical neural network post filters.
* Proposal B: It is proposed to define a valid value range for the two ue(v) coded syntax elements nnpfc\_num\_input\_pics\_minus2 and nnpfc\_interpolated\_pics[ i ]. Currently there is no range specified for these two syntax elements.

[**JVET-AC0074**](https://jvet-experts.org/doc_end_user/current_document.php?id=12276) AHG9: On the VVC use of the NNPFC SEI message for picture rate upsampling, [M. M. Hannuksela (Nokia)](mailto:miska.hannuksela@nokia.com)

The picture rate upsampling purpose was added to the neural-network post-filter characteristics (NNPFC) SEI message in the JVET-AB meeting. A picture rate upsampling post-filter uses two or more input pictures to generate one or more interpolated pictures.

The contribution responds to the following asserted issues:

1. Missing VVC specification text for inputting multiple input pictures for a picture rate upsampling post-filter.
2. The following question that was left for further study in the JVET-AB meeting: Does the selection of the input pictures for a picture rate upsampling post-filter depend on TemporalId values?
3. The use of temporal interleaving frame packing arrangement with a frame rate upsampling post-filter.
4. The use of a picture rate upsampling post-filter when picture widths and/or heights of input pictures for the post-filter vary.

This contribution proposes VVC specification text for inputting multiple input pictures for a picture rate upsampling post-filter and has the following main ingredients:

1. The input sample arrays CroppedYPic[ idx ], CroppedCbPic[ idx ] and CroppedCrPic[ idx ] for increasing order of idx are set in decreasing order of POC regardless of the TemporalId values, where idx equal to 0 corresponds to the current picture where the post-filter is activated.
2. When a frame rate upsampling post-filter is applied with temporal interleaving frame packing arrangement, all the input pictures are selected to be either first constituent frames or second constituent frames.
3. Either of the following is proposed:
   1. Add a constraint that frame rate upsampling can only be used when all input pictures have the same dimensions.
   2. Add a constraint that frame rate upsampling can only be used when all input pictures have the same dimensions after applying super resolution post-filter applicable to the input picture, if any.

[**JVET-AC0075**](https://jvet-experts.org/doc_end_user/current_document.php?id=12277) AHG9: On NNPFC and NNPFA SEI messages for picture rate upsampling post-filter, [M. M. Hannuksela (Nokia)](mailto:miska.hannuksela@nokia.com)

This contribution proposes to indicate a validity range of TemporalId values for a frame rate upsampling post-filter. If the highest temporal sublayer to be decoded is within the indicated validity range of TemporalId values, the frame rate upsampling post-filter is applicable. The validity range is proposed to be indicated either in the neural-network post-filter activation (NNPFA) SEI message (option 1) or in the neural-network post-filter characteristics (NNPFC) SEI message (option 2).

The proposal has the following asserted benefits:

1. The encoder can indicate that no frame rate upsampling should be performed for a sub-bitstream having such a low frame rate where frame rate upsampling is unlikely to provide satisfactory results as judged by the encoder.
2. More than one frame rate upsampling post-filter optimized for different input frame rates can be defined and selectively applied based on the highest temporal sublayer to be decoded.

[**JVET-AC0076**](https://jvet-experts.org/doc_end_user/current_document.php?id=12278) AHG9/AHG15: On the NNPFC SEI message for machine analysis, [M. M. Hannuksela](mailto:miska.hannuksela@nokia.com), F. Cricri, J. I. Ahonen, H. Zhang (Nokia)

The contribution reviews some earlier works on neural-network post-filtering targeted at improving machine analysis precision and concludes, based on the earlier works, that post-filters targeted at machine analysis can have different purposes (i.e., nnpfc\_purpose values) in the neural-network post-filter characteristics (NNPFC) SEI message.

The contribution proposes to add the target usage of a post-processing filter in the NNPFC SEI message to indicate whether the filtered video is suitable for any usage, is intended for user viewing, or is expected to be provided as input to machine analysis.

The proposed target usage is intended to be used for selecting which post-filters are applied as follows:

* When a decoding device displays the video for user viewing rather than performs machine analysis, any post-processing filter that has only machine analysis as the target usage is suggested to be omitted.
* When a decoding device performs machine analysis rather than displays the video, any post-processing filter that has only user viewing as the target usage is suggested to be omitted.

[**JVET-AC0085**](https://jvet-experts.org/doc_end_user/current_document.php?id=12287) AHG9: On neural-network post-filter characteristics (NNPFC) SEI message for temporal upsampling towards machine vision, [S. Wang](mailto:srwang3-c@my.cityu.edu.hk), [J. Chen](mailto:jiechen.cj@alibaba-inc.com), [Y. Ye (Alibaba)](mailto:yan.ye@alibaba-inc.com), [S. Wang (CityU)](mailto:shiqwang@cityu.edu.hk)

There are two proposals in the contribution.

The first one is to impose constraints on the input number (numInputPics) and output number (numOutputPics) of pictures for picture rate upsampling based on the performance and implementation cost consideration. The constraints were included in the proposed design of version 1 of the contribution and was abstracted as a separated proposal in version 2.

The second one is to simplify the current design of picture rate upsampling in NNPFC SEI message by replacing nnpfc\_num\_input\_pics\_minus2 and nnpfc\_interpolated\_pics[ i ] with two new syntax elements to specify the number of interpolated pictures between the current picture and the previous picture in output order minus 1, and half of the number of input pictures minus 1. The simplified design was the same as the one proposed in version 1 of the contribution and was added as proposal 2 in version 3.

[**JVET-AC0127**](https://jvet-experts.org/doc_end_user/current_document.php?id=12329) AHG9: Combination of picture rate upsampling with other NNPF purposes, [Y.-K. Wang](mailto:yekui.wang@bytedance.com), [J. Xu](mailto:xujizheng@bytedance.com), [Y. Li](mailto:yue.li@bytedance.com), [L. Zhang](mailto:lizhang.idm@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [J. Li](mailto:lijunru@bytedance.com), [C. Lin (Bytedance)](mailto:linchaoyi.cy@bytedance.com)

Currently, neural-network post-filter (NNPF) purposes with chroma upsampling only, resolution upsampling only, combination of chroma upsampling and resolution upsampling, and picture rate upsampling only are specified.

However, picture rate upsampling in combination with other types of upsampling can also be used in video applications. This contribution proposes to allow such combination of NNPF purposes.

[**JVET-AC0128**](https://jvet-experts.org/doc_end_user/current_document.php?id=12330) AHG9: On the signalling of complexity information in NNPFC SEI message, [Hendry](mailto:dr.hendry@lge.com), S. Kim (LGE)

In the current VSEI specification, the signalling of complexity information in NNPFC SEI message may be present only in the SEI message that contains base filter (i.e., the first NNPFC SEI message with a particular nnpfc\_id). It is asserted that an NNPFC SEI message that updates a base NNPFC SEI message may have different complexity, which can be higher or lower. Therefore, the current design of the complexity information signalling in NNPFC SEI message puts limitation such that an NNPFC SEI message that updates a base NNPFC SEI message cannot have different complexity.

This contribution proposes 2 options for modifying the signalling of complexity information in NNPFC SEI message, they are:

*Option 1:* Move the signalling related to complexity information outside the nnpfc\_formatting\_and\_purpose\_flag condition. Furthermore, update the semantics for the complexity signalling such when complexity information is not present in an NNPFC SEI message that contains update is not present, it is inferred to be the same as the complexity information signalled in the NNPFC SEI message with the same Id that contains the base filter.

*Option 2:* Move the signalling related to complexity information outside the nnpfc\_formatting\_and\_purpose\_flag condition and change the nnpfc\_complexity\_info\_present\_flag to be nnpfc\_complexity\_info\_present\_idc (2 bits). The semantics of this indication is as follows:

* Value 0 means that syntax elements related to complexity information are not present in the SEI message. If the SEI message is an NNPFC SEI message that is not a base NNPFC SEI message, the complexity information signalled in the base SEI message does not apply to the SEI message.
* Value 1 means that syntax elements related to complexity information are not present in the SEI message. If the SEI message is an NNPFC SEI message that is not a base NNPFC SEI message and the base SEI message associated with this SEI message contains complexity information, the complexity information applies to this SEI message as well.
* Value 2 means that syntax elements related to complexity information are present in the SEI message.
* Value 3 is reserved.

[**JVET-AC0129**](https://jvet-experts.org/doc_end_user/current_document.php?id=12331) AHG9: On the NNPFC SEI message update and activation, [Hendry](mailto:dr.hendry@lge.com), [J. Nam](mailto:junghak.nam@lge.com), H. Jang, S. Kim, J. Lim (LGE)

It is asserted that the current VSEI specification is not clear about which neural-network post-filter (NNPF) applies to a picture when an NNPFC SEI message containing an update filter is present, prior to the picture, in the bitstream after an NNPFC SEI message containing the base filter has been activated. The asserted problem in such situation is which filter to be applied to the picture, the base filter (i.e., the filter in the NNPFC SEI message that is present prior to the NNPFA SEI message) or the updated filter (i.e., the filter in the NNPFC SEI message that is present after the NNPFA SEI message).

In this contribution, three possible solutions were presented as follows:

1. Specifying that the NNPF that applies to a picture is the one that is in the NNPFC that immediately precedes the picture in output order.
2. Specifying that the NNPF that applies to a picture is the one that is in the NNPFC that immediately precedes the NNPFA in output order.
3. Specifying that when an NNPFC SEI message is present, it cancels the persistence of previous NNPFA SEI message with the same ID, if present.

The authors proposed that option 2 (first choice) or option 3 (second choice) are used to solve the asserted problem.

[**JVET-AC0131**](https://jvet-experts.org/doc_end_user/current_document.php?id=12333) AHG9: On NNPFC SEI message repetition, [Hendry](mailto:dr.hendry@lge.com), [J. Nam](mailto:junghak.nam@lge.com), H. Jang, S. Kim, J. Lim (LGE)

In this contribution, it is asserted that the current semantics of NNPFC SEI message has the following problems:

1. It is not clear what it means when the signalling of formatting, purpose, and complexity is not present as there can be two possibilities:
   1. For NNPFC SEI message in which formatting, purpose, and complexity information is not present, those information is inferred from the NNFPC SEI message that contains the base neural-network post-processing filter, or,
   2. For NNPFC SEI message in which formatting, purpose, and complexity information is not present, those information is not known.
2. NNPFC SEI message that contains base filter cannot be repeated since only the first SEI message with a particular nnpfc\_id can have the value of that flag equal to 1. Any SEI message with same nnpfc\_id may be perceived as referencing to the first SEI message (i.e., an update).

To address the above asserted problems, the following modifications are proposed:

1. Specify that the signalling of formatting, purpose, and complexity information is present only in NNPFC SEI message that contains base filter.
   1. This means that when NNPFC SEI message contains base neural-network post-processing filter, the value of nnpfc\_formatting\_and\_purpose\_flag is equal to 1 and likewise, when the value of nnpfc\_formatting\_and\_purpose\_flag is equal to 1, the NNPFC SEI message contains base neural-network post-processing filter.
2. Specify that for NNPFC SEI message that contains the update filter, the formatting and purpose information is inferred to be the same as the information signalled in associated NNPFC SEI message (i.e., NNPFC SEI message that has same value of nnpfc\_id) that contains the base neural-network post-processing filter.
3. NNPFC SEI message that contains base neural-network post-processing filter may be repeated. When repeated, the repeated SEI message contains the same information as the first NNPFC SEI message (i.e., the first one with the particular value of nnpfc\_id).

[**JVET-AC0132**](https://jvet-experts.org/doc_end_user/current_document.php?id=12334)AHG9: On design for region-based neural-network post-filter SEI message, [Hendry](mailto:dr.hendry@lge.com), [J. Nam](mailto:junghak.nam@lge.com), H. Jang, S. Kim, J. Lim (LGE)

This contribution proposes signalling to support region-based application of neural-network post-filters. It is asserted that such capability is desirable for use-cases such as application of the filter for region of interest. The proposed signalling is based on the signalling proposed in JVET-AB0134 with the following modifications:

1. Specify a flag (i.e., nnpfc\_region\_based\_flag) in NNPFC SEI message whether or not the neural-network post-filter (NNPF) may be used for region(s) within picture(s).
2. When nnpfc\_region\_based\_flag is equal to 1, a list of candidate region size may be signalled in the NNPFC SEI message. This list of region size may be referred to from NNPFA SEI message by index.
3. Specify a two-bit indication (i.e., nnpfa\_region\_based\_idc) in NNPFA SEI message with the following semantics:
   1. nnpfa\_region\_based\_idc equals to 0 indicates that the activated NNPF applies to the whole picture.
   2. nnpfa\_region\_based\_idc equals to 1 indicates that the activated NNPF applies to region(s) of the picture. The size of the region is derived from the list of candidate region size signalled in NNPFC SEI message and the location of the region is signalled in NNPFA SEI message.
   3. nnpfa\_region\_based\_idc equals to 2 indicates that the activated NNPF applies to region(s) of the picture. For this mode, the location of the region and its size are signalled in NNPFA SEI message.
   4. nnpfa\_region\_based\_idc equals to 3 is reserved.
4. The following constraints may need to be defined:
   1. When the value of nnpfc\_region\_based\_flag is equal to 0, the value of nnpfa\_region\_based\_idc shall be equal to 0.
   2. When there is no candidate region size signalled in NNPFC SEI message, the value of nnpfa\_region\_based\_idc shall not be equal to 1.

[**JVET-AC0134**](https://jvet-experts.org/doc_end_user/current_document.php?id=12336)AHG9: Signalling of NNPF quality improvement, [C. Lin](mailto:linchaoyi.cy@bytedance.com), [Y.-K. Wang](mailto:yekui.wang@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [J. Li](mailto:lijunru@bytedance.com), [Y. Li](mailto:yue.li@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com)

In current design of the neural network post-filter characteristics (NNPFC) SEI message, the complexity of neural network post filter (NNPF) may be signalled. Given the complexity of NNPF, the decoder might decide whether to apply these post filters. However, if the quality improvement by applying the NNPF is signalled, the decoder could make more educated choice on whether to apply the NNPF. For example, if the complexity is moderate to high, by that information itself the decoder may choose not to apply the NNPF. However, given the same complexity, and at the same time the indicated quality improvement is significant, the decoder may choose to apply the NNPF anyway.

This contribution proposes to enable signalling of information of quality improvement by applying an NNPF. Since this information depends on the set of pictures to which the NNPF is applied. It is proposed to signal the information in the NNPFA SEI message instead of in the NNPFC SEI message. It is proposed to adopt one of the following three options:

1. Signal the value difference of one metric.
2. Signal the value differences of multiple metrics.
3. Signal the values of one metric with and without applying NNPF.

[**JVET-AC0135**](https://jvet-experts.org/doc_end_user/current_document.php?id=12337) AHG9: Separate activation of color components for neural-network post-filter, [C. Lin](mailto:linchaoyi.cy@bytedance.com), [Y. Li](mailto:yue.li@bytedance.com), [Y.-K. Wang](mailto:yekui.wang@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [J. Li](mailto:lijunru@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com)

Currently, all colour components are controlled collectively in the neural-network post-filter activation (NNPFA) SEI message. With such a design, all of, or none of, the colour components should be processed by the neural-network post-processing filter (NNPF). However, the NNPF may degrade the performance of one colour component while improving the performance of the others in certain cases. Therefore, it is beneficial to only activate the target NNPF for certain colour components.

This contribution proposes to provide the capability of specifying activation of NNPF separately for different components (luma or chroma) in the NNPFA SEI message. It is proposed to use one of the following two options:

1. Signal an indicator to indicate which of the components present in the output tensor are used.
2. Signal four flags to indicate which of the components present in the output tensor are used.

[**JVET-AC0136**](https://jvet-experts.org/doc_end_user/current_document.php?id=12338) AHG9: On additional NNPF purposes, [C. Lin](mailto:linchaoyi.cy@bytedance.com), [Y.-K. Wang](mailto:yekui.wang@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [Y. Li](mailto:yue.li@bytedance.com), [J. Li](mailto:lijunru@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com)

Currently, some neural network post-filter (NNPF) purposes, such as quality improvements, resolution upsampling, and picture rate upsampling, are defined. However, it is asserted that more purposes could be added, as indicated by the following sentences in the meeting minutes of the previous JVET meeting:

*"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."*

*"Further study appears necessary to come up with a list of purposes that is useful for the current version of this SEI message."*

For example, machine vision tasks are widely used, and they could be added as additional NNPF purposes. Besides, style transfer and object removal are also considered to be additional NNPF purpose.

This contribution proposes to signal machine vision tasks as an additional purpose for post processing filter in NNPFC SEI message. It is proposed to use one of the following three options:

1. Add style transfer and object removal as additional NNPF purposes.
2. Add machine vision tasks as an additional NNPF purpose and add further indicate the type of the machine vision task.
3. Change the purpose of visual quality improvement to be visual information processing, and further indicate the visual information processing type.

[**JVET-AC0152**](https://jvet-experts.org/doc_end_user/current_document.php?id=12356)AHG9: Bit-masking based representation of nnpfc\_purpose, [J. Xu](mailto:xujizheng@bytedance.com), [Y.-K. Wang](mailto:yekui.wang@bytedance.com), [L. Zhang](mailto:lizhang.idm@bytedance.com), [Y. Li (Bytedance)](mailto:yue.li@bytedance.com)

This contribution proposes bit-masking based representation of nnpfc\_purpose. Specifically, each bit or effective bit of nnpfc\_purpose denotes a primitive format change, which can be chroma format change, resolution change, frame rate change and other primitive format changes.

[**JVET-AC0153**](https://jvet-experts.org/doc_end_user/current_document.php?id=12357) AHG9: Bitdepth increase indication in the NNPFC SEI message, [J. Xu](mailto:xujizheng@bytedance.com), [Y.-K. Wang](mailto:yekui.wang@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com)

This contribution proposes to add a purpose of bitdepth increase to the neural-network post-filter characteristics (NNPFC) SEI message, to indicate the use of bitdepth increase that can be used for support related applications, e.g. SDR to HDR.

[**JVET-AC0154**](https://jvet-experts.org/doc_end_user/current_document.php?id=12358) AHG9: Miscellaneous cleanups of the neural-network post-filter SEI messages, [J. Xu](mailto:xujizheng@bytedance.com), [Y.-K. Wang](mailto:yekui.wang@bytedance.com), [L. Zhang](mailto:lizhang.idm@bytedance.com), [C. Lin (Bytedance)](mailto:linchaoyi.cy@bytedance.com)

This contribution proposes the following miscellaneous cleanup changes to the neural-network post-filter SEI messages:

1. Specify a value range for nnpfc\_num\_input\_pics\_minus2.
2. Specify a value range for nnpfc\_interpolated\_pics[ i ].
3. Add a flag to indicate whether the number of interpolated pictures between every pair of consecutive input pictures is the same.
4. Add a constraint on nnpfc\_interpolated\_pics[ i ] to avoid the all zero case.
5. Add the 4:0:0 format into the case of chroma upsampling purpose.
6. Specify just the range of the input tensor instead of specifying the detailed input process.
7. Fix a potential bug in the syntax about nnpfc\_purpose.
8. Fix a potential bug in equation (83).
9. Fix a few typos.
10. Adjust the value range of nnpfc\_num\_kmac\_operations\_idc and nnpfc\_total\_kilobyte\_size.

[**JVET-AC0174**](https://jvet-experts.org/doc_end_user/current_document.php?id=12378) AHG9: Multiple input pictures for neural-network post-processing filter, [Y. Li](mailto:yue.li@bytedance.com), [Y.-K. Wang](mailto:yekui.wang@bytedance.com), [C. Lin](mailto:linchaoyi.cy@bytedance.com), [J. Li](mailto:lijunru@bytedance.com), [J. Xu](mailto:xujizheng@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com)

Currently, the neural-network post-filter characteristics (NNPFC) SEI message specifies to use only one input picture when the post filter is for visual quality improvement, chroma upsampling, resolution upsampling, or both chroma upsampling and resolution upsampling. However, it might be beneficial to input multiple pictures for these purposes.

This contribution proposes to provide the capability of feeding multiple pictures into the neural network post-processing filter in case of all purposes.

[**JVET-AC0299**](https://jvet-experts.org/doc_end_user/current_document.php?id=12504) AHG9: On indicating processing order in the NNPFC SEI message,

Tong Shao, Arjun Arora, Peng Yin, [Sean McCarthy](mailto:sean.mccarthy@dolby.com), Taoran Lu, Fangjun Pu, Walt Husak (Dolby), [Hendry](mailto:dr.hendry@lge.com), Seung hwan Kim (LGE)

At the 28th Meeting of JVET, it was agreed that there can be several NNPFA SEI messages present for the same picture, for example, when the post-processing filters are meant for different purposes. It is asserted that it would be beneficial to enable signaling of the preferred processing order, as determined by the encoder (i.e., the content producer), for different neural-network post-filters when multiple neural-network post-filters are present. This contribution proposes to add syntax elements to the NNPFC SEI message to enable signalling of preferred processing order.

[JVET-AC0047](https://jvet-experts.org/doc_end_user/current_document.php?id=12232) AHG9: On bugfix for NNPFC and NNPFA SEI messages [T. Chujoh, Y. Yasugi, T. Ikai (Sharp)]

[JVET-AC0061](https://jvet-experts.org/doc_end_user/current_document.php?id=12263) AHG9: Comments on Neural-network Post-filter Characteristics SEI Message [S. Deshpande (Sharp)]

[JVET-AC0062](https://jvet-experts.org/doc_end_user/current_document.php?id=12264) AHG9: On NNPFC SEI Message [S. Deshpande, A. Sidiya (Sharp)]

[JVET-AC0074](https://jvet-experts.org/doc_end_user/current_document.php?id=12276) AHG9: On the VVC use of the NNPFC SEI message for picture rate upsampling [M. M. Hannuksela (Nokia)]

[JVET-AC0075](https://jvet-experts.org/doc_end_user/current_document.php?id=12277) AHG9: On NNPFC and NNPFA SEI messages for picture rate upsampling post-filter [M. M. Hannuksela (Nokia)]

[JVET-AC0076](https://jvet-experts.org/doc_end_user/current_document.php?id=12278) AHG9/AHG15: On the NNPFC SEI message for machine analysis [M. M. Hannuksela, F. Cricri, J. I. Ahonen, H. Zhang (Nokia)]

[JVET-AC0085](https://jvet-experts.org/doc_end_user/current_document.php?id=12287) AHG9: On neural-network post-filter characteristics (NNPFC) SEI message for temporal upsampling towards machine vision [S. Wang, J. Chen, Y. Ye (Alibaba), S. Wang (CityU HK)]

[JVET-AC0127](https://jvet-experts.org/doc_end_user/current_document.php?id=12329) AHG9: Combination of picture rate upsampling with other NNPF purposes [Y.-K. Wang, J. Xu, Y. Li, L. Zhang, K. Zhang, J. Li, C. Lin (Bytedance)]

[JVET-AC0128](https://jvet-experts.org/doc_end_user/current_document.php?id=12330) AHG9: On the signalling of complexity information in NNPFC SEI message [Hendry, S. Kim (LGE)]

[JVET-AC0129](https://jvet-experts.org/doc_end_user/current_document.php?id=12331) AHG9: On the NNPFC SEI message update and activation [Hendry, J. Nam, H. Jang, S. Kim, J. Lim (LGE)]

[JVET-AC0131](https://www.jvet-experts.org/doc_end_user/current_document.php?id=12333) AHG9: On NNPFC SEI message repetition [Hendry, J. Nam, H. Jang, S. Kim, J. Lim (LGE)]

[JVET-AC0132](https://www.jvet-experts.org/doc_end_user/current_document.php?id=12334) AHG9: On design for region-based neural-network post-filter SEI message [Hendry, J. Nam, H. Jang, S. Kim, J. Lim (LGE)]

[JVET-AC0134](https://jvet-experts.org/doc_end_user/current_document.php?id=12336) AHG9: Signalling of NNPF quality improvement [C. Lin, Y.-K. Wang, K. Zhang, J. Li, Y. Li, L. Zhang (Bytedance)]

[JVET-AC0135](https://jvet-experts.org/doc_end_user/current_document.php?id=12337) AHG9: Separate activation of color components for neural-network post-filter [C. Lin, Y. Li, Y.-K. Wang, K. Zhang, J. Li, L. Zhang (Bytedance)]

[JVET-AC0136](https://jvet-experts.org/doc_end_user/current_document.php?id=12338) AHG9: On additional NNPF purposes [C. Lin, Y.-K. Wang, K. Zhang, Y. Li, J. Li, L. Zhang (Bytedance)]

[JVET-AC0152](https://jvet-experts.org/doc_end_user/current_document.php?id=12356) AHG9: Bit-masking based representation of nnpfc\_purpose [J. Xu, Y.-K. Wang, L. Zhang, Y. Li (Bytedance)]

[JVET-AC0153](https://jvet-experts.org/doc_end_user/current_document.php?id=12357) AHG9: Bitdepth increase indication in the NNPFC SEI message [J. Xu, Y.-K. Wang, L. Zhang (Bytedance)]

[JVET-AC0154](https://jvet-experts.org/doc_end_user/current_document.php?id=12358) AHG9: Miscellaneous cleanups of the neural-network post-filter SEI messages [J. Xu, Y.-K. Wang, L. Zhang, C. Lin (Bytedance)]

[JVET-AC0174](https://jvet-experts.org/doc_end_user/current_document.php?id=12378) AHG9: Multiple input pictures for neural-network post-processing filter [Y. Li, Y.-K. Wang, C. Lin, J. Li, J. Xu, K. Zhang, L. Zhang (Bytedance)]

[JVET-AC0299](https://jvet-experts.org/doc_end_user/current_document.php?id=12504) AHG9: On indicating processing order in the NNPFC SEI message [T. Shao, A. Arora, P. Yin, S. McCarthy, T. Lu, F. Pu, W. Husak (Dolby), Hendry, S. Kim (LGE)] [late]

[JVET-AC0344](https://jvet-experts.org/doc_end_user/current_document.php?id=12549) [AHG9] On Neural Network Post Filter Patch Size [S. Deshpande, A. Sidiya (Sharp), M.M. Hannuksela (Nokia), Y.-K. Wang (Bytedance)] [late]

[JVET-AC0353](https://jvet-experts.org/doc_end_user/current_document.php?id=12558) AHG9: On the repetition and activation of NNPFC SEI message Hendry (LGE), Y.-K. Wang (Bytedance), M. Hannuksela, F. Cricri (Nokia), S. Deshpande (Sharplabs) [late]

## AHG9: SEI messages on topics other than NNPF (5)

Contributions in this area were discussed in session 23 at 2145–2300 UTC on Wednesday 18 Jan. 2023 (chaired by JRO), in session 23 at 1330–1415 UTC on Thursday 19 Jan. 2023 (chaired by JRO).

[JVET-AC0058](https://jvet-experts.org/doc_end_user/current_document.php?id=12260) AHG9: On the SEI processing order SEI message [Y. He, M. Coban, M. Karczewicz (Qualcomm)]

This contribution proposes the following changes for the SEI processing order SEI message:

1. It is required that an SEI processing order SEI message shall contain at least two different po\_sei\_processing\_order values.
2. It is proposed to modify the syntax to indicate the processing order of ITU-T T.35 SEI messages when present

Aspect 1. was supported by other experts, no objections were raised.

It was commented that Aspect 2. is valid but might be useful also for other types of SEI messages where prefix information is needed to identify their purpose in cases when several messages of same type (but different purpose) are concatenated and their order is relevant (NNPF or would be an example). It was further commented that there could be alternative ways to resolve the issue that aspect 2. intends to resolve. Further study on this.

It was commented that in general cascading of SEI messages is tricky, as some SEI messages might not have been designed in a way that they expect other input than the decoder output.

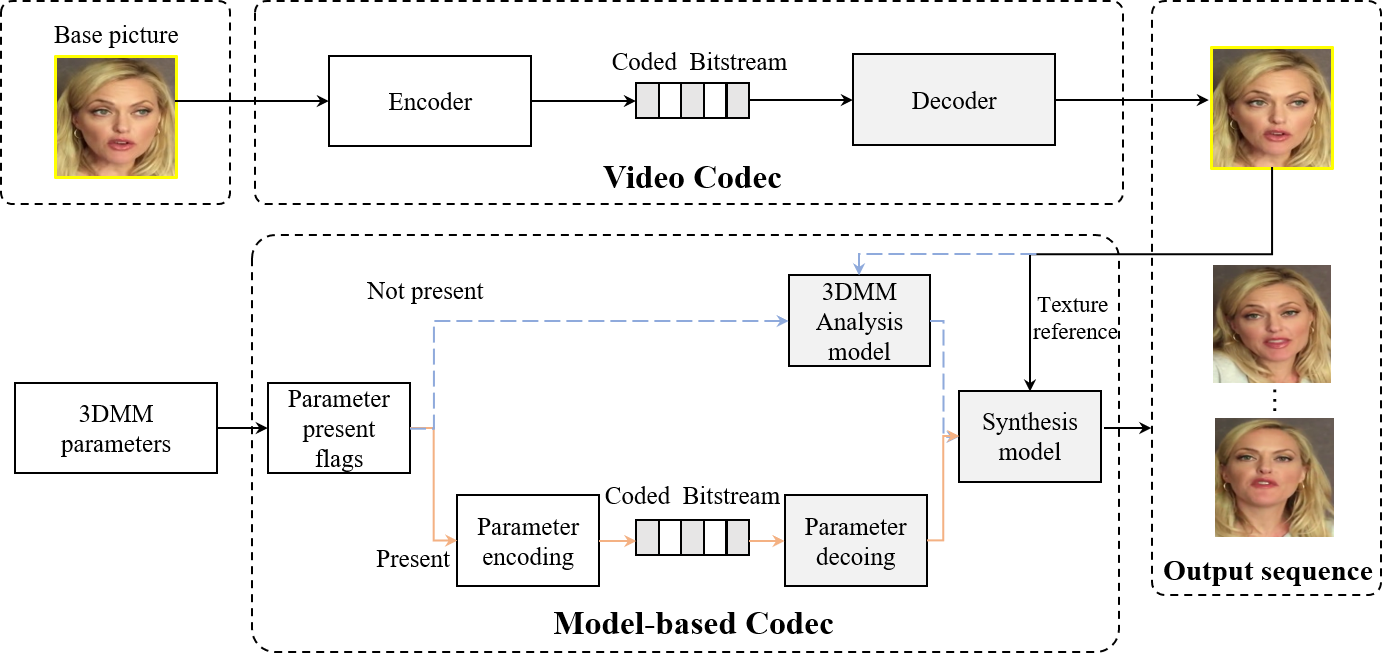
Decision: Adopt JVET-AC0058 (for JVET-AC2027), for Aspect 2. Include an editors’ note that this should be considered to be extended to other types of SEI with prefix information when two of them appear in sequence.

[JVET-AC0077](https://jvet-experts.org/doc_end_user/current_document.php?id=12279) AHG9/AHG15: On bitstreams that are potentially suboptimal for user viewing [M. M. Hannuksela, F. Cricri, H. Zhang (Nokia)]

See under 4.14

[JVET-AC0088](https://jvet-experts.org/doc_end_user/current_document.php?id=12290) AHG9: Generative Face Video SEI Message [B. Chen, J. Chen, S. Wang, Y. Ye (Alibaba), S. Wang (CityU HK)]

This contribution proposes for a generative face video SEI messages. The proposed SEI has two functions: one is to reconstruct high-quality talking face video at ultra-low bitrate and the other is to manipulate the talking face video towards personalized characterization. As such, the proposed SEI is applicable to video conferencing, live entertainment, face animation and metaverse-related functionalities.



1. *Flowchart of generative face video SEI message*

One SEI message would be needed for each time instance that is synthesized. The concept is to send the SEI with the picture from which the other pictures are generated (including the number of them).

Has some relationship with concepts from NNPF temporal picture upsampling, however using extrapolation. This might introduce some problem in real-time applications, unless sending one SEI per time step (which could be done associating them with dummy pictures). This would cost some overhead, how much?

A demo was given, synthesis is performed by a GAN. Size of model? Cannot be answered.

Animating another person’s face was also shown as example.

Synthesis speed 250 frames/minute, 256x256.

It was asked how this would compare to audio-animated face videos.

How much rate used for video, how much for the SEI?

Several experts expressed this is a very interesting approach. Further study recommended.

[JVET-AC0122](https://jvet-experts.org/doc_end_user/current_document.php?id=12324) AHG9: Attenuation Map Information SEI for reducing energy consumption of displays [C.-H. Demarty, F. Aumont, E. Reinhard, L. Blondé, O. Le Meur, Z. Ameur, E. François (InterDigital)]

The energy consumption of a display represents a large part of the total energy consumption when consuming a video content.

Light emitting display technologies allow for controlling the consumption at the pixel level through a pixel-wise reduction of the luminance.

This contribution proposes the signaling of new metadata for enabling the attenuation of each sample’s values (of each color component) of a decoded video at the receiver side. The reported benefits are to allow an intelligent reduction of each pixel luminance value individually so as to reduce the energy consumption, while maintaining or controlling the quality of experience for the end user.

This contribution proposes to transmit the Attenuation Map through an additional type of auxiliary picture. Its further use at the player or display side is guided by the addition of a new SEI message.

It was commented that a comparison with global dimming of backlight (as per green metadata) should be done, to justify the necessity of control at pixel level.

It was also commented that the subjective impact would be more relevant than PSNR. This should be compared with the global dimming at same level of energy saving.

How much additional rate for sending the attenuation map? Is it necessary to send this at the same frame rate as the video? The additional bit rate and decoding also consume power, this should be taken into consideration as well.

Is is independent from display size? Is it not also dependent on local display settings and user behaviour?

It is pointed out that many displays (in particular mobile) already adapt to ambient lighting, and there may also be dependency.

It should be investigated if it would not be possible using the existing alpha map purpose of auxiliary pictures, and specify this in the green metadata spec. It might be desirable to have this in the same place as the global attenuation SEI, might even be possible to combine both (not every display has a local attenuation mechanism).

Further study.

[JVET-AC0141](https://jvet-experts.org/doc_end_user/current_document.php?id=12343) AHG9: Alternative Picture Timing SEI [H.-B. Teo, J. Gao, C.-S. Lim, K. Abe (Panasonic)]

This contribution proposes an Alternative Picture Timing SEI message. In the case if the video bitstream contains slow-motion scenes, this SEI can indicate the actual motion speed of the content at captured time. Two aspects are included. In the first aspect, a time scale factor is introduced. This time scale factor is used to indicate the time scale difference between the Picture Timing SEI and the Alternative Picture Timing SEI. In the second aspect, level\_idc may be added for every temporal sublayer. Second aspect is to allow update of the level\_idc if the decoding level needs to be changed based on the time scale factor. The proposal suggests to adopt aspect 1 and aspect 2 into VVC specification.

The proposal intends that the SEI message takes influence on the output time of VVC decoding.

It was asked if this might lead to the consequence of exceeding maximum framerate of a level. This would only apply to aspect 1, for aspect 2 a lower temporal laye could be decoded.

What would be the application? For example, machine analysis, or user-controlled selection of speed.

It was pointed out that applications like this are already existing.

I was asked if it is necessary to change the decoder output behaviour to achieve this.

Generally, it might be beneficial to know what scale factor is used, e.g. if in a video sequence only one part is contained in “slow motion” mode, as often in sports events.

Further study recommended.

## Film grain synthesis (4)

Contributions in this area were discussed in session 13 at 1440–1510 UTC and session 14 at 1530-1600 on Monday 16 Jan. 2023 (chaired by JRO).

[JVET-AC0043](https://jvet-experts.org/doc_end_user/current_document.php?id=12228) AHG4: Report on AHG meeting on subjective test preparation for film grain synthesis [M. Wien (RWTH)]

[JVET-AC0151](https://jvet-experts.org/doc_end_user/current_document.php?id=12353) AHG13: Proposed text: Film grain synthesis technology for video applications (CD draft) [D. Grois (Comcast), Y. He (Qualcomm), W. Husak (Dolby), P. de Lagrange (InterDigital), A. Norkin (Netflix), M. Radosavljević (Xiaomi), A. Tourapis (Apple), W. Wan (Broadcom)]

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.

Software related to the report is integrated in VTM and HM. It was suggested to refer more precisely to parameter settings of the software.

It was commented that the “artistic” overlay of film grain (e.g. for hiding coding artifacts) is not described practically.

It was pointed that some language is not appropriate for a technical report (e.g. avoid “guidance”, “shall”, “should”, etc.). Section 2 should only contain normative references, all other references should be moved to biography. Avoid “may” in mentioning T.35, and describe the code point that has been registered for AFGS1. See ISO directives part2 and ISO house style documents.

Issue CDTR.

[JVET-AC0181](https://jvet-experts.org/doc_end_user/current_document.php?id=12385) AHG4: experiments in preparation of film grain visual tests [P. de Lagrange (InterDigital)]

[JVET-AC0199](https://jvet-experts.org/doc_end_user/current_document.php?id=12403) AHG13: On Film Grain Synthesis Subjective Evaluation [X. Meng, W. Zhang, S. Labrozzi (Disney Streaming)] [late]

Initial version rejected as “placeholder”.

The subjective evaluation of film grain synthesis is recommended in the last JVET meeting. This document provides the film grain synthesis evaluation results of both the frequency-based method and the auto-regressive based method. The main observations are as follows,

1. In the VTM-19.0, the film grain synthesis (FGS) denoiser doesn’t change the input of the encoder.
   1. A typical (FGS) framework includes denoising, noise modeling and grain synthesis. In the VTM-19.0, the effect of bitrate saving can’t be evaluated.
2. The level of synthesized film grain changes regularly in the temporal domain.
   1. The film grain intensity ramps up every fourth frame. The main reason may be the MCTF and QP adjustment.
3. For videos with relative heavy film grain,
   1. The synthesized grain is coarse-grained, and there are obvious repetitive patterns. Parameters are recommended to control the granularity of synthesized film grain.
   2. The FGS is turned off in some frames with heavy grain. This may due to the denoiser or noise modeling method. More test videos are recommended to cover different content and film grain level.
4. For videos with low film grain, the synthesized grain looks fine.
5. The frequency based FGS method and auto-regressive (AR) based method show different results under different circumstances.
   1. How to select the proper method for a film grain test video is important.

Results are reported with several sequences which are currently not available publicly.

It is commented that the current analysis may not be optimized for any material. The MCTF-based denoising may also not be optimum. Some of these aspects might be appearing with material that has stronger film grain.

A recommendation should be made requesting for more film grain material

It was recommended also having clean content as ground truth.

Some concern is expressed if the TR is ready to go. However, this is more about further improving the software whereas the general concepts described in the TR are probably OK. It is not the intent of reference software to be a market product.

See further discussion under 4.4

## Non-SEI HLS aspects (0)

Contributions in this area were discussed in session X at XXXX–XXXX UTC on XXday X Jan. 2023 (chaired by XXX).

# 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.

Communication and coordination items discussed in session 13, Monday Jan. 16, 1300 UTC

* Joint meetings:
  + With AG5 on visual testing: Mon 01-13 1620 UTC (1 hr.)
  + With WG2, AG5 and VCEG on JVET exploration status: Tue 1300 UTC (1 hr.)
  + With WG4 on VCM and Lenslet: Wed. 1400 UTC (1 hr.)
* Session planning (see calendar)
* Antalya meeting: “Early bird” registration deadline extended until Friday 01-20
* Preparation of DoCRs: HEVC 5th ed. DIS, VVC conformance DAM1, VVC DAM1 (draft DoCR and preliminary text of new edition to be produced)
* WG2 plans to issue a TR on “MPEG AI” – volunteers sought

General plenary wrap-up discussions are recorded under sections 8, 9, and 10.

## Information sharing meetings

Information sharing sessions with other WGs and AGs of the MPEG community were held on Monday 16 Jan. 0900–1130, Wednesday 18 Jan. 0900–1000, and Friday 20 Jan. 2100–2300 UTC. The status of the work in the MPEG WGs and AGs was reviewed at these information sharing sessions.

Add joint meetings with AG 5 referencing notes in 4.4

## Joint meeting with WG2, AG5, and VCEG on JVET exploration status Tuesday 17 Jan. 13:00-XX:XX UTC

Chaired by Igor Curcio, Jens Rainer Ohm, Gary Sullivan and Mathias Wien

Status summary slides were shown by JRO and MW.

NNVC exploration summary:

* NNVC loop filtering showing more than 10% for loop filtering, especially for chroma
* Post-filtering is also showing gain
* NNVC super-resolution is showing average 2-3% gain.
* NNVC intra coding also showing gain
* NNVC result from JVET-AC0014 showing ~9.5% RA PSNR gain in a test, with about ~36× decoder CPU runtimes

ECM exploration study:

* See AHG 6, AHG 12 and EE2 reports, tables from JVET-AC0006.
* ECM shows ~19% PSNR RA gain over VVC, ~6× encode, 8× decode runtimes, typically 1% improvement per meeting cycle (slower than during HEVC & VVC development)
* See JVET-AB2029 for a subjective testing report, UHD and HD tests were performed
* ECM subjective gain in the neighbourhood of 38% average for UHD and 32-33% average for HD was shown in laboratory testing
* Coding tool modifications include intra & inter prediction, SCC tools, transforms, loop filters, not so much partitioning changes
* Subjectively, VVC seems to tend toward blur, while ECM seems to add detail

Some comments from discussion

* At the time when VVC was beginning, the JEM had ~29% PSNR over HEVC
* Subjectively, JEM vs HEVC had been evaluated in JVET-E0129 (using expert viewing, perhaps somewhat less rigorous than the testing performed now for ECM vs. VVC), that contribution says: "Confirm the objective quality advantage (25%-30%) shown by the JEM codec with respect to HM")
* We are closer in time to VVCv1 than the JEM was for HEVCv1
* Of course, some of the tools investigated in the ECM don’t seem very practical, but that was also the case for some things investigated in the days of JEM
* VVC is still relatively early in its deployment process

Potential actions:

* Continued study along the current paths?
* Issue Call for Evidence or Call for Proposals?
* Extension of VVC as additional profile(s)?
* Developing new standard(s)?

It was agreed that it does not yet seem time to start a new project; continued study in the current exploration mode in JVET seems best. This had also been the expressed position from VCEG.

It was commented that additional analysis of practicality would be beneficial in the explorations (e.g., implementable ~5 years in the future). “Micro-optimization” either for coding efficiency or practicality was noted not to be so helpful. The previous AHG 13 on/off testing was noted along that line (not too fine in granularity – see JVET-AC0171 as an example). It was noted that runtime alone is not really a good characterization of complexity – e.g. when considering template matching technology. Analysis of techniques along both investigation paths is also appreciated.

## Joint meeting with WG4 and VCEG on VCM coordination and lenslet video coding Wednesday 18 Jan. 14:00-15:25 UTC

On VCM, there was discussion of metadata development coordination relating to VCM.

* It was noted that the annotated regions (AR) SEI message already includes an indicator to indicate a machine analysis purpose.
* Some machine analysis indication has also been proposed for the NNPF SEI message.

The following approach was suggested and agreed for current coordination purposes:

* If the purpose of the metadata syntax is primarily for VCM, it would be developed in WG 4 and likely handled in video bitstreams (if needed there) by reference, e.g. similar to the green metadata approach for carriage, where a video spec or the VSEI spec would reference a VCM standard.
* For broader-purpose metadata (e.g. the NNPF or AR SEI messages) that is not developed specifically for VCM but might have some small VCM-related aspect (such as a purpose indicator), the small VCM-related aspect could be developed in JVET and specified e.g. in VSEI. A VCM standard could also be developed by WG 4 that further expounds on usage of such data for VCM purposes.

CTC alignment with WG 4 for VCM was then discussed. See additional notes in section 4.14 under JVET-AC0073.

Work in WG 4 on lenslet video coding was also presented. Studies of potential low-level coding tools added or modified in the VVC context for customization for this purpose were reportedly showing some gain (e.g. in the range of 2-7% for plenoptic content). However, it was commented that special purpose extension of VVC could dilute the concept of that standard applying in a versatile manner across a broad variety of applications and that historically, low-level changes associated with profiles for special applications have typically not been widely implemented. It was suggested that the market for lenslet-customized coding may not be sufficiently large to justify special coding tools / profiles for it. It was noted that exploration of more general-purpose modifications also shows benefit (e.g. using the ECM can probably provide more gain over VVC than what was reported for purpose-specific customization). See additional notes in section 4.15 under JVET-AC0334.

## BoGs (2)

The following break-out groups were established at this meeting to conduct discussion and develop recommendations on particular subjects

[JVET-AC0324](https://jvet-experts.org/doc_end_user/current_document.php?id=12529) BoG report on neural-network based post-filtering contributions [G. J. Sullivan (BoG coordinator)]

[JVET-AC0351](https://jvet-experts.org/doc_end_user/current_document.php?id=12556) BoG report on non-normative optimization for machine [C. Hollmann, S. Liu (BoG coordinators)]

* A BoG on …. See section XX and the report JVET-AC0XXX for details.

## Liaison communications

[m62300](https://dms.mpeg.expert/doc_end_user/current_document.php?id=84196&id_meeting=192) Liaison statement from SC 29/WG 1 to WG 2, WG 4, and WG 5 on JPEG AI [WG 1 via SC 29 Secretariat]

A reply was drafted by JVET as WG 5 N 191. This was reviewed in session 24.

# Project planning

## Software timeline

ECM8 software (including all adoptions) was planned to be available 3 weeks after the meeting.

The NNVC 4.0 codebase software was planned to be available 4 weeks after the meeting.

VTM20.0 software was planned to be available on 2022-02-17. (Note that further updates may be released later)

Updates on top of HM17.0 software were not planned, but might be released after merging pending requests, as appropriate.

## Core experiment and exploration experiment planning

An EE on neural network-based video coding was established, as recorded in output document JVET-AC2023.

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-AC2024.

Initial versions of these documents were presented and approved in the first plenary session on Friday 20 January.

## Drafting of specification text, encoder algorithm descriptions, and software

The following agreement has been established: the editorial team has the discretion to not integrate recorded adoptions for which the available text is grossly inadequate (and cannot be fixed with a reasonable degree of effort), if such a situation hypothetically arises. In such an event, the text would record the intent expressed by the committee without including a full integration of the available inadequate text.

## Plans for improved efficiency and contribution consideration

The group considered it important to have the full design of proposals documented to enable proper study.

Adoptions need to be based on properly drafted working draft text (on normative elements) and HM/VTM encoder algorithm descriptions – relative to the existing drafts. Proposal contributions should also provide a software implementation (or at least such software should be made available for study and testing by other participants at the meeting, and software must be made available to cross-checkers in EEs).

Suggestions for future meetings included the following generally-supported principles:

* No review of normative contributions without draft specification text
* VTM algorithm description text is strongly encouraged for non-normative contributions
* Early upload deadline to enable substantial study prior to the meeting
* Using a clock timer to ensure efficient proposal presentations (5 min) and discussions

As general guidance, it was suggested to avoid usage of company names in document titles, software modules etc., and not to describe a technology by using a company name.

## General issues for experiments

It was emphasized that those rules which had been set up or refined during the 12th JVET meeting should be observed. In particular, for some CEs of some previous meetings, results were available late, and some changes in the experimental setup had not been sufficiently discussed on the JVET reflector.

Group coordinated experiments have been planned as follows:

* “Core experiments” (CEs) are the coordinated experiments on coding tools which are deemed to be interesting but require more investigation and could potentially become part of a draft standard by the next meeting or in the near future.
* “Exploration experiments” (EEs) are also coordinated experiments. These are conducted on technology which is not foreseen to become part of a draft standard in the near future. The investigating methodology for assessment of such technology can also be an important part of an EE. (Further general rules for EEs, as far as deviating from the CE rules below, should be discussed in a future meeting. For the current meeting, procedures as described in the EE description document are deemed to be sufficient.)
* A CE is a test of a specific fully described technology in a specific agreed way. It is not a forum for thinking of new ideas (like an AHG). The CE coordinators are responsible for making sure that the CE description is complete and correct and has adequate detail. Reflector discussions about CE description clarity and other aspects of CE plans are encouraged.
* A description of each experiment is to be approved at the meeting at which the experiment plan is established. This should include the issues that were raised by other experts when the tool was presented, e.g., interference with other tools, contribution of different elements that are part of a package, etc. The experiment description document should provide the names of individual people, not just company names.
* Software for tools investigated in a CE will be provided in one or more separate branches of the software repository. Each CE will have a “fork” of the software, and within the CE there may be multiple branches established by the CE coordinator. The software coordinator will help coordinate the creation of these forks and branches and their naming. All JVET members will have read access to the CE software branches (using shared read-only credentials as described below).
* During the experiment, revisions of the experiment plans can be made, but not substantial changes to the proposed technology.
* The CE description must match the CE testing that is done. The CE description needs to be revised if there has been some change of plans.
* The CE summary report must describe any changes that were made in the process of finalizing the CE.
* By the next meeting it is expected that at least one independent cross-checker will report a detailed analysis of each proposed feature that has been tested and confirm that the implementation is correct. Commentary on the potential benefits and disadvantages of the proposed technology in cross-checking reports is highly encouraged. Having multiple cross-checking reports is also highly encouraged (especially if the cross-checking involves more than confirmation of correct test results). The reports of cross-checking activities may (and generally should) be integrated into the CE report rather than submitted as separate documents.
* It is mandatory to report encoder optimizations made for the benefit of a tool, and if an equivalent optimization could be applied on the anchor, a comparison against the improved anchor shall be provided.
* A new proposal can be included in a CE based on group decision, regardless if an independent party has already performed a cross-check in the meeting when it was first proposed.

It is possible to define sub-experiments within particular CEs, for example designated as CEX.a, CEX.b, etc., where X is the basic CE number.

As a general rule, it was agreed that each CE should be run under the same testing conditions using one software codebase, which should be based on the group test model software codebase. An experiment is not to be established as a CE unless there is access given to the participants in (any part of) the CE to the software used to perform the experiments.

The general agreed common conditions for single-layer coding efficiency experiments for SDR video are described in the prior output document JVET-T2010.

Experiment descriptions should be written in a way such that it is understood as a JVET output document (written from an objective “third party perspective”, not a proponent perspective – e.g. not referring to methods as “improved”, “optimized”, “enhanced”, etc.). The experiment descriptions should generally not express opinions or suggest conclusions – rather, they should just describe what technology will be tested, how it will be tested, who will participate, etc. Responsibilities for contributions to CE work should identify individuals in addition to company names.

CE descriptions contain a basic description of the technology under test, but should not contain excessively verbose descriptions of a technology (at least not unless the technology is not adequately documented elsewhere). Instead, the CE descriptions should refer to the relevant proposal contributions for any necessary further detail. However, the complete detail of what technology will be tested must be available – either in the CE description itself or in documents that are referenced in the CE description that are also available in the JVET document archive.

Any technology must have at least one cross-check partner to establish a CE – a single proponent is not enough. It is highly desirable have more than just one proponent and one cross-checker.

The CE development workflow is described at:

<https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/wikis/Core-experiment-development-workflow>

CE read access is available using shared accounts: One account exists for MPEG members, which uses the usual MPEG account data. A second account exists for VCEG members with account information available in the TIES informal ftp area (IFA) system at:

<https://www.itu.int/ifa/t/2017/sg16/exchange/wp3/q06/vceg_account.txt>

Some agreements relating to CE activities were established as follows:

* Only qualified JVET members can participate in a CE.
* Participation in a CE is possible without a commitment of submitting an input document to the next meeting. Participation is requested by contacting the CE coordinator.
* All software, results, and documents produced in the CE should be announced and made available to JVET in a timely manner.
* A JVET CE reflector will be established and announced on the main JVET reflector. Discussion of logistics arrangements, exchange of data, minor refinement of the test plans, and preparation of documents shall be conducted on the JVET CE reflector, with subject lines prefixed by “[CEx: ]”, where “x” is the number of the CE. All substantial communications about a CE other than such details shall take place on main JVET reflector. In the case that large amounts of data are to be distributed, it is recommended to send a link to the data rather than the data itself, or upload the data as an input contribution to the next meeting.

General timeline for CEs

T1= 3 weeks after the JVET meeting: To revise the CE description and refine questions to be answered. Questions should be discussed and agreed on JVET reflector. Any changes of planned tests after this time need to be announced and discussed on the JVET reflector. Initially assigned description numbers shall not be changed later. If a test is skipped, it is to be marked as “withdrawn”.

T2 = Test model software release + 2 weeks: Integration of all tools into a separate CE branch of the VTM is completed and announced to JVET reflector.

* Initial study by cross-checkers can begin.
* Proponents may continue to modify the software in this branch until T3.
* 3rd parties are encouraged to study and make contributions to the next meeting with proposed changes

T3: 3 weeks before the next JVET meeting or T2 + 1 week, whichever is later: Any changes to the CE test branches of the software must be frozen, so the cross-checkers can know exactly what they are cross-checking. A software version tag should be created at this time. The name of the cross-checkers and list of specific tests for each tool under study in the CE plan description shall be documented in an updated CE description by this time.

T4: Regular document deadline minus 1 week: CE contribution documents including specification text and complete test results shall be uploaded to the JVET document repository (particularly for proposals targeting to be promoted to the draft standard at the next meeting).

The CE summary reports shall be available by the regular contribution deadline. This shall include documentation about crosscheck of software, matching of CE description and confirmation of the appropriateness of the text change, as well as sufficient crosscheck results to create evidence about correctness (crosscheckers must send this information to the CE coordinator at least 3 days ahead of the document deadline). Furthermore, any deviations from the timelines above shall be documented. The numbers used in the summary report shall not be changed relative to the description document.

CE reports may contain additional information about tests of straightforward combinations of the identified technologies. Such supplemental testing needs to be clearly identified in the report if it was not part of the CE plan.

New branches may be created which combine two or more tools included in the CE document or the VTM (as applicable).

It is not necessary to formally name cross-checkers in the initial version of the CE description document. To adopt a proposed feature at the next meeting, JVET would like to see comprehensive cross-checking done, with analysis of whether the description matches the software, and a recommendation of the value of the tool and given tradeoffs.

The establishment of a CE does not indicate that a proposed technology is mature for adoption or that the testing conducted in the CE is fully adequate for assessing the merits of the technology, and a favourable outcome of CE does not indicate a need for adoption of the technology into a standard or test model.

Availability of specification text is important to have a detailed understanding of the technology and also to judge what its impact on the complexity of the specification will be. There must also be sufficient time to study this in detail. CE contributions without sufficiently mature draft specification text in the CE input document should not be considered for adoption.

Lists of participants in CE documents should be pruned to include only the active participants. Read access to software will be available to all members.

# Establishment of ad hoc groups

The ad hoc groups established to progress work on particular subject areas until the next meeting are described in the table below. The discussion list for all of these ad hoc groups was agreed to be the main JVET reflector ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de)).

Review of AHG plans was conducted during the plenary on Friday 20 January 2023 at XXXX 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-AC1008, JVET-AC2002, JVET-AC2005, JVET-AC2027, and JVET-AC2032). * 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-AC1004 errata output collection. * Coordinate with the test model software development AhG to address issues relating to mismatches between software and text. * Collect and consider errata reports on the texts. | B. Bross, C. Rosewarne (co-chairs), F. Bossen, A. Browne, S. Kim, S. Liu, J.‑R. Ohm, G. J. Sullivan, A. Tourapis, Y.-K. Wang, Y. Ye (vice-chairs) | N |
| **Test model software development (AHG3)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Coordinate development of test models (VTM, HM, SCM, SHM, HTM, MFC, MFCD, JM, JSVM, JMVM, 3DV-ATM, 360Lib, and HDRTools) software and associated configuration files. * Produce documentation of software usage for distribution with the software. * Enable software support for recently standardized additional SEI messages. * Discuss and make recommendations on the software development process. * Propose improvements to the guideline documents for developments of the test model software. * Perform comparative tests of test model behaviour using common test conditions, including HDR, high bit depth and high bit rate. * Suggest configuration files for additional testing of tools. * Investigate how to minimize the number of separate codebases maintained for group reference software. * Coordinate with AHG on Draft text and test model algorithm description editing (AHG2) to identify any mismatches between software and text, and make further updates and cleanups to the software as appropriate. * Prepare drafts of merged and updated CTC documents for HM and VTM, as applicable. | F. Bossen, X. Li, K. Sühring (co-chairs), E. François, Y. He, K. Sharman, V. Seregin, A. Tourapis (vice‑chairs) | N |
| **Test material and visual assessment (AHG4)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Consider plans for additional verification testing of VVC capability, particularly target 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, in coordination with other AHGs, as appropriate. * Maintain and update the directory structure for the test sequence repository, as necessary. * Collect information about test sequences that have been made available by other organizations. * Prepare and conduct 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))   * Produce and finalize the draft conformance testing for operation range extensions JVET-AC2026 and the additional conformance bitstreams for VVC multilayer configurations JVET-AC2028, and investigate the need for future improvements of conformance testing specifications. * Study the requirements of VVC, HEVC, and AVC conformance testing to ensure interoperability. * Maintain and update the conformance bitstream database, and contribute to report problems, and suggest actions to resolve these. * Study additional testing methodologies to fulfil the needs for VVC conformance testing. | D. Rusanovskyy, I. Moccagatta (co-chairs), F. Bossen, K. Kawamura, T. Ikai, 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-8.0 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) |
| **ECM tool assessment (AHG7)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Discuss JVET-AC0354, and prepare an updated document as input to the next meeting on the methodology of tool assessment, tool grouping and a list of tools in each group to be tested. * Investigate missing tool controls in ECM, and suggest appropriate modifications to software in coordination with AHG6. * Prepare configuration files and generate bitstreams and results of tool-on/tool-off testing. * Prepare reporting of tool assessment results. | X. Li (chair), L.-F. Chen, Z. Deng, J. Gan, E. François, H.-J. Jhu, R.-L. Liao, H. Wang (vice-chairs) | Y (tel., 2 weeks notice) |
| **Optimization of encoders and receiving systems for machine analysis of coded video content (AHG8)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Solicit and study non-normative encoder and receiving systems technologies that enhance performance of machine analysis tasks on coded video content. * Identify and collect test materials that are suitable to be used by JVET for machine analysis tasks. * Develop common test conditions and generate anchors. * Develop an evaluation framework, including evaluation procedures and methodologies. * Coordinate software development, and investigate the possibility of migrating the software basis to newest VTM version. * Coordinate experiments on optimization of encoders and receiving systems for machine analysis of coded video content. * 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 WG 4 VCM AHG on common interests and activities such as common test conditions, test and training materials, and on studying characteristics and requirements of targeted machine analysis tasks, etc. | C. Hollmann, S. Liu, S. Wang, M. Zhou (AHG chairs) | 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 AHG8 and WG 4 to study mechanisms for signalling metadata in the context of machine analysis of coded video content. * Coordinate with AHG3 for software support of SEI messages. | S. McCarthy, Y.-K. Wang (co-chairs), T. Chujoh, S. Deshpande, C. Fogg, Hendry, P. de Lagrange, G. J. Sullivan, A. Tourapis, S. Wenger (vice-chairs) | N |
| **Encoding algorithm optimization (AHG10)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the impact of using techniques such as tool adaptation and configuration, and perceptually optimized adaptive quantization for encoder optimization. * Study the impact of non-normative techniques of preprocessing for the benefit of encoder optimization. * Study encoding techniques of optimization for objective quality metrics and their relationship to subjective quality. * Study optimized encoding for reference picture resampling and scalability modes in VTM. * Study optimized encoding and tool combinations for low latency and low complexity. * Consider neural network-based encoding optimization technologies for video coding standards. * Investigate other methods of improving objective and/or subjective quality, including adaptive coding structures and multi-pass encoding. * Study methods of rate control and rate-distortion optimization and their impact on performance, subjective and objective quality. * Study the potential of defining default or alternate software configuration settings and test conditions optimized for either subjective quality, or higher objective quality, and coordinate such efforts with AHG3 and AHG6. * Study the effect of varying configuration parameters depending on temporal layer, such as those related to deblocking, partitioning, chroma QP. | P. de Lagrange, A. Duenas, R. Sjöberg, A. Tourapis (AHG chairs) | N |
| **Neural network-based video coding (AHG11)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Evaluate and quantify the performance improvement potential of NN-based video coding technologies compared to existing video coding standards such as VVC, including both individual coding tools and novel architectures. * Discuss potential refinements of the test conditions for NN-based video coding in JVET-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. If possible, prepare encodings with combinations of tools included in the NNVC software for visual quality assessment at the next meeting. * Coordinate with AHG14 on items related to NNVC software development, and study the potential for unification of filter sets #0 and #1. | E. Alshina, S. Liu, A. Segall (co‑chairs), F. Galpin, J. Li, R.-L. Liao, D. Rusanovskyy, T. Shao, M. Wien, P. Wu (vice‑chairs) | Y (tel., 2 weeks notice) |
| **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 ECM8 algorithm description JVET-AC2025. * Coordinate with AHG7 to study the performance and complexity tradeoff of these video coding tools. * Coordinate with AHG6 on ECM software development. * Support AHG6 in generating anchors according to the test conditions in JVET-Y2017. * Analyse the results of exploration experiments described in JVET-AC2024 in coordination with the EE coordinators. * Coordinate with AHG11 to study the interaction with neural network-based coding tools. | M. Karczewicz, Y. Ye, L. Zhang (co-chairs), B. Bross, R. Chernyak, X. Li, K. Naser, H. Yang (vice-chairs) | Y (tel., 2 weeks notice) |
| **Film grain technologies (AHG13)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the benefits and characteristics of film grain technologies, including autoregressive and frequency-filtering technologies. * Produce and finalize the JVET-AC2020 draft of the Technical Report on Film grain synthesis technology for video applications and suggest future improvements as necessary. * Study alternative film grain models and their associated documentation. * Coordinate with AHG4 on soliciting additional test materials as needed for further investigations on film grain technologies. * Study preprocessing and encoder technologies for determining values for FGC (Film Grain Characteristics) SEI message syntax elements. * Identify potential need for additional film grain technology and signalling, if needed. * Coordinate development of film grain technology software and configuration files. * Coordinate with AG 5 on finalizing the draft plan for subjective quality testing of the FGC SEI message JVET-AC2022, 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, W. Zhang (vice-chairs) | Y (tel., 2 weeks notice) |
| **NNVC software development (AHG14)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Coordinate development of the NNVC software and associated configuration files. * Prepare and deliver NNVC-4.0 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-AC2019) editors to identify any mismatches between software and description document, suggest further updates to the description document as appropriate. * Investigate combinations of tools included in the NNVC software and coordinate with AHG11 on defining default settings of such combinations. * Develop software guidelines for NNVC and SADL. * Coordinate with AHG11 on items related to NNVC activities. | S. Eadie, F. Galpin, Y. Li, L. Wang, Z. Xie (AHG 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 192) 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 192, as noted in section 9.

[JVET-AC1000](https://jvet-experts.org/doc_end_user/current_document.php?id=12210) Meeting Report of the 29th JVET Meeting [J.-R. Ohm] [WG 5 N 174] (2023-02-17)

Initial versions of the meeting notes (d0 … d9) were made available on a daily basis during the meeting.

[JVET-AC1001](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=5095) Guidelines for HM-based software development [K. Sühring, F. Bossen, X. Li (software coordinators)] (2023-02-03)

Developed from JVET-AC0204

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]

[JVET-AC1004](https://jvet-experts.org/doc_end_user/current_document.php?id=12211) Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR [Y.-K. Wang, B. Bross, I. Moccagatta, C. Rosewarne, G. J. Sullivan] (2023-03-17, near next meeting)

New aspects included the issues pointed out in JVET-AC0311 and JVET-AC0346.

Remains valid – not updated: [JVET-Z1005](https://jvet-experts.org/doc_end_user/current_document.php?id=11707) New levels for HEVC (Draft 3) [T. Suzuki, A. Tourapis, Y.-K. Wang]

The content of this document will be included in a new edition of HEVC. For additional corrigenda items included, see documents JVET-AB1004, as well as notes under JVET-AC0311 and JVET-AC0346. A DoCR on ISO/IEC DIS23008-2 was submitted as WG 5 N178 (reviewed during session 24), and the FDIS was submitted as WG 5 N 179.

JVET-Z1005 can be removed after publication of the new edition of ISO/IEC 23008-2.)

No output: JVET-Axx1006

Remains valid – not updated: [JCTVC-V1007](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=10312) SHVC Test Model 11 (SHM 11) Introduction and Encoder Description [G. Barroux, J. Boyce, J. Chen, M. M. Hannuksela, Y. Ye] [WG 11 N 15778]

[JVET](https://jvet-experts.org/doc_end_user/current_document.php?id=12212)-AC1008 Additional colour type identifiers for AVC, HEVC and Video CICP (Draft 3) [G. J. Sullivan, W. Husak, A. Tourapis] [WG 5 Preliminary WD N186] (2023-02-24)

Modified from JVET-AC0352

[JCTVC-AC1009](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=10572) Common Test Conditions for SHVC [K. Sühring] (2023-02-17)

This requires an update, as the previous version referred to an outdated location of test sequences.

Remains valid – not updated [JCTVC-O1010](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=8511) Guidelines for Conformance Testing Bitstream Preparation [T. Suzuki, W. Wan]

Remains valid – not updated: [JVET-AA1011](https://jvet-experts.org/doc_end_user/current_document.php?id=11944) HEVC multiview profiles supporting extended bit depth (draft 1) [A. Tourapis, W. Husak] [WG 5 preliminary WD N 143] (2022-08-26)

Proponents of the new profiles reported that they are currently developing software and conformance bitstreams.

Remains valid – not updated: JVET-[AB1012](https://jvet-experts.org/doc_end_user/current_document.php?id=12213) Overview of IT systems used in JVET [J.-R. Ohm, I. Moccagatta, K. Sühring, M. Wien]

[JVET-AC1013](http://phenix.int-evry.fr/jct3v/doc_end_user/current_document.php?id=1878) Common Test Conditions of 3DV Experiments [K. Sühring] (2023-02-17)

This requires an update, as the previous version referred to an outdated location of test sequences.

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]

[JCTVC-AC1015](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=10689) Common Test Conditions for Screen Content Coding [K. Sühring] (2023-02-17)

This requires an update, as the previous version referred to an outdated location of test sequences.

Remains valid – not updated: [JCT3V-G1003](http://phenix.int-evry.fr/jct3v/doc_end_user/current_document.php?id=1884) 3D-AVC Test Model 9 [ D. Rusanovskyy, F. C. Chen, L. Zhang, T. Suzuki ] [WG 11 N 14239]

Remains valid – not updated: [JCT3V-K1003](http://phenix.int-evry.fr/jct3v/doc_end_user/current_document.php?id=2499) Test Model 11 of 3D-HEVC and MV-HEVC [Y. Chen, G. Tech, K. Wegner, S. Yea] [WG 11 N 15141]

No output: JVET-Axx1016 through JVET-Axx1099

Remains valid – not updated [JVET-AA1100](https://jvet-experts.org/doc_end_user/current_document.php?id=11944) Common Test Conditions for HM Video Coding Experiments [K. Sühring, K. Sharman]

This specifies only the CTC for non-4:2:0 colour formats. The corresponding document for VVC is JVET-T2013, with no unification yet.

No output: JVET-Axx2001

[JVET-AC2002](https://jvet-experts.org/doc_end_user/current_document.php?id=12214) Algorithm description for Versatile Video Coding and Test Model 19 (VTM 19) [A. Browne, Y. Ye, S. Kim] [WG 5 N 183] (2023-04-14, near next meeting)

Updates from JVET-AC0096 (description of RPR switching control).

It is noted that the list above may not be complete; if some adoption is missing that is recorded somewhere else in the meeting notes it shall also be considered included.

[JVET-AC2003](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=6638) Guidelines for VTM-based software development [F. Bossen, X. Li, K. Sühring] (2023-02-03)

Developed from JVET-AC0166.

Remains valid – not updated: [JVET-T2004](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10542) Algorithm descriptions of projection format conversion and video quality metrics in 360Lib (Version 12) [Y. Ye, J. Boyce]

[JVET-AC2005](https://jvet-experts.org/doc_end_user/current_document.php?id=11946) New level and systems-related supplemental enhancement information for VVC (Draft 4) [E. François, B. Bross, M. Hannuksela, A. Tourapis, Y.-K. Wang] (2023-03-31)

JVET-AC2005 is an update of JVET-AA2005 (previously submitted as WG 5 DAM 1 N 145). For elements included, see documents JVET-AB1004, as well as notes under JVET-AC0311 and JVET-AC0346.

The content of this document was included in a new edition of VVC. A draft DoCR on ISO/IEC 23090-3/DAM1 was submitted as WG 5 N180 (reviewed during session 24), and the preliminary FDIS text was submitted as WG 5 N 181.

Remains valid – not updated: [JVET-AB2006](https://jvet-experts.org/doc_end_user/current_document.php?id=12215) Additional SEI messages for VSEI (Draft 3) [S. McCarthy, T. Chujoh, M. M. Hannuksela, G. J. Sullivan, Y.-K. Wang] [WG 5 DAM N 158]

No output: [JVET-Axx2007](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9679)

Remains valid – not updated: [JVET-X2008](https://jvet-experts.org/doc_end_user/current_document.php?id=11228) Conformance testing for versatile video coding (Draft 7) [J. Boyce, F. Bossen, K. Kawamura, I. Moccagatta, W. Wan]

Remains valid – not updated: [JVET-Y2009](https://jvet-experts.org/doc_end_user/current_document.php?id=11470) Reference software for versatile video coding (Draft 3) [F. Bossen, K. Sühring, X. Li]

Remains valid – not updated [JVET-AB2010](https://jvet-experts.org/doc_end_user/current_document.php?id=12216) VTM and HM common test conditions and software reference configurations for SDR 4:2:0 10 bit video [F. Bossen, X. Li, V. Seregin, K. Sharman, K. Sühring]

[JVET-AC2011](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] (2023-02-03)

Remains valid – not updated: [JVET-U2012](https://jvet-experts.org/doc_end_user/current_document.php?id=10681) JVET common test conditions and evaluation procedures for 360° video [Y. He, J. Boyce, K. Choi, J.-L. Lin]

Remains valid – not updated: [JVET-T2013](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10546) VTM common test conditions and software reference configurations for non-4:2:0 colour formats [Y.-H. Chao, Y.-C. Sun, J. Xu, X. Xu]

Remains valid – not updated: [JVET-Q2014](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9683) JVET common test conditions and software reference configurations for lossless, near lossless, and mixed lossy/lossless coding [T.-C. Ma, A. Nalci, T. Nguyen]

Remains valid – not updated: [JVET-Q2015](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9684) JVET functionality confirmation test conditions for reference picture resampling [J. Luo, V. Seregin]

[JVET-AC2016](https://jvet-experts.org/doc_end_user/current_document.php?id=12217) Common Test Conditions and evaluation procedures for neural network-based video coding technology [E. Alshina, R.-L. Liao, S. Liu, A. Segall] (2023-02-03)

Template has an error to be fixed.

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]

Remains valid – not updated: [JVET-AA2018](https://jvet-experts.org/doc_end_user/current_document.php?id=11949) Common test conditions for high bit depth and high bit rate video coding [A. Browne, T. Ikai, D. Rusanovskyy, X. Xiu, Y. Yu]

[JVET-AC2019](https://jvet-experts.org/doc_end_user/current_document.php?id=12218) Description of algorithms and software in neural network-based video coding (NNVC) version 2 [Y. Li, H. Wang, L. Wang, F. Galpin, J. Ström] [WG 5 N 188] (2023-03-10)

New elements from notes elsewhere in this report:

* Decision: Adopt JVET-AC0194 method 1.1-2
* Decision: Adopt JVET-AC0196 into NNVC 4.0 (EE1-2.2)
* Decision: Adopt JVET-AC0055
* Decision: Adopt JVET-AC0116
* Decision: Adopt JVET-AC0310, but make 1.5.3 the default configuration, and allow disabling 1.7.1 component. Further, the method shall be changed to use the actual reference frame (not the unfiltered decoded frame)
* Decision: Adopt JVET-AC0328, but disable the RDO network in the encoder by default (unless such an approach is generally enabled for both filter sets #0 and #1)

It is noted that the list above may not be complete; if some adoption is missing that is recorded somewhere else in the meeting notes it shall also be considered included.

[JVET-AC2020](https://jvet-experts.org/doc_end_user/current_document.php?id=12219) Film grain synthesis technology for video applications (Draft 4) [D. Grois, Y. He, W. Husak, P. de Lagrange, A. Norkin, M. Radosavljević, A. Tourapis, W. Wan] [WG 5 CDTR N 176] (2023-03-31)

Developed from JVET-AC0151.

[JVET-AC2021](https://jvet-experts.org/doc_end_user/current_document.php?id=12220) Draft verification test plan for VVC multilayer coding [S. Iwamura, P. de Lagrange, M. Wien] (2023-03-03)

See notes under 4.4 for updates.

[JVET-AC2022](https://jvet-experts.org/doc_end_user/current_document.php?id=12221) Draft plan for subjective quality testing of FGC SEI message [P. de Lagrange, W. Husak, M. Radosavljević, M. Wien] (2023-03-03)

See notes under 4.4 for updates.

[JVET-AC2023](https://jvet-experts.org/doc_end_user/current_document.php?id=12208) Exploration Experiment on Neural Network-based Video Coding (EE1) [E. Alshina, F. Galpin, Y. Li, D. Rusanovskyy, M. Santamaria, J. Ström, L. Wang, Z. Xie] [WG 5 N 187] (2023-02-03)

An initial draft of this document was reviewed and approved at 1515-1535 on Friday 20 January.

Categories are enhancement loop filters and inter coding

[JVET-AC2024](https://jvet-experts.org/doc_end_user/current_document.php?id=12209) Exploration Experiment on Enhanced Compression beyond VVC capability (EE2) [V. Seregin, J. Chen, R. Chernyak, K. Naser, J. Ström, F. Wang, M. Winken, X. Xiu, K. Zhang] [WG 5 N 189] (2023-03-17)

An initial draft of this document was reviewed and approved at 1535-1550 on Friday 20 January.

Categories are intra prediction, inter prediction, screen content coding, in-loop filters, and GDR.

It was suggested moving 1.2 after 1.7 to better group the CCCM related proposals.

The number of tests in the intra category is rather large, and will require significant time for review at the next meeting. Proponents are asked to consider withdrawing those tests where they find that the benefit is insignificant.

[JVET-AC2025](https://jvet-experts.org/doc_end_user/current_document.php?id=12222) Algorithm description of Enhanced Compression Model 8 (ECM 8) [M. Coban, F. Le Léannec, R.-L. Liao, K. Naser, J. Ström, L. Zhang] [WG 5 N 190] (2023-03-24)

New elements from notes elsewhere in this report:

* Decision: Adopt JVET-AC0105 (test 1.2)
* Decision: Adopt JVET-AC0119 (test 1.6)
* Decision: Adopt JVET-AC0098 (test 1.8)
* Decision: Adopt JVET-AC0094 (test 1.10)
* Decision: Adopt JVET-AC0054 (test 1.12a)
* Decision: Adopt JVET-AC0147 (test 1.13)
* Decision: Adopt JVET-AC0144 (test 2.2)
* Decision: Adopt JVET-AC0158 (test 2.5a) with the following modification: It was agreed to adopt a solution with pixel-based affine as in test 2.5a (not using pixel based affine whenever OBMC is used in the given block), but without the modification of OBMC flag signalling
* Decision: Adopt JVET-AC0185 test 2.7a
* Decision: Adopt JVET-AC0113 (version 3.5a)
* Decision: Adopt JVET-AC0112 test 3.6d
* Decision: Adopt JVET-AC0115 test 4.1a
* Decision: Adopt JVET-AC0130 (test 4.2b)
* Decision: Adopt JVET-AC0162 (5.1b)
* Decision: Adopt JVET-AC0053
* Decision: Adopt JVET-AC0189 (in CTC, set the PPS flag to 1 for class F and class TGM)

It is noted that the list above may not be complete; if some adoption is missing that is recorded somewhere else in the meeting notes it shall also be considered included.

[JVET-AC2026](https://jvet-experts.org/doc_end_user/current_document.php?id=11477) Conformance testing for VVC operation range extensions (Draft 4) [D. Rusanovskyy, T. Hashimoto, H.-J. Jhu, I. Moccagatta, Y. Yu] (2023-04-14)

JVET-AC2026 is an update of JVET-Y2026 (previously submitted as WG 5 DAM [N 110](https://dms.mpeg.expert/doc_end_user/current_document.php?id=81998&id_meeting=189))

The content of this document was included in a new edition of VVC conformance. A DoCR on ISO/IEC 23090-15/DAM1 was submitted as WG 5 N184 (reviewed during session 24), and the FDIS text was submitted as WG 5 N 185.

[JVET-AC2027](https://jvet-experts.org/doc_end_user/current_document.php?id=12223) SEI processing order SEI message in VVC (draft 3) [S. McCarthy, M. Hannuksela, Y.-K. Wang] [WG 5 preliminary WD 3 N 182] (2023-03-31)

Update from JVET-AC0058

[JVET-AC2028](https://jvet-experts.org/doc_end_user/current_document.php?id=12224) Additional conformance bitstreams for VVC multilayer configurations [S. Iwamura, I. Moccagatta] (2023-04-14)

Remove bitstream that are put into FDIS, and some other wording changes

Remains valid – not updated: [JVET-AB2029](https://jvet-experts.org/doc_end_user/current_document.php?id=12225) Visual quality comparison of ECM/VTM encoding [V. Baroncini, J.-R. Ohm, M. Wien] [AG 5 N75]

JVET-AC2030 Optimization of encoders and receiving systems for machine analysis of coded video content (draft 1) [J. Chen, C. Hollmann, S. Liu] [WG 5 N177)] (2023-03-31)

JVET-AC2031 CTC for optimization of encoders and receiving systems for machine analysis of coded video content [S. Liu, C. Hollmann] (2023-02-17)

JVET-AC2032 Improvements under consideration for neural network post filter SEI messages [S. McCarthy, S. Deshpande, M. Hannuksela, Hendry, G. Sullivan, Y.-K. Wang] [WG 5 N175] (2023-03-03)

For elements included, see BoG report JVET-AC0324 (generate a list from that)

# Future meeting plans, expressions of thanks, and closing of the meeting

Future meeting plans were established according to the following guidelines (assuming face-to-face meetings):

* Meeting under ITU-T SG16 auspices when it meets (ordinarily starting meetings on the Wednesday of the first week and closing it on the Wednesday of the second week of the SG16 meeting – a total of 8 meeting days), and
* Otherwise meeting under ISO/IEC JTC 1/‌SC 29 auspices when its MPEG WGs meet (ordinarily starting meetings on the Friday prior to the main week of such meetings and closing it on the same day as other MPEG WGs – a total of 8 meeting days).

In cases where an exceptionally high workload is expected for a meeting, an earlier starting date may be defined. In cases of online meetings, no sessions should be held on weekend days. This may imply an earlier starting date as well.

Some specific future meeting plans (to be confirmed) were established as follows:

* During Fri. 21 – Fri. 28 April 2023, 30th meeting under ISO/IEC JTC 1/‌SC 29 auspices in Antalya, TR
* During Tue. 11 – Wed. 19 July 2023, 31st meeting under ITU-T SG16 auspices in Geneva, CH
* During Fri. 13 – Fri. 20 October 2023, 32nd meeting under ISO/IEC JTC 1/‌SC 29 auspices, location t.b.d.
* During Wed. 17 – Fri. 19 and Mon. 22 – Fri. 26 January 2024, 33rd meeting under ISO/IEC JTC 1/‌SC 29 auspices, to be held as teleconference meeting.
* During April 2024, 34th meeting under ITU-T SG16 auspices, date and location t.b.d.
* During Fri. 12 – Fri. 19 July 2024, 35th meeting under ISO/IEC JTC 1/‌SC 29 auspices in Sapporo, JP
* During October 2024, 36th meeting under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.
* During January 2025, 37th meeting under ITU-T SG16 auspices, date and location t.b.d.
* During April 2025, 38th meeting under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.
* During Fri. 11 – Fri. 18 July 2025, 39th meeting under ISO/IEC JTC 1/‌SC 29 auspices in Daejeon, KR

The agreed document deadline for the 30th JVET meeting was planned to be Friday 14 April 2023.

It was suggested that it would be interesting to perform subjective quality investigation with the elements adopted to NNVC software in a future meeting (provided that close enough rate/quality matching with VTM anchor is achieved).

The following companies were thanked for preparing new bitstreams to be included in the second edition of VVC conformance specification: Alibaba, InterDigital, KDDI, Kwai, OPPO, Qualcomm, Sharp, Sony.

Marius Preda was thanked for his support in maintaining the document site jvet-experts.org, as well as the document sites of JCT-VC and JCT-3V. Institut Mines-Télécom is thanked for hosting the sites.

The 29th JVET meeting was closed at approximately XXXX hours on XXday 2X January 2023.

# Annex A to JVET report: List of documents

# Annex B1 to JVET report: List of meeting participants attending in person

The participants who were personally present at the meeting site of the twenty-ninth meeting of the JVET, according to an attendance sheet circulated during the meeting sessions (approximately 0 people in total), were as follows (note: Annex kept for future use):

# 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 had attended the meeting in person at least part-time (see annex B1), and not including those who attended only the joint sessions with other groups), were as follows:

# Annex C to JVET report: Recommendations of the 10th meeting of ISO/IEC JTC 1/SC 29/WG 5 MPEG Joint Video Coding Team(s) with ITU-T SG 16

**ISO/IEC JTC 1/SC 29/WG 5 N XXX**