|  |  |
| --- | --- |
| **Joint Video Experts Team (JVET)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29**  27th Meeting, by teleconference, 13–22 July 2022 | Document: JVET-AA\_Notes\_d7 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Meeting Report of the 27th Meeting of the Joint Video Experts Team (JVET), by teleconference, 13–22 July 2022** | | |
| *Status:* | Report document from the chair of JVET | | |
| *Purpose:* | Report | | |
| *Author(s) or Contact(s):* | **Jens-Rainer Ohm** Institute of Communication Engineering RWTH Aachen Melatener Straße 23 D-52074 Aachen | Tel: Email: | +49 241 80 27671 [ohm@ient.rwth-aachen.de](mailto:ohm@ient.rwth-aachen.de) |
| *Source:* | Chair of JVET | | |

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

# Summary

The Joint Video Experts Team (JVET) of ITU-T WP3/16 and ISO/IEC JTC 1/‌SC 29 held its twenty-seventh meeting during 13–22 July 2022 as an online-only meeting. It had previously been planned to be in Cologne, DE, but this plan was changed due to uncertainties resulting from the COVID-19 pandemic. For ISO/IEC purposes, JVET is alternatively designated ISO/IEC JTC 1/‌SC 29/‌WG 5, and this was the eighth 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 0500 hours UTC on Wednesday 13 July 2022. Meeting sessions were held on all days except the weekend days of Saturday and Sunday 16 and 17 July 2022, until the meeting was closed at approximately 00XX hours UTC on Saturday 23 July 2022. Approximately XXX people attended the JVET meeting, and approximately XXX input documents (not counting crosschecks), 13 AHG reports, 2 EE summary reports, and X BoG reports 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 was performed in the interim period since the twenty-sixth JVET meeting in producing the following documents:

* JVET-Z1003 Coding-independent code points for video signal type identification (Draft 1 of 3rd edition)
* JVET-Z1004 Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR
* JVET-Z1005 New levels for HEVC (Draft 3), also issued as WG 5 CDAM
* JVET-Z1008 Additional colour type identifiers for AVC and HEVC (Draft 1)
* JVET-Z2002 Algorithm description for Versatile Video Coding and Test Model 17 (VTM 17)
* JVET-Z2005 New level and systems-related supplemental enhancement information for VVC (Draft 2), also issued as WG 5 CDAM
* JVET-Z2006 Additional SEI messages for VSEI (Draft 1)
* JVET-Z2011 VTM and HM common test conditions and evaluation procedures for HDR/WCG video
* JVET-Z2016 Common Test Conditions and evaluation procedures for neural network-based video coding technology
* JVET-Z2023 Exploration Experiment on neural network-based video coding (EE1)
* JVET-Z2024 Exploration Experiment on enhanced compression beyond VVC capability (EE2)
* JVET-Z2025 Algorithm description of Enhanced Compression Model 5 (ECM 5)

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

* JVET-Z1003 Coding-independent code points for video signal type identification (Draft 1 of 3rd edition)
* JVET-Z1004 Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR
* JVET-Z1005 New levels for HEVC (Draft 3), also issued as WG 5 CDAM
* JVET-Z1008 Additional colour type identifiers for AVC and HEVC (Draft 1)
* JVET-Z2002 Algorithm description for Versatile Video Coding and Test Model 17 (VTM 17)
* JVET-Z2005 New level and systems-related supplemental enhancement information for VVC (Draft 2), also issued as WG 5 CDAM
* JVET-Z2006 Additional SEI messages for VSEI (Draft 1)
* JVET-Z2011 VTM and HM common test conditions and evaluation procedures for HDR/WCG video
* JVET-Z2016 Common Test Conditions and evaluation procedures for neural network-based video coding technology
* JVET-Z2023 Exploration Experiment on neural network-based video coding (EE1)
* JVET-Z2024 Exploration Experiment on enhanced compression beyond VVC capability (EE2)
* JVET-Z2025 Algorithm description of Enhanced Compression Model 5 (ECM 5)

For the organization and planning of its future work, the JVET established 13 “ad hoc groups” (AHGs) to progress the work on particular subject areas. At this meeting, 2 Exploration Experiments (EE) were defined. The next eight JVET meetings were planned for 21 – 28 October 2022 under ITU-T SG16 auspices in Antalya, TR; during 11 – 20 January 2023 under ISO/IEC JTC 1/‌SC 29 auspices, to be conducted as a teleconference meeting; during April 2023 under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.; during July 2023 under ITU-T SG16 auspices, date and location t.b.d.; during October 2023 under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.; during January 2024 under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.; during April 2024 under ITU-T SG16 auspices, date and location t.b.d.; and during July 2024 under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.

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

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

# Administrative topics

## Organization

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

The Joint Video Experts Team (JVET) of ITU-T WP3/16 and ISO/IEC JTC 1/‌SC 29 held its twenty-seventh meeting during 13–22 July 2022 as an online-only meeting, using the Zoom teleconferencing tool. For ISO/IEC purposes, JVET is alternatively designated ISO/IEC JTC 1/‌SC 29/‌WG 5, and this was the eighth meeting as WG 5. The JVET meeting was held under the chairmanship of Dr Jens-Rainer Ohm (RWTH Aachen/Germany).

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

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

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

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

This report contains three important annexes, as follows:

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

## Meeting logistics

Information regarding logistics arrangements for the meeting had been provided via the email reflector [jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de) and at <http://wftp3.itu.int/av-arch/jvet-site/2022_07_AA_Virtual/>.

## Primary goals

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

* JVET-Z1003 Coding-independent code points for video signal type identification (Draft 1 of 3rd edition)
* JVET-Z1004 Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR
* JVET-Z1005 New levels for HEVC (Draft 3), also issued as WG 5 CDAM
* JVET-Z1008 Additional colour type identifiers for AVC and HEVC (Draft 1)
* JVET-Z2002 Algorithm description for Versatile Video Coding and Test Model 17 (VTM 17)
* JVET-Z2005 New level and systems-related supplemental enhancement information for VVC (Draft 2), also issued as WG 5 CDAM
* JVET-Z2006 Additional SEI messages for VSEI (Draft 1)
* JVET-Z2011 VTM and HM common test conditions and evaluation procedures for HDR/WCG video
* JVET-Z2016 Common Test Conditions and evaluation procedures for neural network-based video coding technology
* JVET-Z2023 Exploration Experiment on neural network-based video coding (EE1)
* JVET-Z2024 Exploration Experiment on enhanced compression beyond VVC capability (EE2)
* JVET-Z2025 Algorithm description of Enhanced Compression Model 5 (ECM 5)

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

## Documents and document handling considerations

### General

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

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

The document registration and upload times and dates listed in Annex A and in headings for documents in this report are in Paris/Geneva time. Dates mentioned for purposes of describing events at the meeting (other than as contribution registration and upload times) follow the UTC timezone.

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, 6 July 2022. Any documents uploaded after 1159 hours Paris/Geneva time on Thursday 7 July 2022 were considered “officially late”, with a grace period of 12 hours (to accommodate those living in different time zones of the world). The deadline does not apply to AHG reports and other such reports which can only be produced after the availability of other input documents.

All contribution documents with registration numbers higher than JVET-AA0150 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-AA0XXX (a proposal on …), uploaded 07-XX,
* … .

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

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

* JVET-AA0XXX (a document on …), uploaded 07-XX,
* … .

All cross-verification reports at this meeting were registered late, and/or uploaded late. In the interest of brevity, these are not specifically identified here. Initial upload times for each document are recorded in Annex A of this report.

At some previous meetings, some cross-verification reports had not been uploaded yet by the time when the meeting ended, and neither were they provided within two weeks after the meeting: This case did not happen at this meeting.

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-AA0041, JVET-AA0049, JVET-AA0050, JVET-AA0060, JVET-AA0068, JVET-AA0077, JVET-AA0180, JVET-AA0197, JVET-AA0198, JVET-AA0199.

“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-AA0094 and JVET-AA0143, which were both categorized as late in the list above, based on the time of the first reasonable document upload.

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

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

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

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

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

### Outputs of the preceding meeting

All output documents of the previous meeting, particularly the meeting report JVET-Z1000, the Coding-independent code points for video signal type identification (Draft 1 of 3rd edition) JVET-Z1003, the Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR JVET-Z1004, the New levels for HEVC (Draft 3) JVET-Z1005, the Additional colour type identifiers for AVC and HEVC (Draft 1) JVET-Z1008, the Algorithm description for Versatile Video Coding and Test Model 17 (VTM 17) JVET-Z2002 (updated version 2 delivered on the second day of the current meeting), the New level and systems-related supplemental enhancement information for VVC (Draft 2), JVET-Z2005, the Additional SEI messages for VSEI (Draft 1) JVET-Z2006, the VTM and HM common test conditions and evaluation procedures for HDR/WCG video JVET-Z2011, the Common Test Conditions and evaluation procedures for neural network-based video coding technology JVET-Z2016, the Description of the EE on Neural Network-based Video Coding JVET-Y2023, the Description of the EE on Enhanced Compression beyond VVC capability JVET-Y2024, and the Algorithm description of Enhanced Compression Model 5 (ECM 5) JVET-Z2025, had been completed and were approved. The software implementations of VTM (version2 16.1, 16.2, and 17.0), and ECM (version 5.0) were also approved.

Only minor editorial issues were found in the meeting report JVET-Z1000; 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 the 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.
* 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.
* 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 attendance records.
* Turn on the chat window and watch for chair communication and side commentary there as well as by audio.
* Generally do not use video for the teleconferencing calls in order to avoid overloading people’s internet connections; enable only voice and screen sharing.
* Extensive use of screen sharing is encouraged, to enable participants to view the presented material and the meeting notes. At times, multiple sources of screen sharing may be enabled, so it may be necessary for participants to understand that this is happening and to understand how to select which one they want to watch.

## Agenda

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

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

The agenda was approved as suggested

The plans for the times of meeting sessions were established as follows, in UTC (which for this meeting was 2 hours behind the time in Geneva and Paris; 7 hours ahead of the time in Los Angeles, etc.). No session was scheduled to last longer than 2 hrs.

* 0500–0700 1st “morning” session [break after 2 hours]
* 0720–0920 2nd “morning” session
* [“midday” break – nearly 4 hours]
* 1300–1500 1st “afternoon” session [break after 2 hours]
* 1520–1720 2nd “afternoon” 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 24th JVET meeting (which had been the seventh meeting conducted as an online meeting)

* 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 is used in the process of designing the VTM software, and for evaluating proposals for technology to be potentially included in the design. This software or parts thereof might be published by ITU-T and ISO/IEC as an example implementation of a future video coding standard and for use as the basis of products to promote adoption of such technology.

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

These considerations apply to the 360Lib video conversion software and HDRTools as well. It is noted that the SADL package for neural network-based video coding uses the same licensing terms.

Software packages that had been developed in prior work of the JCT-VC 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 1182 (as of 11 July 2022). All discussions (including those on AVC, HEVC, VVC, CICP, etc.) shall be conducted on the JVET reflector rather than any of the old reflectors (including JVT, JCT-VC, and JCT-3V) which are retained for archiving purposes.

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

## Terminology

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

## Opening remarks

Remarks during the opening session of the meeting Wednesday 20 April at 0500 UTC were as follows.

* Timing and organization of online meetings, calendar posting of session plans
* Standards, TRs, supplements and technical papers approval and publication status
  + AVC
    - H.264 V14 Consented at 22nd meeting on 2021-04-30 (with annotated regions, shutter interval, and miscellaneous corrections), approved 2021-08-22, published 2021-10-13
    - ISO/IEC 14496-10:2020 (Ed. 9) FDIS ballot closed 2020-11-27, published 2020-12-15
    - ISO/IEC 14496-10:202X (Ed. 10), had been forwarded from DIS directly for publication 2022-01-21 (with annotated regions, shutter interval, and miscellaneous corrections) with an editing period, submitted to ITTF in 2022-05 after consultation with ISO staff on format of graphics files, upgraded to “DIS approved for registration” in ISO Project system 2022-07-04, currently in consultation between ISO EPM and project editors
    - Preliminary draft text for YCgCo-Re and YCgCo-Ro issued at 26th meeting (not yet formally requested as a project)
    - Conformance testing
      * H.264.1 V6 Approved 2016-02-13, published 2016-06-17
      * Various amendments of ISO/IEC 14496-4:2004, including:
        + ISO/IEC 14496-4:2004/AMD 6:2005 Advanced Video Coding conformance
        + ISO/IEC 14496-4:2004/AMD 9:2006 AVC fidelity range extensions conformance
        + ISO/IEC 14496-4:2004/AMD 30:2009 Conformance testing for new profiles for professional applications
        + ISO/IEC 14496-4:2004/AMD 31:2009 Conformance testing for SVC profiles
        + ISO/IEC 14496-4:2004/AMD 38:2010 Conformance testing for Multiview Video Coding
        + ISO/IEC 14496-4:2004/AMD 41:2014 Conformance testing of MVC plus depth extension of AVC
        + ISO/IEC 14496-4:2004/AMD 42:2014 Conformance testing of Multi-Resolution Frame Compatible Stereo Coding extension of AVC
        + ISO/IEC 14496-4:2004/AMD 43:20153D-AVC conformance testing
        + ISO/IEC 14496-4:2004/AMD 45:2016 Conformance Testing for the Multi-resolution Frame Compatible Stereo Coding with Depth Maps Extension of AVC
    - Reference software
      * H.264.2 V7 Approved 2016-02-13, published 2016-05-30
      * Various amendments of ISO/IEC 14496-5:2001, including:
        + ISO/IEC 14496-5:2001/AMD 6:2005 Advanced Video Coding (AVC) and High Efficiency Advanced Audio Coding (HE AAC) reference software
        + ISO/IEC 14496-5:2001/AMD 8:2006 AVC fidelity range extensions reference software
        + ISO/IEC 14496-5:2001/AMD 15:2010 Reference software for Multiview Video Coding
        + ISO/IEC 14496-5:2001/AMD 18:2008 Reference software for new profiles for professional applications
        + ISO/IEC 14496-5:2001/AMD 19:2009 Reference software for Scalable Video Coding
        + ISO/IEC 14496-5:2001/AMD 33:2015 Reference software for MVC plus depth extension of AVC
        + ISO/IEC 14496-5:2001/AMD 34:2014 Reference software of the multi-resolution frame compatible stereo coding of AVC
        + ISO/IEC 14496-5:2001/AMD 35:2015 3D-AVC Reference software
        + ISO/IEC 14496-5:2001/AMD 39:2016 Reference software for the Multi-resolution Frame Compatible Stereo Coding with Depth Maps of AVC
        + ISO/IEC 14496-5:2001/AMD 42:2017 Reference software for the alternative depth information SEI message extension of AVC
  + HEVC
    - H.265 V7 approved 2019-11-29, published 2020-01-10
    - ISO/IEC 23008-2:2020 (Ed. 4) FDIS ballot closed 2020-07-16, published 2020-08-27
    - H.265 V8 Consented at the 22nd meeting (shutter interval information SEI message and miscellaneous corrections), published 2020-10-13
    - ISO/IEC 23008-2:2020/AMD 1:2021 (shutter interval information SEI message) published 2021-07-12
    - ISO/IEC 23008-2:202x (Ed. 5) began as CDAM 2 High-range levels output of 25th meeting of January 2022, CDAM ballot closed 2022-04-15, conversion to 5th edition with miscellaneous corrections planned at 26th meeting of April 2022, text submitted for DIS ballot 2022-07-10
    - Preliminary draft text for YCgCo-Re and YCgCo-Ro issued at 26th meeting (not yet formally requested as a project)
    - Conformance testing
      * H.265.1 V3 approved 2018-10-14, published 2019-01-15
      * ISO/IEC 23008-8:2018 (Ed. 2) Conformance specification for HEVC, published 2018-08
      * ISO/IEC 23008-8:2018/AMD 1:2019 Conformance testing for HEVC screen content coding (SCC) extensions and non-intra high throughput profiles, published 2019-10
    - Reference software
      * H.265.2 V4 approved 2016-12-22, published 2017-04-10
      * ISO/IEC 23008-5:2017 (Ed. 2) Reference software for high efficiency video coding, published 2017-02
      * ISO/IEC 23008-5:2017/AMD 1:2017 Reference software for screen content coding extensions, published 2017-10
  + VVC
    - H.266 V1 approved 2020-08-29, published 2020-11-10
    - ISO/IEC 23090-3:2021 (Ed. 1) published 2021-02-16
    - H.266 V2 with operation range extensions, Consented 2022-01-28, Last Call began 2022-04-01, Approved 2022-04-29, pre-published 2022-06-06, published 2022-07-12
    - ISO/IEC 23090-3:202x (Ed. 2) with operation range extensions, approval at WG level to proceed to FDIS 2022-01-21, FDIS ballot initiated 2022-06-29, voting to close 2022-08-24
    - ISO/IEC 23090-3:202x (Ed. 2) / CDAM 1 New level and systems-related supplemental enhancement information issued from 26th meeting, ballot to close 2022-07-14
    - Conformance testing
      * H.266.1 V1 Consented 2022-01-28, Last Call began 2022-04-01, Approved 2022-04-29, pre-published 2022-05-17, published 2022-07-12
      * ISO/IEC 23090-15 V1 approval at WG level to proceed to FDIS 2022-10-15, upgraded to “DIS approved for registration” in ISO Projects system 2021-10-24, upgraded to “FDIS registered for formal approval” 2022-07-11
      * ISO/IEC 23090-15 DAM 1 Operation range extensions – DAM issued from 25th meeting 2022-01-21, upgraded to “CD approved for registration as DIS” status in ISO Projects system 2022-05-31, upgraded to “DIS registered” 2022-06-22, DAM ballot pending, no action at his meeting
    - Reference software
      * H.266.2 V1 Consented 2022-01-28, Last Call began 2022-04-01, Approved 2022-04-29, pre-published 2022-05-17, published 2022-07-12
      * ISO/IEC 23090-16 V1 approval at WG level to proceed to FDIS 2022-01-21, FDIS ballot pending, upgraded to “DIS approved for registration” status in ISO Projects system 2022-04-21, upgraded to “FDIS registered for formal approval” 2022-04-22
  + VSEI
    - H.274 V1 approved 2020-08-29, published 2020-11-10
    - ISO/IEC 23002-7:2021 (Ed. 1) published 2021-01-28
    - H.274 V2 Consented 2022-01-28, Last Call began 2022-04-01, Approved 2022-05-22 (after 1 Last Call comment and Additional Review), pre-published 2022-06-17
    - ISO/IEC 23002-7:202x (Ed. 2) approval at WG level to proceed to FDIS 2022-01-21, upgraded to “DIS approved for registration” status in ISO Projects system 2022-05-05 and “FDIS registered for formal approval” 2022-05-08, FDIS ballot pending
    - ISO/IEC 23002-7:202x (2nd Ed.) Amd.1 Request for new edition and WD for additional SEI messages issued at 26th meeting 2022-04-29
  + 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 WD for YCgCo-Re and YCoCg-Ro issued at 26th meeting 2022-04-29
  + Conversion and coding practices for HDR/WCG Y′CbCr 4:2:0 video with PQ transfer characteristics
    - H.Sup15 V1, approved 2017-01-27, published 2017-04-12
    - ISO/IEC TR 23008-14:2018 published 2018-08
  + Signalling, backward compatibility and display adaptation for HDR/WCG video coding
    - H.Sup18 V1, approved 2017-10-27, published 2018-01-18
    - ISO/IEC TR 23008-15:2018 published 2018-08
  + Usage of video signal type code points
    - H.Sup19 V3 approved 2021-04-30, published 2021-06-04
    - ISO/IEC TR 23091-4 (Ed. 3) published 2021-05-23
  + Working practices using objective metrics for evaluation of video coding efficiency experiments
    - HSTP-VID-WPOM V1: approved 2020-07-03, published 2020-11
    - ISO/IEC TR 23002-8 (Ed. 1) published 2021-05-20
  + Film grain synthesis technologies for video applications
    - ISO/IEC TR 23002-9 Request for subdivision and WD 1 issued at 25th meeting 2022-01-21, uploaded 2022-04-20, no action taken at 26th meeting
  + The following freely available standards are published here in ISO/IEC:  
    <https://standards.iso.org/ittf/PubliclyAvailableStandards/index.html>
    - ISO/IEC 14496-10:2020 (Ed. 9) AVC
    - ISO/IEC 23002-7:2021 (Ed. 1) VSEI
    - ISO/IEC 23008-2:2020 (Ed. 4) HEVC
    - ISO/IEC 23090-3:2021 (Ed. 1) VVC
  + The following standards that have been intended by JVET to be publicly available were not available at <https://standards.iso.org/ittf/PubliclyAvailableStandards/index.html> as of 2022-07-12. (Please see below for record of previously issued requests.)
    - ISO/IEC 23091-2:2021 (Ed. 2) Video CICP (was requested in April 2021, and the 2019 previous edition was also not made available there)
    - ISO/IEC 23008-2:2020 (Ed. 4) Amd.1:2021: Shutter interval information SEI message (has not been requested but separate publication may not be necessary if it is promptly included in a next edition)
    - ISO/IEC 14496-10:202X – AVC 10th edition – final text issued and public availability requested at the 25th meeting (January 2022)
    - ISO/IEC 23002-7:202X – VSEI 2nd edition – FDIS issued and public availability requested at the 25th meeting (January 2022)
    - ISO/IEC 23090-3:202X – VVC 2nd edition – FDIS issued and public availability requested at the 25th meeting (January 2022)
    - ISO/IEC 23090-15:202X – VVC conformance – FDIS issued and public availability requested at the 24th meeting (October 2021)
    - ISO/IEC 23090-16:202X – VVC reference software – FDIS issued and public availability requested at the 25th meeting (January 2022)
* Draft standards progression status
  + New levels (from JVET-Z1005) – ISO/IEC 23008-2 DIS of new edition of HEVC was issued from 26th meeting, incorporating Amd.1 and corrigenda items (expecting FDIS in January 2023, ITU-T consent in October 2022); note that Amd.1 = shutter interval SEI is already included in latest ITU-T edition of H.265
  + New level and systems-related supplemental enhancement information (from JVET-Z2005) –VVC CDAM was issued from 26th meeting, ballot to close 2022-07-14, DAM (JVET draft 3) to be issued
  + Additional SEI messages (from JVET-Z2006) –VSEI CDAM (JVET draft 2) to be issued from current meeting (request made at 26th meeting)
  + Film grain synthesis technology for video applications (from JVET-Y2020) – JVET draft 2 to be issued at the current meeting, also ISO/IEC 23002-9 DTR (request made at 25th meeting)
  + VVC Conformance testing for operation range extensions – (from JVET-Y2026) – plan for ISO/IEC 23090-16 FDAM and ITU-T consent in October – new JVET draft at current meeting? New edition ISO/IEC in October?
  + Video CICP new edition with for YCgCo-Re and YCgCo-Ro (from JVET-Z1003) – JVET draft 2 to be issued at the current meeting, also ISO/IEC 23091-2 CD (request made at 26th meeting), target ITU-T consent in October
  + 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 if it becomes included in next edition
* The meeting logistics, agenda, working practices, policies, and document allocation considerations were reviewed.
  + The meeting was conducted using Zoom.
  + Having text and software available is crucial (and not just arriving at the end of the meeting).
  + There were no objections voiced in the opening plenary to the consideration of late contributions.
* The results of the previous meeting and the meeting report JVET-Z1000 were reviewed. The following small issues in the meeting report were noted and were not considered sufficient to warrant issuing a revision:
  + In 1 (summary), the information that JVET-Z1005 would have been issued as WG 5 CDAM is wrong. The CDAM had already been issued from JVET-Y1005 of the 25th meeting.
  + In 2.13, the session 0720-0920 should have been designated as 2nd “morning” session instead of “afternoon” session.
* There was somewhat less of a problem of late non-cross-check documents; two “placeholders” were rejected in their initial versions, and updates were not made before the deadline (see section 2.4.2).
* 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
  + Errata
  + Checking the status on conformance testing for version 2 of VVC (under DAM ballot)
  + New levels for HEVC (DIS under ballot)
  + New level and systems-related SEI for VVC (DAM)
  + Additional SEI in VSEI (CDAM)
  + Preparation of TR for film grain (draft 2)
  + New edition video CICP (CD)
  + Additional color type identifiers for AVC and HEVC (Draft 2) – should be included in new edition of H.265 (23008-2 DIS needs ballot comment), no action for AVC at this moment.
  + Exploration Experiments
    - Neural network-based video coding
    - Enhanced compression beyond VVC
* Liaison communication: tbd. -> JPEG?
* The number of documents was lower than for the previous meeting (135->115)
* 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.
* Meeting plans need to be discussed

## Scheduling of discussions

The plans for the times of meeting sessions were established as follows, in UTC (which for this meeting was 2 hours behind the time in Geneva and Paris; 7 hours ahead of the time in Los Angeles, etc.). No session should last longer than 2 hrs.

* 0500–0700 1st “morning” session [break after 2 hours]
* 0720–0920 2nd “morning” session
* [“midday” break – nearly 4 hours]
* 1300–1500 1st “afternoon” session [break after 2 hours]
* 1520–1720 2nd “afternoon” 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. 13 July, 1st day
  + Session 1:
    - 0500–0550 Opening remarks, review of practices, agenda, IPR reminder (section 2)
    - 0550–0700 Reports of AHGs 1–4 (section 3)
  + Session 2:
    - 0720–0850 Reports of AHGs 5–13 (section 3)
    - 0850–0920 Review of EE1 summary (section 5.2.1)
  + Session 3:
    - 1300–1500 Review of EE2 summary (section 5.3.1)
  + Session 4:
    - 1520–1720 Review of EE2 summary (section 5.3.1)
* Thu. 14 July, 2nd day
  + Session 5:
    - 0500–0700 Review of SEI (section 6.1)
  + Session 6:
    - 0720–0920 Review of SEI (sections 6.1, 6.2)
  + Session 7:
    - 1300–1500 Continue review of EE2 summary, and EE2 related
  + Session 8:
    - 1520–1720 Continue review of EE1 summary and EE1 related
* Fri. 15 July, 3rd day
  + Session 9:
    - 0500–0700 Review of SEI (section 6.2)
  + Session 10:
    - 0720–0920 Review of HLS (section 6.x)
  + Session 11:
    - 1300–1500 Continue review of EE2 related (5.3.3) and other ECM related (5.3.4)
  + Session 12:
    - 1520–1720 Continue review of EE1 (5.2.2)
* Mon. 18 July, 4th day
  + 0500–0730 MPEG information sharing session
  + Session 13:
    - 0750–0920 Review of EE1 related (5.2.3)
  + 1400-1500 Joint meeting with AG 5: 4.3, 4.5
  + Session 14:
    - 1520–1720 Review of ECM related (5.3.4)
* Tue. 19 July, 5th day
  + 0500-0700 BoG on NN based VC (E. Alshina, A. Segall)
  + Session 15:
    - 0500–0700 Review project development (4.x)
  + Session 16:
    - 0720–0920 BoG report; Review NN based VC beyond EE1 (5.2.4)
  + Session 17 (chaired by Y. Ye from 1400):
    - 1300–1400 HLS (6.x),
    - 1400-1500 Review remaining 5.3.4
  + Session 18 (chaired by Y. Ye):
    - 1520–1720 Review remaining 5.3.4
  + 1520-1720 BoG on NN based VC (E. Alshina, A. Segall)
* Wed. 20 July, 6th day
  + 0500–0600 MPEG information sharing session
  + Session 19:
    - 0600–0700 Remaining doc reviews 6.4, 4.3
  + Session 20:
    - 0720–0920 NNVC BoG report, remaining doc reviews 4.7, 4.8, 4.9, 4.10, 4.12 5.3.4
  + Session 21:
    - 1300–1500 Remaining doc reviews 5.2.4
  + Session 22:
    - 1520–1720 Remaining doc reviews 5.3.4
  + 1520-1720 BoG on NN based VC (E. Alshina, A. Segall)
* Thu. 21 July, 7th day
  + Session 23:
    - 0500–0705 BoG report, Remaining doc reviews, revisits, further planning
  + Session 24:
    - 0730–0920 Revisits, further planning
  + Session 25:
    - 1300–1430 Class G side activity report, , output document planning, software timelines
* Fri. 22 July, 8th day
  + Plenary:
    - 0500–XXXX AHG, output document timelines, review/approval of DoCR
    - XXXX–XXXX Output document reviews and approvals (section 10), draft recommendations, meeting planning, AoB
  + 2100–2300 MPEG information sharing session
  + XXXX–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 may not be completely accurate):

* AHG reports (13) (section 3)
* Project development (section 4)
  + Deployment and advertisement of standards (2)
  + Text development and errata reporting (2)
  + Test conditions (6)
  + Verification testing (0)
  + Test Material (1)
  + Quality assessment (0)
  + Conformance test development (1)
  + Software development (1)
  + Implementation studies and complexity analysis (1)
  + AHG7: Low latency and constrained complexity (3)
  + Encoding algorithm optimization (1)
  + Profile/tier/level specification (1)
  + Proposed modification of system interface (0)
* 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 (1) (section 5.1)
  + AHG11 and EE1: Neural network-based video coding (24) (section 5.2)
  + AHG12 and EE2: Enhanced compression beyond VVC capability (59) (section 5.3) 2 TBP
* High-level syntax (HLS) proposals (section 6) with subtopics
  + AHG9: SEI message studies and proposals (5) (section 6.1)
  + Neural-network post filter (8) (section 6.2)
  + Film grain synthesis (3) (section 6.3)
  + Non-SEI HLS aspects (1) (section 6.4)
* Joint meetings, plenary discussions, BoG and viewing reports (0), summary of actions (section 0)
* Project planning (section 8)
* Establishment of AHGs (section 9)
* Output documents (section 10)
* Future meeting plans and concluding remarks (section 11)

The document counts above do not include cross-checks and summary reports.

# AHG reports (13)

These reports were discussed Wednesday 13 July 2022 in session 1 during 0610–0700 and in session 2 0720–0845 UTC (chaired by JRO).

[JVET-AA0001](https://jvet-experts.org/doc_end_user/current_document.php?id=11827) JVET AHG report: Project management (AHG1) [J.-R. Ohm, G. J. Sullivan (AHG chairs)]

The work of the JVET overall had proceeded well in the interim period with slightly lower number of input documents (as compared to the previous meeting) submitted to the current meeting. Intense discussion had been carried out on the group email reflector, and 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_04_Z_Virtual/>). It is noted that the previous document sites <http://phenix.int-evry.fr/jvet/>, <http://phenix.int-evry.fr/jct/>, and <http://phenix.int-evry.fr/jct3v/> are still accessible, but were converted to read-only.

The list of documents produced included the following, particularly:

* The meeting report (JVET-Z1000) [Posted 2022-05-27, also submitted as WG 5 N 124]
* Coding-independent code points for video signal type identification (Draft 1 of 3rd edition) (JVET-Z1003) [Posted 2022-07-08, also submitted as WG 5 N 132]
* Errata report items for VVC, HEVC, AVC, Video CICP, and CP usage TR (JVET-Z1004) [Posted 2022-06-15]
* New levels for HEVC (Draft 3) (JVET-Z1005) [Posted 2022-06-15, last update 2022-07-11]
* Additional color type identifiers for AVC and HEVC (Draft 1) (JVET-Z1008) [Posted 2022-07-08]
* Algorithm description for Versatile Video Coding and Test Model 17 (VTM 17) (JVET-Z2002) [Posted 2022-06-30, last update 2022-07-14, also submitted as WG 5 N 130]
* New level and systems-related supplemental enhancement information for VVC (Draft 2) (JVET-Z2005) [Posted 2022-05-18, last update 2022-05-20, also submitted as WG 5 CDAM N 129]
* Additional SEI messages for VSEI (Draft 1) (JVET-Z2006) Posted 2022-06-21, last update 2022-03-21, also submitted as WG 5 WD N 126]
* VTM and HM common test conditions and evaluation procedures for HDR/WCG video (JVET-Z2011) [Posted 2022-05-19]
* Common Test Conditions and evaluation procedures for neural network-based video coding technology (JVET-Z2016) [Posted 2022-05-20]
* Exploration experiment on Neural Network-based Video Coding (EE1) (JVET-Z2023) [Posted 2022-04-28, last update 2022-05-20, also submitted as WG 5 N 133]
* Exploration experiment on Enhanced Compression beyond VVC capability (EE2) (JVET-Z2024) [Posted 2022-04-29, last update 2022-05-28, also submitted as WG 5 N 134]
* Algorithm description of Enhanced Compression Model 5 (ECM 5) (JVET-Z2025) [Posted 2022-07-04, also submitted as WG 5 N 135]

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

* Recommendations of the 7th WG 5 meeting (WG 5 N 123)
* Request for ISO/IEC 23002-7:202x (2nd Ed.) Amd.1 Additional SEI messages (WG 5 N 125)
* Disposition of comments received on ISO/IEC 23008-2 CDAM2 (WG 5 N 127)
* Text of ISO/IEC DIS 23008-2:202X High efficiency video coding (5th edition) (WG 5 N 128)
* Request for ISO/IEC 23091-2:202x Coding-independent code points – Part 2: Video (3rd edition) (WG 5 N 131)
* List of AHGs established at the 7th WG 5 meeting (WG 5 N 136)

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

Due to uncertainties associated with the COVID-19 pandemic, a conversion of the meeting to be conducted only online was again necessitated.

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

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

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

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

A preliminary basis for the document subject allocation and meeting notes for the 27th meeting had been made publicly available on the ITU-hosted ftp site <http://wftp3.itu.int/av-arch/jvet-site/2022_07_AA_Virtual/>.

The AHG recommends its continuation.

[JVET-AA0002](https://jvet-experts.org/doc_end_user/current_document.php?id=11828) JVET AHG report: Draft text and test model algorithm description editing (AHG2) [B. Bross, C. Rosewarne, F. Bossen, J. Boyce, A. Browne, S. Kim, S. Liu, J.-R. Ohm, G. J. Sullivan, A. Tourapis, Y.-K. Wang, Y. Ye (AHG chairs)]

**Output documents produced**

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

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

**JVET-Z1004 Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR**

This document contains a list of reported errata items for VVC, VSEI, HEVC, AVC, Video CICP, and the TR on usage of video signal type code points, for tracking purposes. Some of the items have been confirmed by the JVET and have been agreed to require fixing, while some other items have not yet been confirmed. This document also provides publication status backgrounds of these standards.

Incorporated items at the JVET-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.

**JVET-Z1005 New levels for HEVC (Draft 3)**

This document contains the draft text for changes to the High Efficiency Video Coding (HEVC) standard (ITU‑T H.265 | ISO/IEC 23008-2), for the addition of Levels 6.3, 7, 7.1, and 7.2. In addition, the document also contains some corrections to the previous edition of HEVC.

Draft 3 incorporated items:

* Increasing the max bit rate values for the High tier of levels 7.1 and 7.2 ([JVET-Z0060](https://jvet-experts.org/doc_end_user/current_document.php?id=11495))
* Clarifications on the alpha blending text description (JVET-Z0119)
* Text changes for the unlimited level for video profiles (JVET-Z0122)
* Upon ballot comment US-006, it was decided to increase the maximum frame rate to 960 fps (same value as in the corresponding VVC levels) in the high tier for levels 7.0 and higher.
* Correction of derived luma picture size numbers for 11520x6480 and 12288x6480 in Tables A-11, A-12, and A-13
* Further aspects of errata reports per JVET-Z1004

One issue noticed during the editing of this document is that the maximum luma picture size for new levels 7, 7.1 and 7.2 and the maximum luma sample rate for levels 7.1 and 7.2 may have been determined incorrectly. These values in the text were proposed in JVET-Y0072 (from A. Tourapis, D. Singer, and K. Kolarov), saying that the motivation was to enable “up to video resolutions of 8192x4320 at 120fps” (120 fps for level 7.2, and presumably half that frame rate for level 7.1). However, the proposed values for the specified limits are higher than the values necessary for achieving that goal, as they appear to have been computed by rounding up the luma picture width and height to multiples of 128. In fact, rounding only to multiples of 64 is used to derive all other values in Annex A (and values of MinCbSizeY greater than 64 are prohibited in HEVC). As can be observed in Table A.12 of the JVET-Y0072 proposal document, the drafted values thus support slightly higher frame rates than what is described in the document’s abstract (60.4 and 120.8 rather than 60 and 120 fps for levels 7.1 and 7.2, respectively). A maximum luma picture size of 141 557 760 would appear sufficient, but 142 606 336 was proposed and adopted into the draft text. To support the 8192x4320 picture size at maximum frame rates of 60 and 120 fps, the maximum luma sample rate limits could be reduced from 8 556 380 160 to 8 493 465 600 for level 7.1 and from 17 112 760 320 to 16 986 931 200 for level 7.2. The editors refrained from taking action on this issue since overprovisioning is not necessarily an error and since no input had been received on this issue in the previous meeting cycle. (Incorrect values derived from specified limits were corrected as noted above, but the apparent overprovisioning of the limits was left as-is.)

**JVET-Z1008 Additional color type identifiers for AVC and HEVC (Draft 1)**

This document contains the draft text for the specification of additional colour type identifiers for AVC (Rec. ITU-T H.264 | ISO/IEC 14496-10) and HEVC (Rec. ITU-T H.265 | ISO/IEC 23008-2). Text modifications are provided for specification of code point identifiers for YCoCg-R colour representation with equal luma and chroma bit depths. The new code points are referred to as YCgCo-Re and YCgCo-Ro, where the number of bits added to a source RGB bit depth is 2 (i.e., even) and 1 (odd), respectively. Revision marking is provided to show modifications relative to the basis texts (based on the 2021-08 edition of Rec. ITU-T H.264 and the 2020 edition of ISO/IEC 23008-2, respectively). Equation numbers and their cross-references that are maintained automatically have been updated without revision marking.

**JVET-Z2002 Algorithm description for Versatile Video Coding and Test Model 17 (VTM 17)**

The JVET established the VVC Test Model 17 (VTM17) algorithm description and encoding method at its 26th meeting (20 – 29 April 2022, teleconference). This document serves as a source of general tutorial information on the VVC design and also provides an encoder-side description of VTM17. 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 17 (VTM17) algorithm description and encoding method v1

* Incorporated JVET-Z0046: VTM Software Implementation for GREEN-MPEG SEI Messaging
* Incorporated JVET-Z0072: AHG10/AHG12: Enhanced reference picture structures for ECM and VTM
* Incorporated JVET-Z0099: AHG10: Deblocking in RDO and beta offset minus 2 for VTM
* Incorporated JVET-Z0111: Adaptively bypass affine ME in VTM
* Incorporated JVET-Z0120: AHG9: Shutter interval information SEI message for VSEI
* General editorial improvements throughout

At the time of preparing this AHG2 report, JVET-Z2002-v1 was available, and it was noted that a v2 version was expected shortly during the 27th meeting. The v2 version was then uploaded on 2022-07-14 with the additional modifications listed below:

To be integrated into VVC Test Model 17 (VTM17) algorithm description and encoding method v2

* Incorporated JVET-Z0047: AHG13: Improvements of film grain analysis
* Incorporated JVET-Z0244: AHG9: NN post-filter SEI

**JVET-Z2005 New level and systems-related supplemental enhancement information for VVC (Draft 2)**

This document contains a draft amendment for changes to the Versatile Video Coding (VVC) standard (ITU‑T H.266 | ISO/IEC 23090-3). This amendment includes the support of an additional level of capability for the profiles that have previously been defined, thus expanding the range of application requirements addressed by the standard. This amendment also includes the support of two systems-related supplemental enhancement information (SEI) messages, for signalling of “green metadata” as to be specified in ISO/IEC 23001-11 and of an alternative video decoding interface for immersive media as to be specified in ISO/IEC 23090-13.

Draft 2 incorporated items:

* Updated text changes for the unlimited level for video profiles (JVET-Z0122)
* Fixes for tickets [#1548](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1548) and [#1549](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1549)
* Fixes for tickets [#1544](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1544) and [#1547](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1547)

**JVET-Z2006 Additional SEI messages for VSEI (Draft 1)**

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 and SEI messages to facilitate neural-network-based post-processing filtering.

**Related input contributions**

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

* JVET-AA0048 Some VVC text changes

**Remaining bug tickets**

* [#1555](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1555) Typo in 7.4.3.5 (pps\_num\_exp\_tile\_rows\_minus1)
* [#1556](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1556) Typo in clause 7.4.2.4.5 (Order of VCL NAL units and their association to coded pictures)
* [#1558](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1558) Mismatch between spec & VTM for MipChromaDirectFlag derivation in YUV444
* [#1560](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1560) Typo regarding "maxTbSize"

**Recommendations**

The AHG recommends to:

* Approve JVET-Z1003, JVET-Z1004, JVET-Z1005, JVET-Z1008, JVET-Z2002, JVET-Z2005, and JVET-Z2006 documents as JVET outputs,
* Compare the VVC documents with the VVC software and resolve any discrepancies that may exist, in collaboration with the software AHG,
* Encourage the use of the issue tracker to report issues with the text of both the VVC specification text and the algorithm and encoder description,
* Continue to improve the editorial consistency of VVC text specification and Test Model documents,
* Ensure that, when considering changes to VVC, properly drafted text for addition to the VVC Test Model and/or the VVC specification text is made available in a timely manner,
* Review AHG2 related contributions, bug tickets, and other AHG2 related inputs and act on them if found to be necessary.

During the discussion, it is noted that the over-provision of capabilities mentioned above under JVET-Z1005 could be changed by DIS ballot comments.

#1558 should be considered in JVET-AA2005 (VVC extension).

[JVET-AA0003](https://jvet-experts.org/doc_end_user/current_document.php?id=11829) JVET AHG report: Test model software development (AHG3) [F. Bossen, X. Li, K. Sühring, Y. He, K. Sharman, V. Seregin, A. Tourapis (AHG chairs)]

**Software development**

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

The server is located at:

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

The registration and development workflow are documented at:

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

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

**VTM related activities**

The VTM software can be found at

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

The software development continued on the GitLab server. VTM versions 16.1 and 16.2 were tagged on May 3, and VTM version 17.0 was tagged on Jun. 3. VTM 17.1 is expected during the 27th JVET meeting.

VTM 16.1 was tagged on May 3, 2022. Changes include:

* Fix: Check the value of NoOutputBeforeRecoveryFlag prior to marking no output
* JVET-Y0072: Unlimited level support for all video profiles
* JVET-S0078: No output of prior pics flag
* Avoid potential buffer overflow in illegal bitstream
* Fix #1546: Tune range for m\_cpbRemovalDelayDelta
* Change video\_decoding\_interface to VDI\_sei\_envelope
* JVET-Y0044: Signal Green metadata and VDI SEI messages
* Fix #1542: Correct QPA rate allocation for bit depths other than 10
* Add checks for activated APS ALF flags
* Fix #1543: BIM index out of bounds error
* Cleanup PU::getIntraMPMs and add const qualifiers
* Remove unused field
* Remove commented defaults and unnecessary nullptr cast
* Use nullptr instead of NULL
* Remove unnecessary braces
* Remove unused code, fix braces
* Fix #1541

VTM 16.2 was tagged on May 3, 2022. Changes include:

* Remove macros from previous cycle

VTM 17.0 was tagged Jun. 3, 2022. Changes include:

* Slightly improved readability for fix #1552
* Fix range check of vps\_num\_ptls\_minus1 when it is not signaled
* JVET-Z0244: NN post-filter SEI
* Revert unnecessary includes from !2225
* JVET-Z0120: SII SEI support and illustration of use case "Backwards-compatible HFR video"
* Allocate large buffer on heap instead of stack
* Fix allocation size for VPS olsHrdParams
* Remove variable that is never read
* Remove old HEVC code and clean up related functions
* JVET-Z0046: Green Metadata SEI
* Clean up weighted pred code
* Clean up
* JVET-Z0099: enabling deblocking in RDO and adjusting tc and beta offset for random access and low delay fix
* JVET-Z0072: enhanced referencing configs for random access and low delay
* Clean up variable names, spacing, and a few other things
* Clean up variable naming
* JVET-Z0111: Adaptive bypassing affine ME (default off)

VTM 17.1 is expected to be tagged during the 27th JVET meeting. Changes include so far:

* Remove unnecessary mrgTypeNeighbours field
* JVET-Z0047: Improved film grain analysis
* Fix which ptl\_mulltlayer\_enabled\_flag is used
* Fix setNumOutputLayerSets() to accept value 256
* Use chromaFormatIDC stored inside Picture
* Fix: lintf undef
* Avoid floating-point operations when computing HAD
* Clean up variable names, indentation, braces
* Rename BcwIdx to bcwIdx, indentation fixes
* Fix #1557: change dependencies for writing ccv\_max\_luminance\_value and ccv\_avg\_luminance\_value
* Fix #1551: check coefficient range
* Add missing braces
* Avoid using std::vector and use static\_vector instead
* Fix variable names
* Avoid goto in SEIFilmGrainAnalyzer.cpp
* Reduce size of temporary memory for deblocking
* Precalculate fixed-point weights for MSE
* Y4M support at both encoder and decoder (#206)
* Fix variable names
* Fix variable names
* Consistently use helper functions and define them inline
* Don't use reference for simple bool parameters
* Clean up InterPrediction::xPredAffineBlk and associated functions
* Enable inlining of simple PelStorage methods
* Replace #define for constants in interpolation filter with constexpr
* Use constexpr instead of const for compile time constants
* Branchless version of Area::contains
* Clean up filter management for MC interpolation
* Clean up SIMD interpolation filter code and add support of horizontal 6-tap filter

***CTC Performance***

The following tables still show **VTM 15.0** performance over **HM 16.25** for a fair comparison with aligned tool configurations. The upcoming version HM 16.26 is expected to be aligned with VTM 17.0:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **All Intra Main10** |  |  |
|  |  |  | **Over HM-16.25** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | -29.03% | -32.17% | -34.07% | 1564% | 171% |
| Class A2 | -29.29% | -23.92% | -21.06% | 2483% | 183% |
| Class B | -21.73% | -26.96% | -30.76% | 2731% | 182% |
| Class C | -22.54% | -18.95% | -22.70% | 3923% | 187% |
| Class E | -25.75% | -25.91% | -24.45% | 2209% | 171% |
| **Overall** | -25.06% | -25.37% | -26.85% | 2563% | 179% |
| Class D | -18.46% | -13.31% | -13.41% | 4463% | 167% |
| Class F | -39.33% | -39.73% | -42.22% | 5299% | 171% |
|  |  |  |  |  |  |
|  |  |  | **Random access Main10** |  |  |
|  |  |  | **Over HM-16.25** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | -39.85% | -39.48% | -46.15% | 670% | 164% |
| Class A2 | -43.19% | -40.54% | -39.68% | 746% | 178% |
| Class B | -36.30% | -48.60% | -47.20% | 740% | 166% |
| Class C | -33.16% | -34.83% | -36.95% | 997% | 168% |
| Class E |  |  |  |  |  |
| **Overall** | -37.55% | -41.49% | -42.75% | 787% | 169% |
| Class D | -31.45% | -31.40% | -31.26% | 1125% | 165% |
| Class F | -45.76% | -49.18% | -50.10% | 580% | 147% |
|  |  |  |  |  |  |
|  |  |  | **Low delay B Main10** |  |  |
|  |  |  | **Over HM-16.25** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -29.17% | -34.97% | -32.34% | 761% | 166% |
| Class C | -25.87% | -17.39% | -17.85% | 910% | 164% |
| Class E | -28.93% | -33.29% | -26.46% | 371% | 136% |
| **Overall** | -28.01% | -28.69% | -26.04% | 675% | 157% |
| Class D | -25.08% | -12.54% | -12.25% | 961% | 168% |
| Class F | -40.20% | -41.56% | -41.87% | 509% | 136% |
|  |  |  |  |  |  |
|  |  |  | **Low delay P Main10** |  |  |
|  |  |  | **Over HM-16.25** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -33.74% | -37.63% | -34.94% | 697% | 175% |
| Class C | -27.71% | -17.50% | -17.86% | 831% | 175% |
| Class E | -32.18% | -37.21% | -31.09% | 358% | 147% |
| **Overall** | -31.34% | -30.82% | -28.28% | 626% | 167% |
| Class D | -26.32% | -11.70% | -10.99% | 876% | 170% |
| Class F | -39.97% | -41.10% | -41.48% | 542% | 143% |

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 17.0** performance compared to **VTM 16.0**:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **All Intra Main10** |  |  |
|  |  |  | **Over VTM-16.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | 0.00% | 0.00% | 0.00% | 99% | 98% |
| Class A2 | 0.00% | 0.00% | 0.00% | 100% | 96% |
| Class B | 0.00% | 0.00% | 0.00% | 93% | 94% |
| Class C | 0.00% | 0.00% | 0.00% | 100% | 97% |
| Class E | 0.00% | 0.00% | 0.00% | 103% | 97% |
| **Overall** | 0.00% | 0.00% | 0.00% | 98% | 96% |
| Class D | 0.00% | 0.00% | 0.00% | 99% | 100% |
| Class F | 0.00% | 0.00% | 0.00% | 102% | 97% |
|  |  |  |  |  |  |
|  |  |  | **Random access Main10** |  |  |
|  |  |  | **Over VTM-16.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | -1.06% | -1.76% | -1.63% | 111% | 98% |
| Class A2 | -0.95% | -1.45% | -1.31% | 111% | 102% |
| Class B | -1.19% | -1.84% | -1.48% | 110% | 100% |
| Class C | -1.04% | -0.89% | -1.17% | 108% | 96% |
| Class E |  |  |  |  |  |
| **Overall** | -1.08% | -1.49% | -1.39% | 110% | 99% |
| Class D | -0.68% | -0.94% | -1.03% | 109% | 99% |
| Class F | -0.49% | -0.09% | 0.11% | 106% | 96% |
|  |  |  |  |  |  |
|  |  |  | **Low delay B Main10** |  |  |
|  |  |  | **Over VTM-16.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -1.23% | -3.52% | -2.93% | 105% | 101% |
| Class C | -2.88% | -4.32% | -3.98% | 106% | 100% |
| Class E | -2.30% | -4.18% | -5.17% | 95% | 88% |
| **Overall** | -2.05% | -3.95% | -3.84% | 103% | 97% |
| Class D | -2.36% | -4.65% | -4.65% | 105% | 98% |
| Class F | -2.84% | -3.01% | -2.90% | 100% | 94% |
|  |  |  |  |  |  |
|  |  |  | **Low delay P Main10** |  |  |
|  |  |  | **Over VTM-16.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -0.67% | -0.70% | -0.77% | 106% | 103% |
| Class C | -0.54% | -0.28% | -0.25% | 106% | 102% |
| Class E | -0.73% | -2.26% | -2.44% | 95% | 90% |
| **Overall** | -0.64% | -0.95% | -1.02% | 103% | 99% |
| Class D | -0.45% | -0.55% | -0.59% | 106% | 100% |
| Class F | -0.16% | -0.02% | 0.26% | 100% | 98% |

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

JVET-AA0194 reports coding results for HDR CTC and proposes to verify performance impact for non-normative changes also HDR CTC. The AHG recommends reviewing the contribution.

***Issues in VTM 16 affecting conformance***

The following issues in VTM master branch (Jan. 11, 2022) affect conformance:

* Missing HLS features (see sections below)

There are no known issues in VTM that affect processing of current VVC v1 conformance bitstreams.

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

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

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

**HM related activities**

HM 16.26 is expected to be tagged during or shortly after the 27th JVET meeting. Changes include so far:

* JVET-Y0155: Fixes for motion-compensated temporal prefilter
* JVET-Y0105: An improved VVC rate control scheme
* JVET-Y0077: Block Importance Mapping
* Silence compiler warning
* Update copyright date to include 2022
* fix ticket #1516

The following MRs are pending:

* [JCTVC-AD0021(JVET-T0056) SEI manifest & SEI prefix indication](https://vcgit.hhi.fraunhofer.de/jvet/HM/-/merge_requests/51)
* Mark the current picture as short-term ref (for SCM)

The HM SCC branch was not updated for recent HM versions. It appears possible to merge the SCC branch into the main HM branch. 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.

**360Lib related activities**

The latest 360Lib software (360Lib-13.2) can be found at <https://vcgit.hhi.fraunhofer.de/jvet/360lib>

The following table is for the projection formats comparison using VTM-17.0 according to 360-degree video CTC (JVET-U2012) compared to that using VTM-16.0 (VTM-16.0 as anchor).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **PERP: VTM-17.0 over VTM-17.0** | | | | | |
|  | **End-to-end**  **WS-PSNR** | | | **End-to-end**  **S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | -0.48% | -2.87% | -2.63% | -0.49% | -2.87% | -2.63% |
| Class S2 | -0.65% | -2.56% | -2.48% | -0.65% | -2.56% | -2.49% |
| **Overall** | -0.55% | -2.74% | -2.57% | -0.56% | -2.75% | -2.58% |

The following table compares generalized cubemap (GCMP) coding and padded equi-rectangular projection (PERP) coding using VTM-17.0 (PERP as anchor).

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

The following tables are for PERP and GCMP coding comparison between VTM-17.0 and HM-16.22 (HM as anchor), respectively.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **VTM-17.0 PERP Over HM-16.22 PERP (anchor)** | | | | | |
|  | **End-to-end WS-PSNR** | | | **End-to-end S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | -30.91% | -38.80% | -41.29% | -30.90% | -38.85% | -41.29% |
| Class S2 | -36.88% | -37.35% | -39.69% | -36.87% | -37.39% | -39.75% |
| **Overall** | -33.30% | -38.22% | -40.65% | -33.29% | -38.27% | -40.67% |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **VTM-17.0 GCMP Over HM-16.22 PCMP (anchor)** | | | | | |
|  | **End-to-end WS-PSNR** | | | **End-to-end S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | -35.61% | -40.88% | -42.79% | -35.57% | -40.82% | -42.74% |
| Class S2 | -39.84% | -39.56% | -41.51% | -39.85% | -39.56% | -41.55% |
| **Overall** | -37.30% | -40.35% | -42.28% | -37.28% | -40.31% | -42.27% |

**SCM related activities**

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

**SHM related activities**

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

**HTM related activities**

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

The next release will include the following changes:

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

**HDRTools related activities**

There had not been any updates of HDRTools.

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

**JM, JSVM, JMVM related activities**

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

**Bug tracking**

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

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

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

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

Bug tracking for HDRTools is located at:

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

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

**Software repositories**

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

**CTC alignment and merging**

JVET-Z2011 was produced as JVET output document containing the merged VTM and HM CTC for SDR/WCG video.

Merging of HM RExt CTC into the appropriate VVC CTC was investigated, but proper comparable HM configuration files were not yet available by the beginning of this meeting. It is planned to provide an input to the 28th meeting.

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

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

This should be corrected to:

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

**Recommendations**

The AHG recommends to:

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

It is noted that JVET-AA0130 proposes a draft for merging CTC in range extensions

It was reported by an expert that the HDR results from JVET-AA0194 had been cross-verified. It is suggested to include in an updated version of the AHG3 report.

[JVET-AA0004](https://jvet-experts.org/doc_end_user/current_document.php?id=11830) JVET AHG report: Test material and visual assessment (AHG4) [V. Baroncini, T. Suzuki, M. Wien, S. Liu, G. Martin-Cocher, A. Segall, P. Topiwala, S. Wenger, J. Xu, Y. Ye (AHG chairs)]

1. Activities
   1. *Test sequences*

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

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

AhG meeting on development of a gaming-type CTC class was held. The results are reported in JVET-AA0046. Based on the input contributions to this meeting (JVET-AA123), the suggested content features are covered as follows:

Matrix of intended content features:

|  |  |  |
| --- | --- | --- |
| Feature | 3840×2160 | 1920×1080 |
| 1st-person perspective | Fontainebleau\_Cinematic\_3840x2160\_60p\_420\_10\_bits | Darktree\_1920x1080\_60p\_420\_10\_bits  Fontainebleau\_Cinematic\_1920x1080\_60p\_420\_10\_bits  Fontainebleau\_FPV\_1920x1080\_60p\_420\_10\_bits |
| 3rd-person perspective | 3D\_game\_kit\_3840x2160\_60p\_420\_10bits  Racing \_3840x2160\_60p\_420\_10bits | 3D\_game\_kit\_Level-1\_1920x1080\_60p\_420\_10bits  3D\_game\_kit\_Level-2\_1920x1080\_60p\_420\_10\_bits |
| Other perspectives (name them) | -/- | -/- |
| Scores | Racing \_3840x2160\_60p\_420\_10bits (Speed) | -/- |
| Chat | -/- | -/- |
| Graphical elements | -/- | -/- |
| Persons / characters | 3D\_game\_kit\_3840x2160\_60p\_420\_10bits | 3D\_game\_kit\_Level-1\_1920x1080\_60p\_420\_10bits  3D\_game\_kit\_Level-2\_1920x1080\_60p\_420\_10\_bits  Darktree\_1920x1080\_60p\_420\_10\_bits |
| Vehicles | Racing \_3840x2160\_60p\_420\_10bits | Darktree\_1920x1080\_60p\_420\_10\_bits |
| Naturalistic rendering | Fontainebleau\_Cinematic\_3840x2160\_60p\_420\_10\_bits | Darktree\_1920x1080\_60p\_420\_10\_bits  Fontainebleau\_Cinematic\_1920x1080\_60p\_420\_10\_bits  Fontainebleau\_FPV\_1920x1080\_60p\_420\_10\_bits |
| Synthetic rendering | 3D\_game\_kit\_3840x2160\_60p\_420\_10bits  Racing \_3840x2160\_60p\_420\_10bits | 3D\_game\_kit\_Level-1\_1920x1080\_60p\_420\_10bits  3D\_game\_kit\_Level-2\_1920x1080\_60p\_420\_10\_bits |

1. Related contributions

|  |  |  |
| --- | --- | --- |
| [JVET-AA0046](https://jvet-experts.org/doc_end_user/current_document.php?id=11721) | [AhG4] Report on AhG4 meeting on development of a gaming-type CTC class | [M. Wien (AHG chair)](mailto:wien@lfb.rwth-aachen.de) |
| [JVET-AA0123](https://jvet-experts.org/doc_end_user/current_document.php?id=11799) | [AHG-7] Update on gaming sequences from InterDigital | [T. Poirier](mailto:Tangi.Poirier@InterDigital.com), [S. Puri](mailto:saurabh.puri@interdigital.com), [G. Martin-Cocher](mailto:Gaelle.Martin-Cocher@InterDigital.com), [E. Faivre d'Arcier (InterDigital)](mailto:Etienne.FaivredArcier@InterDigital.com) |

Note: JVET-AA0123 is considered to be relevant for both, AHG4 and AHG7.

1. Recommendations

The AHG recommends:

* To study the new test sequences proposed in JVET-AA0123.
* To discuss the definition of a new CTC class for gaming type content and contribution JVET-AA0123 in a joint session with SC29/AG5.
* 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 review the set of newly proposed test sequences for potential inclusion in Common Test conditions and for the verification tests.
* To continue to collect new test sequences available for JVET with licensing statement.

It is reported by an expert that investigations on scalable VVC (which could be an interesting item for another verification test) are ongoing. Participation of more companies in such efforts would be desirable. Issue a recommendation?

[JVET-AA0005](https://jvet-experts.org/doc_end_user/current_document.php?id=11831) JVET AHG report: Conformance testing (AHG5) [D. Rusanovskyy, I. Moccagatta, F. Bossen, K. Kawamura, T. Hashimoto, H.-J. Jhu, K. Sühring, Y. Yu (AHG chairs)]

1. **Timeline**

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

* **VVCv1 conformance:**
  + ISO/IEC FDIS 23090-15 issued from 2021-10 meeting, pending FDIS ballot
  + H.266.1 V1 Consent 2022-01, last call to end 2022-04-28
* **VVCv2 conformance:**
  + ISO/IEC 23090-15/Amd.1 CDAM: 2021-10
  + ISO/IEC 23090-15/Amd.1 DAM: 2022-01
  + ISO/IEC 23090-15/Amd.1 FDAM: 2022-07
  + ISO/IEC 23090-15/Amd.1 AMD: 2023-01
  + H.266.1 V2 Consent 2022-10

1. **Status on bitstream submission**

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

* conformance bitstreams for VVC:
  + 104 bitstream categories have been identified
  + At least one bitstream has been submitted in each identified category
  + 283 total bitstreams have been provided, checked, and made available
  + One bitstream has been re-generated. No other changes between 26th and 27th meeting.
* conformance bitstreams for VVC operation range extensions:
  + 57 bitstream categories have been identified
  + Volunteers have been identified to generate bitstreams in all categories
  + Volunteers have been identified to cross-check bitstreams in all categories
  + All bitstream descriptions have been provided
  + 127 bitstreams of 56 identified categories have been cross-checked and uploaded
  + 3 packages have been re-generated
  + 1 bitstream of 1 identified category must still be re-generated.

1. **Activities and Discussion**

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

VVC activities:

One bitstream (HRD\_A\_Fujitsu) has been re-generated because it causes a CPB underflow in Type II conformance testing. The recommendation is to include the fix in the v2 revision. The re-generated package (HRD\_A\_Fujitsu\_4.zip) is available in <https://www.itu.int/wftp3/av-arch/jvet-site/bitstream_exchange/VVCv2/under_test/VTM-17.0/>

There are not currently any known issues with the other provided VVC conformance bitstream packages. All provided bitstreams can be decoded using VTM17.0 w/o and with range extension support.

A version of the packages with removed company names from filenames and removed email addresses from text file descriptions has been generated as requested by ISO and is available in https://www.itu.int/wftp3/av-arch/jvet-site/bitstream\_exchange/VVC/FDIS\_r1.

VVC operation range extensions activities:

Three packages (16b444SPepp\_A\_Sharp, 16b444Iepp\_A\_Sharp, 16b444epp\_A\_Sharp) have been re-generated following last meeting decision to keep yuv.md5 file naming convention for all tests to avoid introducing a new file naming convention and confusion.

One bitstream of 1 identified category (12b420SPvvc1\_A\_KDDI\_1.zip) does still need to be re-generated following VTM16.0 changes. That is the only bitstream in its category, so we solicit its re-generation.

All provided bitstreams can be decoded using VTM17.0.

A version of the DAM packages with removed company names from filenames and removed email addresses from text file descriptions has been generated as requested by ISO and is available in <https://www.itu.int/wftp3/av-arch/jvet-site/bitstream_exchange/VVCv2/under_test/VTM-14.0_r1>.

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**

There is one related contribution, the editors’ update to the conformance testing specification for VVC operation range extension

JVET-AA0109, Editors' update on conformance testing for VVC operation range extensions [D. Rusanovskyy, I. Tomohiro, H.-J. Jhu, I. Moccagatta, Y. Yu]

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:



1. **Recommendations**

The AHG recommends the following:

* Solicit the re-generation of 12b420SPvvc1\_A\_KDDI.
* Include the re-generated HRD\_A\_Fujitsu (HRD\_A\_Fujitsu\_4.zip) in v2 revision.
* Port VVC operation range extensions editorial changes made to ISO/IEC version of the document to JVET document.
* Review Editors' update on conformance testing for VVC operation range extensions.

New version in ITU-T in October, also incorporating the v1 bit stream correction. DAM ballot will close before the January meeting (Nov. 15), an FDIS of a new edition could be issued to take the same action on ISO/IEC side. Issue an errata report (also WG 5 document) that could be referred in the ballot responses.

The changes made in JVET-AA0109 have already been included (as an editorial action) in the DAM ballot. Not necessary to issue an output document, as the changes are purely editorial.

The modified bit stream packages are already in the corresponding ftp site and will be submitted along with the final version.

[JVET-AA0006](https://jvet-experts.org/doc_end_user/current_document.php?id=11716) JVET AHG report: ECM software development (AHG6) [V. Seregin, J. Chen, F. Le Léannec, K. Zhang (AHG 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.

The following adopted aspects were integrated into ECM-4.0:

JVET-Z0131: Block vector difference binarization

JVET-Z0127 (option 2): GPM intra-inter mode

JVET-Z0050 (Test 1.3b): CCLM slop adjustment, DIMD chroma with fusion mode

JVET-Z0054 (Test 2.1b): Block level reference picture reordering

JVET-Z0136 (test EE2 - 2.2a): Enhanced bi-prediction with out-of-boundary prediction samples

JVET-Z0061 (test EE2 – 2.3): Template matching based OBMC

JVET-Z0117 (version where no switching between 4-tap and 6-tap filters): 6-tap chroma interpolation filters

JVET-Z0118: GDR

JVET-Z0056 (test EE2 - 2.4): Template matching based reordering for GPM split modes

JVET-Z0139 (version EE2-2.7c): History based and non-adjacent affine candidates

JVET-Z0153 (test 3.2): IBC reference area extension

JVET-Z0075 (test 3.3): Enlarged HMVP table for IBC

JVET-Z0084 (test 3.4): Template Matching for IBC

JVET-Z0160 (test 3.5b): Replacement of zero-padding candidates

JVET-Z0135 (test 4.3b): Temporal CABAC, weighted states, windows adjustment

JVET-Z0085: Enabling AMVP-merge mode when DMVD is off

JVET-Z0102: No ARMC for Zero candidates for Regular, TM and BM merge modes

JVET-Z0105 (not CTC): Virtual boundary handling for in-loop filters

JVET-Z0083 (solution 1): Fix for MHP parsing condition

JVET-Z0072: Enhanced reference picture structures (encoder)

JVET-Z0099: Enable RDO-DBF in RA, LDB, and LDP (encoder)

JVET-Z0150: Print memory usage

JVET-Z0067/MR105: RPR fixes

A fix for invalid memory access when EncDbOpt=1 was also ported into ECM software from the latest VTM software.

ECM-5.0 was tagged on June 2 2022.

GDR implementation requires software optimization, so GDR macro was disabled in ECM-5.0 release. It is planned to enable it in ECM-5.1.

***CTC Performance***

In this section, ECM-5.0 test results following ECM CTC configuration descried in JVET-Y2017 are summarized.

The below tables show ECM-5.0 performance over ECM-4.0 anchor.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | |
|  | **Over ECM-4.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -0.26% | -2.34% | -2.28% | 109% | 104% |
| Class A2 | -0.20% | -2.09% | -1.52% | 113% | 105% |
| Class B | -0.23% | -2.29% | -2.55% | 110% | 107% |
| Class C | -0.24% | -1.78% | -1.83% | 111% | 104% |
| Class E | -0.27% | -1.79% | -1.63% | 110% | 108% |
| **Overall** | -0.24% | -2.07% | -2.02% | 111% | 106% |
| Class D | -0.21% | -1.43% | -1.30% | 111% | 106% |
| Class F | -5.49% | -6.34% | -6.50% | 135% | 104% |
| Class TGM | -11.42% | -11.32% | -11.10% | 125% | 101% |
|  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | |
|  | **Over ECM-4.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -2.14% | -2.79% | -2.97% | 123% | 113% |
| Class A2 | -2.27% | -4.15% | -3.99% | 124% | 118% |
| Class B | -2.72% | -4.63% | -4.96% | 122% | 114% |
| Class C | -2.58% | -5.42% | -5.20% | 124% | 116% |
| Class E |  |  |  |  |  |
| **Overall** | -2.48% | -4.38% | -4.43% | 123% | 115% |
| Class D | -2.71% | -6.45% | -6.98% | 120% | 114% |
| Class F | -5.27% | -6.80% | -7.06% | 126% | 107% |
| Class TGM | -12.55% | -12.67% | -12.89% | 117% | 102% |
|  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | |
|  | **Over ECM-4.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -2.87% | -4.83% | -4.61% | 122% | 113% |
| Class C | -4.92% | -8.15% | -7.89% | 130% | 118% |
| Class E | -3.82% | -5.26% | -5.70% | 121% | 110% |
| **Overall** | -3.79% | -6.04% | -5.97% | 124% | 114% |
| Class D | -4.43% | -7.91% | -7.81% | 127% | 125% |
| Class F | -7.08% | -7.91% | -8.99% | 129% | 114% |
| Class TGM | -14.86% | -15.50% | -15.64% | 119% | 105% |
|  |  |  |  |  |  |
|  | **Low delay P Main 10** | | | | |
|  | **Over ECM-4.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -1.83% | -2.32% | -2.41% | 118% | 112% |
| Class C | -1.60% | -3.77% | -3.27% | 117% | 112% |
| Class E | -1.94% | -2.43% | -2.21% | 115% | 109% |
| **Overall** | -1.78% | -2.83% | -2.65% | 117% | 111% |
| Class D | -2.63% | -4.06% | -3.85% | 117% | 114% |
| Class F | -4.70% | -6.17% | -6.40% | 131% | 111% |
| Class TGM | -9.31% | -9.36% | -9.82% | 122% | 103% |

Next tables show ECM-5.0 performance over VTM-11.0ecm anchor, the software is located at <https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tree/VTM11_ANC>**.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | |
|  | **Over VTM-11.0ecm5** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -6.84% | -14.30% | -19.53% | 390% | 266% |
| Class A2 | -6.42% | -15.04% | -16.00% | 380% | 259% |
| Class B | -5.92% | -16.12% | -15.78% | 433% | 276% |
| Class C | -7.11% | -10.83% | -11.38% | 424% | 254% |
| Class E | -7.92% | -13.65% | -14.49% | 401% | 277% |
| **Overall** | -6.75% | -14.05% | -15.25% | 409% | 267% |
| Class D | -5.92% | -9.33% | -8.84% | 421% | 268% |
| Class F | -15.79% | -20.74% | -20.14% | 355% | 270% |
| Class TGM | -26.19% | -28.03% | -27.29% | 317% | 275% |
|  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | |
|  | **Over VTM-11.0ecm5** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -16.80% | -18.70% | -24.21% | 436% | 536% |
| Class A2 | -17.87% | -24.31% | -26.27% | 415% | 626% |
| Class B | -15.58% | -23.66% | -22.40% | 432% | 550% |
| Class C | -17.51% | -21.09% | -20.65% | 454% | 555% |
| Class E |  |  |  |  |  |
| **Overall** | -16.80% | -22.11% | -23.07% | 435% | 563% |
| Class D | -18.49% | -22.02% | -22.71% | 444% | 579% |
| Class F | -18.31% | -23.02% | -23.03% | 403% | 371% |
| Class TGM | -24.41% | -28.17% | -28.50% | 425% | 299% |
|  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | |
|  | **Over VTM-11.0ecm5** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -14.25% | -27.64% | -25.57% | 396% | 419% |
| Class C | -15.40% | -24.03% | -23.37% | 449% | 434% |
| Class E | -12.99% | -18.62% | -20.64% | 400% | 317% |
| **Overall** | -14.32% | -24.18% | -23.61% | 414% | 396% |
| Class D | -17.44% | -25.35% | -24.26% | 416% | 489% |
| Class F | -16.96% | -24.65% | -24.21% | 399% | 328% |
| Class TGM | -23.33% | -30.18% | -29.91% | 415% | 294% |
|  |  |  |  |  |  |
|  | **Low delay P Main 10** | | | | |
|  | **Over VTM-11.0ecm5** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -14.09% | -28.08% | -26.41% | 355% | 418% |
| Class C | -15.67% | -23.82% | -23.15% | 333% | 419% |
| Class E | -13.11% | -18.84% | -20.34% | 344% | 317% |
| **Overall** | -14.37% | -24.35% | -23.81% | 345% | 390% |
| Class D | -18.52% | -25.59% | -24.04% | 312% | 431% |
| Class F | -17.47% | -25.22% | -24.88% | 372% | 308% |
| Class TGM | -23.20% | -30.41% | -30.28% | 418% | 290% |

The Excel files with the complete ECM 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.

It is noted that some of the gains over ECM4 are due encoder optimization. In comparison to VTM, corresponding optimizations were implemented for fair comparison

.

[JVET-AA0007](https://jvet-experts.org/doc_end_user/current_document.php?id=11832) JVET AHG report: Low latency and constrained complexity (AHG7) [A. Duenas, T. Poirier, S. Liu, L. Wang, J. Xu (AHG chairs)]

1. **Related Contributions**

[JVET-AA0117](https://jvet-experts.org/doc_end_user/current_document.php?id=11793): **AHG-7: refining low delay configuration for cloud gaming**

This contribution proposes an update of the low delay configuration that was agreed in the last meeting ([JVET-Z0114](https://jvet-experts.org/doc_end_user/current_document.php?id=11561)) by further removing tools that inherently cause delays in the processing pipeline on the decoder.

Two set of refinements of the LLCC baseline configuration are proposed. For the first set, tools that use template matching are disabled. For the second set, tools that use template without refinement are further disabled in addition to the template matching tools.

[JVET-AA0123](https://jvet-experts.org/doc_end_user/current_document.php?id=11799): **[AHG-7] Update on gaming sequences from InterDigital**

This contribution proposes an update on the new gaming sequences initially proposed in JVET-Y0041.

1. **Recommendations**

The AHG recommends reviewing input contributions and:

* to coordinate with AhG4 so that new sequences corresponding to low delay and to low latency and controlled complexity scenarios are added to the JVET CTCs.
* to refine configurations for low latency and controlled complexity, taking in account some of the suggestions made on the reflector.

[JVET-AA0008](https://jvet-experts.org/doc_end_user/current_document.php?id=11833) JVET AHG report: High bit depth, high bit rate, and high frame rate coding (AHG8) [A. Browne, T. Ikai, D. Rusanovskyy, X. Xiu, Y. Yu (AHG chairs)]

1. **Activities**

Previously, the AHG has used the main JVET reflector, jvet@lists.rwth-aachen.de, with [AHG8] in message headers, however no correspondence marked as AHG8 was sent between the 26th and 27th meetings.

The major areas of work of the AHG in this meeting cycle have targeted a combined CTC for VTM and HM, and updates to the conformance test descriptions. For each of these areas of work a contribution has been registered for the 27th meeting.

1. **Contributions**

In total three contributions relevant to the study of high bit depth, high bit rate or high frame rate coding have been registered for the 27th meeting.

[JVET-AA0109](https://jvet-experts.org/doc_end_user/current_document.php?id=11785) “Editors' update on conformance testing for VVC operation range extensions”, D. Rusanovskyy (Qualcomm), T. Ikai (Sharp), H.-J. Jhu (Kwai), I. Moccagatta (Intel), Y. Yu (Oppo)

[JVET-AA0130](https://jvet-experts.org/doc_end_user/current_document.php?id=11806) “AHG8: Draft VTM and HM common test conditions for high bit depth and high bit rate video coding”, A. Browne (Sony), T. Ikai (Sharp), D. Rusanovskyy (Qualcomm), X. Xiu (Kwai), Y. Yu (Oppo)

[JVET-AA0194](https://jvet-experts.org/doc_end_user/current_document.php?id=11883) “On VTM results for HDR content”, A. Browne, S. Keating, K. Sharman (Sony)

1. **Benchmarks**

The two adoptions from the last meeting that affected the results as reported are JVET-Z0072 and JVET-Z0099.

***Standard QP Range***

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access** | | | | | | | | | |
|  | **Over VTM16.0** | | | | | | | | | |
|  |  |  | **wPSNR** |  |  | **PSNR** |  |  |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | 4.51% | -0.60% | -0.52% | 8.24% | 5.63% | -0.57% | 0.20% | -1.15% | 108% | 101% |
| Class H2 |  |  |  |  |  | -0.78% | -1.12% | -1.96% | 108% | 101% |
| **Overall** | 4.51% | -0.60% | -0.52% | 8.24% | 5.63% | -0.67% | -0.46% | -1.56% | 108% | 101% |
|  |  |  |  |  |  |  |  |  |  |  |
|  | **All Intra** | | | | | | | | | |
|  | **Over VTM16.0** | | | | | | | | | |
|  |  |  | **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% | 0.00% | 0.00% | 99% | 99% |
| Class H2 |  |  |  |  |  | 0.00% | 0.00% | 0.00% | 99% | 99% |
| **Overall** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 99% | 99% |

The differences between PSNR and wPSNR results for RA are discussed further in JVET-AA0194.

***Low QP Range***

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **HDR PQ** |  |  | **AI** |  |  |  |  |  |
|  |  |  | **Over VTM16.0** |  |  |  |  |  |
|  | wPsnrY | wPsnrU | wPsnrV | psnrY | psnrU | psnrV | EncT | DecT |
| PQ444 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 101% |
| PQ422 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 102% |
| **Overall** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 101% |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | **LDB** |  |  |  |  |  |
|  |  |  | **Over VTM16.0** |  |  |  |  |  |
|  | wPsnrY | wPsnrU | wPsnrV | psnrY | psnrU | psnrV | EncT | DecT |
| PQ444 | -0.40% | -0.35% | -0.97% | -0.33% | -0.27% | -0.80% | 107% | 99% |
| PQ422 | -0.26% | -0.68% | -1.05% | -0.19% | -0.57% | -0.89% | 107% | 100% |
| **Overall** | -0.33% | -0.52% | -1.01% | -0.26% | -0.42% | -0.85% | 107% | 99% |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | **RA** |  |  |  |  |  |
|  |  |  | **Over VTM16.0** |  |  |  |  |  |
|  | wPsnrY | wPsnrU | wPsnrV | psnrY | psnrU | psnrV | EncT | DecT |
| PQ444 | 0.00% | 0.00% | -0.01% | 0.00% | 0.00% | -0.01% | 106% | 100% |
| PQ422 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 106% | 100% |
| **Overall** | 0.00% | 0.00% | -0.01% | 0.00% | 0.00% | -0.01% | 106% | 100% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **HDR HLG** |  |  | **AI** |  |  |
|  |  |  | **Over VTM16.0** |  |  |
|  | psnrY | psnrU | psnrV | EncT | DecT |
| HLG444 | 0.00% | 0.00% | 0.00% | 100% | 101% |
| HLG422 | 0.00% | 0.00% | 0.00% | 100% | 101% |
| **Overall** | 0.00% | 0.00% | 0.00% | 100% | 101% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **LDB** |  |  |
|  |  |  | **Over VTM16.0** |  |  |
|  | psnrY | psnrU | psnrV | EncT | DecT |
| HLG444 | -0.28% | -0.54% | -0.84% | 104% | 98% |
| HLG422 | -0.40% | -0.79% | -1.20% | 106% | 98% |
| **Overall** | -0.34% | -0.66% | -1.02% | 105% | 98% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **RA** |  |  |
|  |  |  | **Over VTM16.0** |  |  |
|  | psnrY | psnrU | psnrV | EncT | DecT |
| HLG444 | 0.00% | -0.01% | -0.01% | 105% | 101% |
| HLG422 | 0.00% | 0.00% | 0.00% | 105% | 100% |
| **Overall** | 0.00% | 0.00% | 0.00% | 105% | 101% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **SVT RGB** |  |  | **AI** |  |  |
|  |  |  | **Over VTM16.0** |  |  |
|  | psnrG | psnrB | psnrR | EncT | DecT |
| SVT16 | 0.00% | 0.00% | 0.00% | 100% | 100% |
| SVT12 | 0.00% | 0.00% | 0.00% | 100% | 100% |
| **Overall** | 0.00% | 0.00% | 0.00% | 100% | 100% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **LDB** |  |  |
|  |  |  | **Over VTM16.0** |  |  |
|  | psnrG | psnrB | psnrR | EncT | DecT |
| SVT16 | 0.06% | -0.10% | -0.10% | 105% | 100% |
| SVT12 | -0.03% | 0.01% | 0.02% | 102% | 99% |
| **Overall** | 0.01% | -0.04% | -0.04% | 103% | 100% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **RA** |  |  |
|  |  |  | **Over VTM16.0** |  |  |
|  | psnrG | psnrB | psnrR | EncT | DecT |
| SVT16 | 0.00% | 0.00% | 0.00% | 103% | 102% |
| SVT12 | 0.02% | 0.01% | 0.01% | 101% | 101% |
| **Overall** | 0.01% | 0.00% | 0.00% | 102% | 102% |

1. **Recommendations**

The AHG recommends the following:

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

[JVET-AA0009](https://jvet-experts.org/doc_end_user/current_document.php?id=11834) JVET AHG report: SEI message studies (AHG9) [S. McCarthy, Y.-K. Wang, T. Chujoh, S. Deshpande, C. Fogg, P. de Lagrange, G. J. Sullivan, A. Tourapis, S. Wenger (AHG chairs)]

1. **Related contributions**

A total of 14 contributions are identified relating to the mandates of AHG9. Some contributions relate to more than one mandate. “AHG9” is part of the title of one contribution, JVET-AA0099, which appears to relate to non-SEI high level syntax.

The number of contributions relating to each mandate is as follows

* 11 contributions relate to the mandate to study the SEI messages in VSEI, VVC, HEVC, and AVC;
  + 8 contributions relate to the neural-network post-filter SEI messages
  + 1 contribution relates to the film grain characteristics SEI message
  + 1 contribution relates to the decoded picture hash SEI message
  + 1 contribution relates to the alpha channel information SEI message
* 1 contribution relates to the mandate to collect software and showcase information for SEI messages, including encoder and decoder implementations and bitstreams for demonstration and testing;
* 4 contributions relate to the mandate to identify potential needs for additional SEI messages;
* 1 contribution relates to the mandate to investigate the possible need of mandatory post processing in the context of SEI messages; and
* 0 contributions relate to the mandate to study SEI messages defined in HEVC and AVC for potential use in the VVC context.

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**

[JVET-AA0054](https://jvet-experts.org/doc_end_user/current_document.php?id=11730) AHG9: On Neural-network Post-filter Characteristics SEI Message [S. Deshpande (Sharp)]

[JVET-AA0055](https://jvet-experts.org/doc_end_user/current_document.php?id=11731) AHG9: Comments on Neural-network Post-filter Characteristics SEI Message [S. Deshpande, A. Sidiya (Sharp)]

[JVET-AA0056](https://jvet-experts.org/doc_end_user/current_document.php?id=11732) AHG9: On syntax gating in the neural-network post-filter characteristics SEI message [M. M. Hannuksela, F. Cricri, M. Santamaria (Nokia), T. Chujoh, Y. Yasugi, T. Ikai (Sharp), S. McCarthy, A. Arora, T. Shao, P. Yin, T. Lu, F. Pu, W. Husak (Dolby)]

[JVET-AA0067](https://jvet-experts.org/doc_end_user/current_document.php?id=11743) AHG9: Some specification improvements for neural-network post-filter characteristics SEI message [T. Chujoh, Y. Yasugi, T. Ikai (Sharp), M. Hannuksela, F. Cricri (Nokia), S. McCarthy, A. Arora, T. Shao, P. Yin, T. Lu, F. Pu, W. Husak (Dolby)]

[JVET-AA0083](https://jvet-experts.org/doc_end_user/current_document.php?id=11759) AHG9: NNR post-filter SEI message extension for flexible decoding capabilities [F. Galpin, T. Dumas, P. Bordes, E. François (InterDigital)]

(JVET-AA0083 also relates to the mandate to collect software and showcase information )

[JVET-AA0100](https://jvet-experts.org/doc_end_user/current_document.php?id=11776) AHG9: On auxiliary input and separate colour description in the neural-network post-filter characteristics SEI message [T. Shao, A. Arora, P. Yin, S. McCarthy, T. Lu, F. Pu, W. Husak (Dolby), M. M. Hannuksela, F. Cricri, M. Santamaria Gomez (Nokia), T. Chujoh, Y. Yasugi, T. Ikai (Sharp)]

[JVET-AA0101](https://jvet-experts.org/doc_end_user/current_document.php?id=11777) AHG9: On processing order in the neural-network post-filter activation SEI message [T. Shao, A. Arora, P. Yin, S. McCarthy, T. Lu, F. Pu, W. Husak (Dolby)]

[JVET-AA0145](https://jvet-experts.org/doc_end_user/current_document.php?id=11821) AHG9: On decoupling neural-network post-filter activation SEI message [H.-B. Teo, J. Gao, C.-S. Lim, K. Abe, V. Drugeon (Panasonic)]

**Film grain characteristics SEI message**

[JVET-AA0052](https://jvet-experts.org/doc_end_user/current_document.php?id=11728) AHG13: On VSEI film grain profiles [Y. He, M. Karczewicz (Qualcomm), M. Radosavljević (Xiaomi)]

(JVET-AA0052 also relates to the mandate to investigate the possible need of mandatory post processing in the context of SEI messages)

**Decoded picture hash SEI message**

[JVET-AA0079](https://jvet-experts.org/doc_end_user/current_document.php?id=11755) AHG9: Decoded picture hash SEI message extension [P. Bordes, F. Galpin, P. DeLagrange, E. Francois (InterDigital)]

**Alpha channel information SEI message**

[JVET-AA0105](https://jvet-experts.org/doc_end_user/current_document.php?id=11781) AHG9: Metadata for display on transparent screens based on ACI SEI messages [E. Thomas, P. Andrivon, F. Le Léannec, M. Radosavljević, M.-L. Champel (Xiaomi)]

(JVET-AA0105 also relates to the mandate to identify potential needs for additional SEI messages)

***Collect software and showcase information for SEI messages, including encoder and decoder implementations and bitstreams for demonstration and testing***

[JVET-AA0083](https://jvet-experts.org/doc_end_user/current_document.php?id=11759) AHG9: NNR post-filter SEI message extension for flexible decoding capabilities [F. Galpin, T. Dumas, P. Bordes, E. François (InterDigital)]

(JVET-AA0083 also relates to the mandate to study the SEI messages in VSEI, VVC, HEVC, and AVC)

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

[JVET-AA0091](https://jvet-experts.org/doc_end_user/current_document.php?id=11767) AHG9: Resolution Change Information SEI message [V. Drugeon, K. Abe, T. Toma (Panasonic)]

[JVET-AA0102](https://jvet-experts.org/doc_end_user/current_document.php?id=11778) AHG9: SEI processing order SEI message [P. Yin, S. McCarthy, W. Husak, K. Konstantinos, T. Lu, F. Pu, A. Arora, T. Shao (Dolby)]

[JVET-AA0105](https://jvet-experts.org/doc_end_user/current_document.php?id=11781) AHG9: Metadata for display on transparent screens based on ACI SEI messages [E. Thomas, P. Andrivon, F. Le Léannec, M. Radosavljević, M.-L. Champel (Xiaomi)]

(JVET-AA0052 also relates to the mandate to study the SEI messages in VSEI, VVC, HEVC, and AVC)

[JVET-AA0110](https://jvet-experts.org/doc_end_user/current_document.php?id=11786) AHG9: SEI message with sample phase indication for consistent rendering [J. Samuelsson-Allendes, S. Deshpande (Sharp)]

***Investigate the possible need of mandatory post processing in the context of SEI messages***

[JVET-AA0052](https://jvet-experts.org/doc_end_user/current_document.php?id=11728) AHG13: On VSEI film grain profiles [Y. He, M. Karczewicz (Qualcomm), M. Radosavljević (Xiaomi)]

(JVET-AA0052 also relates to the mandate to study the SEI messages in VSEI, VVC, HEVC, and AVC)

***Non-SEI HLS***

[JVET-AA0099](https://jvet-experts.org/doc_end_user/current_document.php?id=11775) AHG9: On subpictures order [Hendry, S. Kim, S. Lee (LGE)]

***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.

***Recommendations***

The AHG recommends to:

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

[JVET-AA0010](https://jvet-experts.org/doc_end_user/current_document.php?id=11835) JVET AHG report: Encoding algorithm optimization (AHG10) [P. de Lagrange, A. Duenas, R. Sjöberg, A. Tourapis (AHG chairs)]

**Related *contributions***

A total of 4 contributions, not including cross-checks, are identified relating to AHG10, and summarized in the following sections.

***Neural network-based encoding optimization technologies***

**JVET-*AA0113* – EE1-1.6-related: RDO Considering Deep In-Loop Filter with SADL**

This contribution proposes an encoder optimization technique to include in the RDO loop the EE1-1.6 deep learning-based in-loop filtering method combined with deblocking.

To reduce complexity, RDO relies on simplified CNNs (with no filter selection), trained for the purpose, and is not applied for blocks larger than 64.

Partial results (C and D classes) show gain around 0.5% with 10% encoding time increase.

**JVET-*AA0122* – EE1-1.3: On BaseQP adjustment in CNNLF**

This contribution proposes a QP-RDO technique for EE1-1.4. Encoder tests -5, 0, +5 QP offsets for the EE1-1.4 NN-based loop filter. To reduce complexity, search can be restricted to one CTU out of 4, with the 3 neighbors reusing the same QP offset

Gain is around 0.3%, with ~5% encoding time increase.

***Encoding techniques of optimization for objective quality metrics***

**JVET-*AA0127* – EE2-1 related: Encoder optimization of EE2-1.2 and 1.3**

This contribution proposes encoder speed-ups for EE2-1.2 (gradient linear model) and EE2-1.3a (combination of convolutional cross-component model and gradient linear model). In consists of software optimization (removal of redundant operations), and early termination in RDO involving GLM: combination of GLM with MDLM is skipped if GLM alone has higher cost than CCLM.

Gains are mostly unchanged (0.8% RA), while the encoding time increase vs anchor is divided by 3.

**JVET-*AA0129* – Non-EE2/AHG10: Improved inter hash RDO considering OBMC off in ECM5.0**

This contribution focusses on screen content, and proposes to consider OBMC-off hash-based ME in RDO, in addition to OBMC-on hash-based ME; early skips are implemented to control complexity.

Gains for TGM is around 1% (RA) and 1.5% (LD), while encoding time is not impacted.

***Recommendation***

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

[JVET-AA0011](https://jvet-experts.org/doc_end_user/current_document.php?id=11836) JVET AHG report: Neural network-based video coding (AHG11) [E. Alshina, S. Liu, A. Segall, F. Galpin, J. Pfaff, S. S. Wang, Z. Wang, M. Wien, P. Wu, J. Xu (AHG chairs)]

1. **Activities**

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

***EE Coordination***

The AHG finalized, conducted and discussed the EE on NN based video coding. The final version of the EE description was uploaded to the document repository on May 17, 2022.

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

|  |  |  |
| --- | --- | --- |
| JVET-AA0023 | EE1: Summary of Exploration Experiments on Neural Network-based Video Coding | E. Alshina, W. Chen, F. Galpin, Y. Li, Z. Ma, H. Wang, L. Wang |

***Anchor Encoding***

Anchors for the NN-based video coding activity were unchanged from those released on August 2, 2021. The anchors were also made available on the Git repository used for the AHG activity: https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/nnvc-ctc/-/tree/master.

***Study and Maintain SADL***

The SADL library was updated during the AHG period. Updates were announced on May 4, 2022; May 19, 2022; June 2, 2022; June 10, 2022; and July 12, 2022. The software is available at https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/sadl.

Change logs include:

*Commit 43515f670d294367d70114ab3aaac62dc9179e9a*

*Add dynamic quantization documentation*

*Commit 1e27c61ec38c0ef75c69482925662be251b417ca*

*Changes:*

*change the Add layer quantizer policy: use the smallest quantizer*

*instead of the 2nd input one. BiasAdd layer keeps the same policy.*

*Commit 7876017ebb70c67916da06d757ddac1974428b7a*

*Corrections:*

*- carry over quantizer values for Expand and transpose layers (JVET-Z-EE-1.6)*

*- add conditions for special tensors multiplication case (JVET-Z0082)*

*- added leakyRelu case in naive\_qunatization sample code (JVET-Z-EE-1.5)*

*Improvements:*

*- SSE42 for convolution in mod8 case (speedup JVET-Z0082)*

*Commit 29b68a048bc3ae968b8d1aa3e08319d5d37bd4e3*

*New:*

*- naive quantization example*

*- associated documentation*

*Improvements*

*- check on quantizer values for some layers*

*- better inference debug info*

*- in count\_mac: also simulate an overflow*

*Commit 3d258f089761b1c906e94c51ec54c8ffd788a85c*

*New:*

*- split sample programs: count\_mac, debug\_model, sample\_xxx*

*- added clang-format file*

*Improvements:*

*- improve SIMD support for float (AVX512, conv1x1 etc.)*

*- go back to HWCD loop order. Some gains on big models.*

*- nbOutputs() method in Model*

*Corrections:*

*- init is now re-entrant: it is possible to do inference with multiple input sizes by calling init with the new size*

*- remove all MAC counters in SIMD paths + add warning message*

*Cleaning:*

*- clean the naming for conv2d (simd, dispatch etc.)*

*- clean messages for SIMD support*

*- clean macro for debug*

*- clean sample\_test.sh script*

*- clang format everything*

*- copyright year*

*- README update*

*- other minor cleanings*

***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 27th meeting is provided below:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Title** | **Common Test Conditions** | **Results** | | | **Training Data** | |
| **RA** | **LDB** | **AI** | **CTC** | **Additional** |
| **Loop Filter** | | | | | | | |
| [JVET-AA0080](file:////Users/asegall/Downloads/current_document.php%3fid=11756) | AHG11: Complexity reduction on neural-network loop filter | Yes | No | No | Yes | - | DIV2K |
| **Post Filtering** | | | | | | | |
| [JVET-AA0054](file:////Users/asegall/Downloads/current_document.php%3fid=11730) | AHG9: On Neural-network Post-filter Characteristics SEI Message | No | No | No | No | - | - |
| [JVET-AA0055](file:////Users/asegall/Downloads/current_document.php%3fid=11731) | AHG9: Comments on Neural-network Post-filter Characteristics SEI Message | No | No | No | No | - | - |
| [JVET-AA0056](file:////Users/asegall/Downloads/current_document.php%3fid=11732) | AHG9: On syntax gating in the neural-network post-filter characteristics SEI message | No | No | No | No | - | - |
| [JVET-AA0067](file:////Users/asegall/Downloads/current_document.php%3fid=11743) | AHG9: Some specification improvements for neural-network post-filter characteristics SEI message | No | No | No | No | - | - |
| [JVET-AA0083](file:////Users/asegall/Downloads/current_document.php%3fid=11759) | AHG9: NNR post-filter SEI message extension for flexible decoding capabilities | No | No | No | No | - | - |
| [JVET-AA0100](file:////Users/asegall/Downloads/current_document.php%3fid=11776) | AHG9: On auxiliary input and separate colour description in the neural-network post-filter characteristics SEI message | No | No | No | No | - | - |
| [JVET-AA0101](file:////Users/asegall/Downloads/current_document.php%3fid=11777) | AHG9: On processing order in the neural-network post-filter activation SEI message | No | No | No | No | - | - |
| [JVET-AA0145](file:////Users/asegall/Downloads/current_document.php%3fid=11821) | AHG9: On decoupling neural-network post-filter activation SEI message | No | No | No | No | - | - |
| **Super-Resolution** | | | | | | | |
| [JVET-AA0065](file:////Users/asegall/Downloads/current_document.php%3fid=11741) | AHG11: CNN Filter for Super-Resolution with RPR functionality in VVC | Yes | Yes | No | Yes | BVI-DVC | DIV2K |
| [JVET-AA0076](file:////Users/asegall/Downloads/current_document.php%3fid=11752) | AHG11: RPR-Based Super-Resolution Guided by Partition Information | Yes | Yes | No | Yes | BVI-DVC | DIV2K |
| [JVET-AA0084](file:////Users/asegall/Downloads/current_document.php%3fid=11760) | AHG11: Neural Network based Super Resolution for Video Coding Using Multiple Side Information | Yes | Yes | No | Yes | BVI-DVC, TVD | DIV2K |
| **Inter-Prediction** | | | | | | | |
| [JVET-AA0082](file:////Users/asegall/Downloads/current_document.php%3fid=11758) | AHG11: Deep Reference Frame Generation for Inter Prediction Enhancement | Yes | Yes | Yes | No | BVI-DVC | VIMEO |
| **End-to-End** | | | | | | | |
| [JVET-AA0063](file:////Users/asegall/Downloads/current_document.php%3fid=11739) | AHG11: A hybrid codec using E2E image coding combined with VVC video coding | No | - | - | - | - | - |

1. **Input contributions**

There are 43 input contriubtions related to the AHG mandates. Twenty-six of the contributions are part of the EE activity, while the remaining 17 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-AA0023 | EE1: Summary of Exploration Experiments on Neural Network-based Video Coding | E. Alshina, W. Chen, F. Galpin, Y. Li, Z. Ma, L. Wang |
| **EE Technology** | | |
| [JVET-AA0059](file:////Users/asegall/Downloads/current_document.php%3fid=11735) | EE1-3.1: Supplementary experiments based on JVET-Z0077 | [Q. Qin](mailto:kippqin@stu.xidian.edu.cn), [C. Jung (Xidian Univ.)](mailto:zhengzk@xidian.edu.cn), [D. Zou](mailto:zoudan@oppo.com), [M. Li (Oppo)](mailto:myron.li@oppo.com) |
| [JVET-AA0066](file:////Users/asegall/Downloads/current_document.php%3fid=11742) | EE1-1.7: Content-adaptive post-filter based on SADL inference | R. Yang, M. Santamaria, F. Cricri, H. Zhang, J. Lainema, R. G. Youvalari, M. M. Hannuksela (Nokia) |
| [JVET-AA0071](file:////Users/asegall/Downloads/current_document.php%3fid=11747) | EE1-2.1: A CNN-based Super Resolution Method Combined with GOP Level Adaptive Resolution | S. Peng, C. Fang, [D. Jiang](mailto:jiang_dong@dahuatech.com), [J. Lin](mailto:lin_jucai@dahuatech.com), X. Zhang (Dahua), [J. Nam](mailto:junghak.nam@lge.com), S. Yoo, J. Lim, S. Kim (LGE) |
| [JVET-AA0081](file:////Users/asegall/Downloads/current_document.php%3fid=11757) | EE1-1.2: NN intra model without attention, partitioning and boundary strength | [J. Ström](mailto:jacob.strom@ericsson.com), [D. Liu](mailto:du.liu@ericsson.com), [M. Damghanian](mailto:mitra.damghanian@ericsson.com), [K. Andersson](mailto:kenneth.r.andersson@ericsson.com), [Y. Li](mailto:yun.y.li@ericsson.com), [P. Wennersten](mailto:per.wennersten@ericsson.com), [R. Yu (Ericsson)](mailto:ruoyang.yu@ericsson.com) |
| [JVET-AA0085](file:////Users/asegall/Downloads/current_document.php%3fid=11761) | EE1-1.1: The Performance of Single-Model Filter Trained on the VTM and ECM Reconstruction | [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-AA0087](file:////Users/asegall/Downloads/current_document.php%3fid=11763) | EE1-1.4: Neural network based in-loop filter with 2 models | [L. Wang](mailto:liqiangwang@tencent.com), [S. Lin](mailto:barneylin@tencent.com), [X. Xu](mailto:xiaozhongxu@tencent.com), [S. Liu (Tencent)](mailto:shanl@tencent.com), [F. Galpin (InterDigital)](mailto:franck.galpin@interdigital.com) |
| [JVET-AA0088](file:////Users/asegall/Downloads/current_document.php%3fid=11764) | EE1-1.5: Neural network based in-loop filter with a single model | [L. Wang](mailto:liqiangwang@tencent.com), [S. Lin](mailto:barneylin@tencent.com), [X. Xu](mailto:xiaozhongxu@tencent.com), [S. Liu (Tencent)](mailto:shanl@tencent.com), [F. Galpin (InterDigital)](mailto:franck.galpin@interdigital.com) |
| [JVET-AA0111](file:////Users/asegall/Downloads/current_document.php%3fid=11787) | EE1-1.6: Deep In-Loop Filter With Fixed Point Implementation | [Y. Li](mailto:yue.li@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [J. Li](mailto:lijunru@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com), [H. Wang](mailto:hongtaow@qti.qualcomm.com), M. Coban, A.M. Kotra, M. Karczewicz (Qualcomm), [F. Galpin (InterDigital)](mailto:Franck.galpin@interdigital.com), [K. Andersson](mailto:kenneth.r.andersson@ericsson.com), J. Ström, D. Liu, R. Sjöberg (Ericsson) |
| [JVET-AA0122](file:////Users/asegall/Downloads/current_document.php%3fid=11798) | EE1-1.3: On BaseQP adjustment in CNNLF | [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) |
| **EE Related** | | |
| [JVET-AA0074](file:////Users/asegall/Downloads/current_document.php%3fid=11750) | [EE1-related] Lighter WCDANN: CNN Based In-Loop Filters | [H. Zhang](mailto:13227706628@163.com), [C. Jung (Xidian Univ.)](mailto:zhengzk@xidian.edu.cn), [D. Zou](mailto:zoudan@oppo.com), [M. Li (Oppo)](mailto:myron.li@oppo.com) |
| [JVET-AA0089](file:////Users/asegall/Downloads/current_document.php%3fid=11765) | EE1-related: More refinements on EE1-1.4 and EE1-1.5 | [L. Wang](mailto:liqiangwang@tencent.com), [S. Lin](mailto:barneylin@tencent.com), [X. Xu](mailto:xiaozhongxu@tencent.com), [S. Liu (Tencent)](mailto:shanl@tencent.com), [Z. Xie](mailto:xiezhihuang@oppo.com), [Y. Yu](mailto:yue.yu@oppo.com), [H. Yu](mailto:v-yuhaoping@oppo.com), [D. Wang (Oppo)](mailto:wangdong7@oppo.com) |
| [JVET-AA0090](file:////Users/asegall/Downloads/current_document.php%3fid=11766) | EE1-related: One luma model with IPB and skip for filtering intra and inter luma slices | [D. Liu](mailto:du.liu@ericsson.com), [J. Ström](mailto:jacob.strom@ericsson.com), [M. Damghanian](mailto:mitra.damghanian@ericsson.com), [P. Wennersten](mailto:per.wennersten@ericsson.com), [Y. Li (Ericsson)](mailto:yun.y.li@ericsson.com) |
| [JVET-AA0094](file:////Users/asegall/Downloads/current_document.php%3fid=11770) | EE1-related: Deep In-Loop Filter in EE1-1.6 with Adaptive Input Samples | [C. Zhou](mailto:chuan.zhou@vivo.com), [Z. Lv](mailto:zhuoyi.lv@vivo.com), [J. Zhang (vivo)](mailto:jirong.zhang@vivo.com), W. Chen, J. Guo, B. Ai (BJTU) |
| [JVET-AA0112](file:////Users/asegall/Downloads/current_document.php%3fid=11788) | EE1-1.6-related: 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-AA0113](file:////Users/asegall/Downloads/current_document.php%3fid=11789) | EE1-1.6-related: RDO Considering Deep In-Loop Filter with SADL | [J. Li](mailto:lijunru@bytedance.com), [Y. Li](mailto:yue.li@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com) |
| [JVET-AA0115](file:////Users/asegall/Downloads/current_document.php%3fid=11791) | EE1-1.6-related: ALF with Samples before Deep In-Loop Filter | [J. Li](mailto:lijunru@bytedance.com), [K. Zhang](mailto:zhangkai.video@bytedance.com), [Y. Li](mailto:yue.li@bytedance.com), [L. Zhang (Bytedance)](mailto:lizhang.idm@bytedance.com) |
| [JVET-AA0131](file:////Users/asegall/Downloads/current_document.php%3fid=11807) | EE1-related: CNN based in-loop filtering with large activation layer | [H. Wang](mailto:hongtaow@qti.qualcomm.com), [S. Eadie](mailto:seadie@qti.qualcomm.com), [M. Coban](mailto:mcoban@qti.qualcomm.com), [M. Karczewicz (Qualcomm)](mailto:martak@qti.qualcomm.com) |
| **Cross Checks** | | |
| [JVET-AA0172](file:////Users/asegall/Downloads/current_document.php%3fid=11860) | Cross-check of JVET-AA0081 (EE1-1.2: NN intra model without attention, partitioning and boundary strength) | [M. Santamaria](mailto:maria.santamaria_gomez@nokia.com), F. Cricri (Nokia) |
| [JVET-AA0174](file:////Users/asegall/Downloads/current_document.php%3fid=11862) | [EE1] Crosscheck of training stage for EE1-1.5 and EE1-1.6 tests | [J. Sauer](mailto:johannes.sauer@huawei.com), [B. Wang](mailto:biao.wang@huawei.com), [E.Alshina (Huawei)](mailto:elena.alshina@huawei.com) |
| [JVET-AA0178](file:////Users/asegall/Downloads/current_document.php%3fid=11866) | Crosscheck of JVET-AA0088 (EE1-1.5: Neural network based in-loop filter with a single model) | [H. Wang (Qualcomm)](mailto:hongtaow@qti.qualcomm.com) |
| [JVET-AA0179](file:////Users/asegall/Downloads/current_document.php%3fid=11867) | Crosscheck of JVET-AA0085 (EE1-1.1: The Performance of Single-Model Filter Trained on the VTM and ECM Reconstruction) | [C. Zhou (vivo)](mailto:chuan.zhou@vivo.com) |
| [JVET-AA0181](file:////Users/asegall/Downloads/current_document.php%3fid=11869) | Cross-check of JVET-AA0111 (EE1-1.6: Deep In-Loop Filter With Fixed Point Implementation) | [K. Lin](mailto:kailin@pku.edu.cn), [C. Jia](mailto:cmjia@pku.edu.cn), [S. Wang](mailto:sswang@pku.edu.cn) |
| [JVET-AA0185](file:////Users/asegall/Downloads/current_document.php%3fid=11873) | Crosscheck of JVET-AA0122 (EE1-1.3: On BaseQP adjustment in CNNLF) | [L. Wang (Tencent)](mailto:liqiangwang@tencent.com) |
| [JVET-AA0186](file:////Users/asegall/Downloads/current_document.php%3fid=11874) | Crosscheck of JVET-AA0131 (EE1-related: CNN based in-loop filtering with large activation layer) | [L. Wang (Tencent)](mailto:liqiangwang@tencent.com) |
| [JVET-AA0200](file:////Users/asegall/Downloads/current_document.php%3fid=11889) | Crosscheck of JVET-AA0094 (EE1-related: Deep In-Loop Filter in EE1-1.6 with Adaptive Input Samples) | [Y. Li](mailto:yue.li@bytedance.com) |

***Non-EE Input Contributions***

|  |  |  |
| --- | --- | --- |
| **Reporting** | | |
| [JVET-AA0011](file:////Users/asegall/Downloads/current_document.php%3fid=11836) | JVET AHG report: Neural network-based video coding (AHG11) | E. Alshina, S. Liu, A. Segall, F. Galpin, J. Pfaff, S. S. Wang, Z. Wang, M. Wien, P. Wu, J. Xu (AHG chairs) |
| [JVET-AA0047](file:////Users/asegall/Downloads/current_document.php%3fid=11722) | [AHG 11] Brief information about JPEG AI CfP status | [E. Alshina](mailto:elena.alshina@huawei.com), [J. Ascenso](mailto:joao.ascenso@lx.it.pt), [T. Ebrahimi](mailto:Touradj.Ebrahimi@epfl.ch), [F. Pereira](mailto:fp@lx.it.pt), [T. Richter](mailto:thorfdbgiis@gmail.com) |
| **Loop Filtering** | | |
| [JVET-AA0080](file:////Users/asegall/Downloads/current_document.php%3fid=11756) | AHG11: Complexity reduction on neural-network loop filter | [J. N. Shingala](mailto:jay.shingala@ittiam.com), S. Kadaramandalgi, A. Shyam (Ittiam), T. Shao, A. Arora, [P. Yin](mailto:pyin@dolby.com), F. Pu, T. Lu, S. McCarthy (Dolby) |
| **Post Filtering** | | |
| [JVET-AA0054](file:////Users/asegall/Downloads/current_document.php%3fid=11730) | AHG9: On Neural-network Post-filter Characteristics SEI Message | [S. Deshpande (Sharp)](mailto:sdeshpande@sharplabs.com) |
| [JVET-AA0055](file:////Users/asegall/Downloads/current_document.php%3fid=11731) | AHG9: Comments on Neural-network Post-filter Characteristics SEI Message | [S. Deshpande](mailto:sdeshpande@sharplabs.com), A. Sidiya (Sharp) |
| [JVET-AA0056](file:////Users/asegall/Downloads/current_document.php%3fid=11732) | AHG9: On syntax gating in the neural-network post-filter characteristics SEI message | [M. M. Hannuksela](mailto:miska.hannuksela@nokia.com), F. Cricri, M. Santamaria (Nokia), [T. Chujoh](mailto:chujoh.takeshi@sharp.co.jp), Y. Yasugi, T. Ikai (Sharp), [S. McCarthy](mailto:sean.mccarthy@dolby.com), A. Arora, T. Shao, P. Yin, T. Lu, F. Pu, W. Husak (Dolby) |
| [JVET-AA0067](file:////Users/asegall/Downloads/current_document.php%3fid=11743) | AHG9: Some specification improvements for neural-network post-filter characteristics SEI message | [T. Chujoh](mailto:chujoh.takeshi@sharp.co.jp), Y. Yasugi, T. Ikai (Sharp), [M. Hannuksela](mailto:miska.hannuksela@nokia.com), F. Cricri (Nokia), [S. McCarthy](mailto:sean.mccarthy@dolby.com), A. Arora, T. Shao, P. Yin, T. Lu, F. Pu, W. Husak (Dolby) |
| [JVET-AA0083](file:////Users/asegall/Downloads/current_document.php%3fid=11759) | AHG9: NNR post-filter SEI message extension for flexible decoding capabilities | [F. Galpin](mailto:franck.galpin@interdigital.com), [T. Dumas](mailto:thierry.dumas@interdigital.com), [P. Bordes](mailto:philippe.bordes@interdigital.com), [E. François (InterDigital)](mailto:edouard.francois@interdigital.com) |
| [JVET-AA0100](file:////Users/asegall/Downloads/current_document.php%3fid=11776) | AHG9: On auxiliary input and separate colour description in the neural-network post-filter characteristics SEI message | [T. Shao](mailto:tong.shao@dolby.com), [A. Arora](mailto:arjun.arora@dolby.ocm), [P. Yin](mailto:pyin@dolby.com), [S. McCarthy](mailto:sean.mccarthy@dolby.com), [T. Lu](mailto:taoran.lu@dolby.com), [F. Pu](mailto:fangjun.pu@dolby.com), [W. Husak (Dolby)](mailto:wjh@dolby.com), [M. M. Hannuksela](mailto:miska.hannuksela@nokia.com), F. Cricri, M. Santamaria Gomez (Nokia), [T. Chujoh](mailto:chujoh.takeshi@sharp.co.jp), Y. Yasugi, T. Ikai (Sharp) |
| [JVET-AA0101](file:////Users/asegall/Downloads/current_document.php%3fid=11777) | AHG9: On processing order in the neural-network post-filter activation SEI message | [T. Shao](mailto:tong.shao@dolby.com), [A. Arora](mailto:arjun.arora@dolby.ocm), [P. Yin](mailto:pyin@dolby.com), [S. McCarthy](mailto:sean.mccarthy@dolby.com), [T. Lu](mailto:taoran.lu@dolby.com), [F. Pu](mailto:fangjun.pu@dolby.com), [W. Husak (Dolby)](mailto:wjh@dolby.com) |
| [JVET-AA0145](file:////Users/asegall/Downloads/current_document.php%3fid=11821) | AHG9: On decoupling neural-network post-filter activation SEI message | [H.-B. Teo](mailto:hanboon.teo@sg.panasonic.com), J. Gao, C.-S. Lim, [K. Abe](mailto:abe.kiyo@jp.panasonic.com), [V. Drugeon (Panasonic)](mailto:virginie.drugeon@eu.panasonic.com) |
| **Super Resolution** | | |
| [JVET-AA0065](file:////Users/asegall/Downloads/current_document.php%3fid=11741) | AHG11: CNN Filter for Super-Resolution with RPR functionality in VVC | [S. Huang](mailto:shimin_huang2022@163.com), [C. Jung (Xidian Univ.)](mailto:zhengzk@xidian.edu.cn), [Y. Liu](mailto:serena@oppo.com), [M. Li (Oppo)](mailto:myron.li@oppo.com) |
| [JVET-AA0076](file:////Users/asegall/Downloads/current_document.php%3fid=11752) | AHG11: RPR-Based Super-Resolution Guided by Partition Information | [Q. Han](mailto:hanqihui2013@163.com), [C. Jung (Xidian Univ.)](mailto:zhengzk@xidian.edu.cn), [Y. Liu](mailto:serena@oppo.com), [M. Li (Oppo)](mailto:myron.li@oppo.com) |
| [JVET-AA0084](file:////Users/asegall/Downloads/current_document.php%3fid=11760) | AHG11: Neural Network based Super Resolution for Video Coding Using Multiple Side Information | [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) |
| **Inter Prediction** | | |
| [JVET-AA0082](file:////Users/asegall/Downloads/current_document.php%3fid=11758) | AHG11: Deep Reference Frame Generation for Inter Prediction Enhancement | [Z. Liu](mailto:zizhengliu@tencent.com), [X. Xu](mailto:xiaozhongxu@tencent.com), [S. Liu (Tencent)](mailto:shanl@tencent.com), [J. Jia](mailto:jiajh2021@whu.edu.cn), [Z. Chen (Wuhan Univ.)](mailto:zzchen@whu.edu.cn) |
| **End-to-End** | | |
| [JVET-AA0063](file:////Users/asegall/Downloads/current_document.php%3fid=11739) | AHG11: A hybrid codec using E2E image coding combined with VVC video coding | [Y. He](mailto:yu.he1@huawei.com), [B. Wang](mailto:biao.wang@huawei.com), [E. Alshina](mailto:elena.alshina@huawei.com), [J. Sauer](mailto:johannes.sauer@huawei.com) |
| **Software** | | |
| [JVET-AA0086](file:////Users/asegall/Downloads/current_document.php%3fid=11762) | AHG11: Small Ad-hoc Deep-Learning Library (SADL) update | [F. Galpin](mailto:franck.galpin@interdigital.com), [T. Dumas](mailto:thierry.dumas@interdigital.com), [P. Bordes](mailto:philippe.bordes@interdigital.com), [E. François (InterDigital)](mailto:edouard.francois@interdigital.com) |
| **Cross Checks** | | |
|  |  |  |

1. **Recommendations**

The AHG recommends:

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

[JVET-AA0012](https://jvet-experts.org/doc_end_user/current_document.php?id=11837) JVET AHG report: Enhanced compression beyond VVC capability (AHG12) [M. Karczewicz, Y. Ye, L. Zhang, B. Bross, X. Li, K. Naser, H. Yang (AHG chairs)]

1. **Activities**

The primary activity of the AHG was the “Exploration experiment on enhanced compression beyond VVC capability” (JVET-Z2024). The combined improvements of the ECM-5.0 over VTM-11.0ecm anchorfor AI, RA and LB configurations are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | All Intra Main10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -6.84% | -14.30% | -19.53% | 390% | 266% |
| Class A2 | -6.42% | -15.04% | -16.00% | 380% | 259% |
| Class B | -5.92% | -16.12% | -15.78% | 433% | 276% |
| Class C | -7.11% | -10.83% | -11.38% | 424% | 254% |
| Class E | -7.92% | -13.65% | -14.49% | 401% | 277% |
| Overall | -6.75% | -14.05% | -15.25% | 409% | 267% |
| Class D | -5.92% | -9.33% | -8.84% | 421% | 268% |
| Class F | -15.79% | -20.74% | -20.14% | 355% | 270% |
| Class TGM | -26.19% | -28.03% | -27.29% | 317% | 275% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Random Access Main 10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -16.80% | -18.70% | -24.21% | 436% | 536% |
| Class A2 | -17.87% | -24.31% | -26.27% | 415% | 626% |
| Class B | -15.58% | -23.66% | -22.40% | 432% | 550% |
| Class C | -17.51% | -21.09% | -20.65% | 454% | 555% |
| Class E |  |  |  |  |  |
| Overall | -16.80% | -22.11% | -23.07% | 435% | 563% |
| Class D | -18.49% | -22.02% | -22.71% | 444% | 579% |
| Class F | -18.31% | -23.02% | -23.03% | 403% | 371% |
| Class TGM | -24.41% | -28.17% | -28.50% | 425% | 299% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Low Delay B Main 10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -14.25% | -27.64% | -25.57% | 396% | 419% |
| Class C | -15.40% | -24.03% | -23.37% | 449% | 434% |
| Class E | -12.99% | -18.62% | -20.64% | 400% | 317% |
| Overall | -14.32% | -24.18% | -23.61% | 414% | 396% |
| Class D | -17.44% | -25.35% | -24.26% | 416% | 489% |
| Class F | -16.96% | -24.65% | -24.21% | 399% | 328% |
| Class TGM | -23.33% | -30.18% | -29.91% | 415% | 294% |

The rate reduction for natural sequences over VTM for RA configuration for {Y, U, V} increased from {-15.83%, -19.85%, -20.78%} to {-16.80%, -22.11%, -23.07%}. For SCC sequences (class TGM) the rate reduction for RA configuration increased from {-16.11%, -19.04%, -19.39%} to {-24.41%, -28.17%, -28.50%}.

1. **Contributions**

In addition to 21 EE2 contributions, 38 (same as last meeting) related contributions were received which can be subdivided as follows:

***In Loop Filters (2)***

JVET-AA0139, "Non-EE2: Longer deblocking filter for luma", K. Andersson J. Enhorn (Ericsson)

JVET-AA0147, "EE2-Related: Extended Offline-Filtering Taps for ALF", W. Yin, K. Zhang, Z. Deng, L. Zhang (Bytedance)

***Intra (13)***

JVET-AA0043, "IntraTMP Adaptation for Camera Captured Contents", K. Naser, T. Poirier, F. Galpin, A. Robert (InterDigital)

JVET-AA0044, "IntraTMP for chroma Components", K. Naser, T. Dumas, T. Poirier, F. Galpin (InterDigital)

JVET-AA0045, "[EE2-1.4 related] Reduced Complexity Spatial GPM", K. Naser, Y. Chen, A. Robert, K. Reuze (InterDigital)

JVET-AA0103, "EE2-related: CCLM with non-linear term", X. Li, Y. Ye, R.-L. Liao, J. Chen (Alibaba)

JVET-AA0104, "Non-EE2: On planar horizontal mode and planar vertical mode", X. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba)

JVET-AA0114, "[EE2-related] Division-free operation and mean-compensation for convolutional cross-component model (CCCM)", A. Aminlou, J. Lainema, P. Astola (Nokia)

JVET-AA0120, "Non-EE2: Template-based multiple reference line intra prediction", L. Xu, Y. Yu, H. Yu, D. Wang (OPPO)

JVET-AA0121, "Non-EE2: Template-based MIP", Z. Xie, Y. Yu, H. Yu, D. Wang, Y. Liu, M. Li (OPPO), J. Huo, W. Qiao, X. Hao, Y. Ma, F. Yang (Xidian University),

JVET-AA0127, "EE2-1 related: Encoder optimization of EE2-1.2 and 1.3", C.-W. Kuo, H.-J. Jhu, X. Xiu, N. Yan, W. Chen, X. Wang (Kwai)

JVET-AA0136, "Non-EE2: On CCCM improvement", Y.-J. Chang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AA0137, "Non-EE2: Intra Prediction Fusion", K. Cao, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AA0138, "EE2-related: On Gradient Linear Model (GLM)", X. Li, Y. Ye, R.-L. Liao, J. Chen (Alibaba)

JVET-AA0140, "EE2-related: Self-Aware Filter Estimation for CCLM", K. Zhang, Z. Deng, L. Zhang (Bytedance)

***Inter (12)***

JVET-AA0069, "Non-EE2: AmvpMerge for low delay", H.Jang, J.Nam, N.Park, J.Lim, S.Kim (LGE)

JVET-AA0075, "Non-EE2: Template matching based BCW index derivation for merge mode", R.-L. Liao, J. Chen, Y. Ye, X. Li (Alibaba)

JVET-AA0119, "EE2-1.4a-related: Modifications of Spatial GPM", F. Wang, Y. Yu, H. Yu, D. Wang (OPPO)

JVET-AA0124, "Non-EE2: Enable amvpMerge mode on scaled reference pictures when DMVD is disabled", Z. Zhang, H. Huang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AA0128, "EE2-related: On regression based affine candidate derivation", W. Chen, X. Xiu, C.-W. Kuo, H.-J. Jhu, N. Yan, X. Wang (Kwai)

JVET-AA0134, "Non-EE2: POC based BCW weights derivation", Z. Zhang, H. Huang, C.-C. Chen, V. Seregin, M. Karczewicz (Qualcomm)

JVET-AA0141, "Non-EE2: Enhanced temporal motion information derivation", L. Zhao, K. Zhang, L. Zhang (Bytedance)

JVET-AA0142, "AHG12/Non-EE2: Picture-Level Geometry Transform", W. Jia, K. Zhang, Y. Wang, T. Fu, Y. Li, L. Zhang (Bytedance)

JVET-AA0143, "Non-EE2: Simplification methods for OBMC", K. Kim, D. Kim, J.-H. Son, J.-S. Kwak (WILUS)

JVET-AA0144, "Non-EE2: DMVR for affine merge coded blocks", J. Chen, R.-L. Liao, X. Li, Y. Ye (Alibaba)

JVET-AA0148, "Non-EE2: On MHP (Multi-Hypothesis Prediction)", K. Sato, Y. Yu, H. Yu, D. Wang (OPPO)

JVET-AA0149, "EE2-1.4 related: Improvements on Spatial GPM", A. Natesan, J. N. Shingala, J. R. Arumugam, V. Valvaiker (Ittiam), T. Lu, P. Yin, F. Pu, T. Shao, A. Arora, S. McCarthy (Dolby)

***Transform Coding (2)***

JVET-AA0064, "AHG12: A study on non-separable primary transform", J. Choi, M. Koo, J. Lim, J. Zhao, S. Kim (LGE)

JVET-AA0073, "Non-EE2: Modification of LFNST for MIP coded block", J.-Y. Huo, W.-H. Qiao, X. Hao, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.), J. Ren, M. Li (OPPO)

***Entropy Coding (1)***

JVET-AA0150, "AHG12: On CIPF (CABAC Initialization from the Previous Frame)", K. Sato, Y. Yu, H. Yu, D. Wang (OPPO)

***Screen Content Coding (3)***

JVET-AA0053, "AHG12: Using block vector derived from IntraTMP for IBC", W. Lim, D. Kim, S.-C. Lim, J. S. Choi (ETRI)

JVET-AA0077, "AHG12: On BVD coding for IBC", A. Filippov, V. Rufitskiy (Ofinno)

JVET-AA0108, "AHG12: IBC AMVP candidates clustering", D. Ruiz Coll, V. Warudkar (Ofinno)

***360 Video (1)***

JVET-AA0146, "AHG12/Non-EE2: Fixes on ECM for 360-degree video coding", Y. Wang, K. Zhang, Z. Deng, L. Zhang (Bytedance)

***ECM Encoder and Test Conditions (4)***

JVET-AA0097, "ECM fix for block-level out-of-bound checking", F. Le Léannec, P. Andrivon, M. Radosavljevć (Xiaomi)

JVET-AA0098, "AHG12: Encoder configuration proposal to reduce worst case encoding time", F. Le Léannec, P. Andrivon, M. Radosavljevć (Xiaomi)

JVET-AA0129, "Non-EE2/AHG10: Improved inter hash RDO considering OBMC off in ECM5.0", X. Xiu, C.-W. Kuo, H.-J. Jhu, W. Chen, N. Yan, X. Wang (Kwai)

JVET-AA0132, "AHG6: ECM software configuration parameters for template matching tools", C.-C. Chen, H. Huang, V. Seregin, M. Karczewicz (Qualcomm)

1. Recommendations

The AHG recommends to:

* To review all the related contributions.

It is pointed out that a late contribution (JVET-A0191) is relating to modified test conditions.

[JVET-AA0013](https://jvet-experts.org/doc_end_user/current_document.php?id=11838) JVET AHG report: Film grain technologies (AHG13) [W. Husak, M. Radosavljević, W. Wan, D. Grois, Y. He, P. de Lagrange, A. Segall, A. Tourapis (AHG chairs)]

1. **Discussion**

***Conformance***

The topic of conformance for film grain SEI implementations was explored. A key aspect of the discussion was the history of film grain implementations in other groups. Another video codec group explored the concept of making conformance mandatory but realized that any implementation would be dependent on the post-decode video processing platform. Codec conformance generally ends at the decode buffer while further video processing may occur downstream. This is the same issue ISO/IEC/ITU faces where a potential reference model would also need to specify downstream video pipeline details – something that is typically out of scope of JVET.

Film grain that is modeled by the encoder is targeted to a nominal resolution that is the anticipated decoded video output. The display device will then adapt the decoded output to the display using brand-specific and/or display technology-specific video pipelines. Some example video processing steps that may impact conformance are internal working color spaces, video pixel representations, chroma sub-sampling, and image resizing.

Another relevant point is VVC v1 has been commercially available but without any specified film grain conformance and therefore a class of decoders will exist that cannot inherently meet any post facto conformance specification. These reasons result in an acknowledgement that hard conformance will not be practical for film grain. Soft conformance is currently undefined but more attractive for discussion within the context of SEI post processing. Application/implementation level specifications might be better places to define or specify conformance due to the above constraints.

The following conformance definitions were discussed:

**Hard conformance** – an implementer must follow a specific set of guidelines for an expected output.

**Soft conformance** – an implementer should follow specific guidelines but may have other options for implementation. The frame packing SEI is a good example of an application that could have soft conformance. HDR SEIs can also be a model due to display adaptation.

**Not required** – implementations are completely optional.

The artifact masking use case is a context where strict conformance could be difficult to define and may be unnecessary since the desired outcome is subjective image improvement.

***Implementations (Identification)***

The definition differences between a reference model and technology example were also discussed. The reference model is traditionally a working example used throughout the development of the standard and/or report. In the film grain case, there are equations that could benefit from a software model. It is believed that the “reference model/implementation” and “example” are interchangeable for this application. The group recommends there is language highlighting that implementers can use the reference implementation, but they may choose to use alternative implementations.

***Classification/Categorization***

Classification was discussed and the result was any technology could be categorized and/or classified by use case. Three use cases identified are:

* Film grain insertion (creative intent)
* Film grain removal/replacement
* Artifact masking.

More use cases could be identified at a later date. Also explored was transcoding between autoregressive and frequency-based models although that could be a different aspect and requires more research. Other differences in models, such as implementation complexity, may also be important.

An important question is whether there are technology examples that favor one or more of the use cases? The main differences are in the implementations themselves – different use cases may need more detailed parameter sets. An example is bitrate and application can create a dependency for at least artifact masking (less bitrate may need more film grain energy). Another example is mobile devices could require less grain due screen size or even implementation complexity. There could be other parameters that are unique to the content (motion etc.) or the display/environment.

Different implementation examples may be chosen based on the target application in order to favor specific parameters. Automatic parameter estimation or manual parameter settings may be better done using one implementation example over another. The encoder and decoder should be a matched set (or closely related). There may need to be a translation step required in this type of scenario. This may not be a big issue but should be further explored.

Additional thoughts explored are:

* What happens with display devices compositing local text or graphics?
* HDR video pipeline design could present a problem for film grain insertion.
* Arbitrary image scaling for final display could change the look of the grain.

This could lead to a note in the technical report.

1. **Related contributions**

Three contributions related to AHG13 were identified as of 07/12/2022:

* One was the AHG report:
  + JVET-AA0013 JVET AHG report: Film grain technologies (AHG13)
* Two were related to the technical report and explore profiling:
  + JVET-AA0051 Film grain synthesis technology for video applications (Draft 2)
  + JVET-AA0052 AHG13: On VSEI film grain profiles

***Contributions***

**JVET-AA0051 Film grain synthesis technology for video applications (Draft 2)**

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 end-to-end processing steps for film grain synthesis, that 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-AA0052 AHG13: On VSEI film grain profiles**

This contribution proposes two VSEI profiles, frequency filtering film grain synthesis profile and auto-regressive film grain synthesis profile, for the film grain characteristics (FGC) SEI message. The profiles specify constraints on FGC SEI and the requirements for the capabilities needed to enable the use of FGC SEI.

1. **Recommendation**

The AHG recommends:

* the related input contributions are reviewed;
* the draft TR text be edited during the JVET meeting period;
* the draft TR includes notes list above; and
* continue the study of film grain technologies in JVET.

It is pointed out that conformance should not refer to an implementation itself, but rather to the output that a conformant implementation reproduces. Conformance should somehow be quantified, which should be

JVET-AA0051 should not be considered to be draft 2, but is rather a proposed text. It is suggested to change the title to “Proposed draft text …”. It should also add changemarks relative to draft 1.

# Project development (19)

## Deployment and advertisement of standards (2)

Contributions in this area were discussed in session 15 at 0700–0710 UTC on Tuesday 19 July 2022 (chaired by JRO).

[JVET-AA0020](https://jvet-experts.org/doc_end_user/current_document.php?id=11726) Deployment status of the HEVC standard [G. J. Sullivan (Microsoft)]

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-Z0020-v1 of April 2022.

New announcements:

**Mac Studio** includes hardware-accelerated HEVC support

**Chromium** browser on **Chromebooks using the Google Chrome OS** with Intel Gen-11 processors supports HEVC for protected content playback

**GoPro Max** 360° camera (introduced in 2019 or 2020 and waterproof to 5 m) uses HEVC for all modes, using 5.6K capture (ordinarily with 1080p export)

The **Canon R5 C** (January 2022) is a mirrorless cinema and photo hybrid camera with HEVC support for up to 8K 10 bit UHD at 30 fps and 4K UHD at 60 fps with support for very high bit rates and HDR

**Panasonic Lumix GH5 Mark II micro four-thirds prosumer mirrorless camera** (firmware 2.0 or higher) with 4K 10 bit and HLG HDR support up to 60 fps, May 2021, **GH6**, with 4:2:2 as well, February 2022

The **JVC KY-PZ510** series (announced April 2022) of pan-tilt-zoom cameras support HEVC up to 4K at 60 fps

**Colorfront** offers a **Colorfront Streaming Server** with remote color grading capability using HEVC, as of April 2022

**IETF RFC 7798** (March 2016) is a standards track RFC that specifies a Real-time Transport Protocol (RTP) payload format for HEVC. The RFC includes a media type registration (a.k.a. MIME type registration) and parameters for use with the Session Description Protocol (SDP)

[JVET-AA0021](https://jvet-experts.org/doc_end_user/current_document.php?id=11937) Deployment status of the VVC standard [[G. J. Sullivan (Microsoft)](mailto:garysull@microsoft.com)]

(abstract)

Reference to JVET-AA0154 only.

## Text development and errata reporting (2)

Contributions in this area were discussed in session 15 at 0710–0625 UTC on Tuesday 19 July 2022 (chaired by JRO).

[JVET-AA0048](https://jvet-experts.org/doc_end_user/current_document.php?id=11723) Some VVC text changes [Y.-K. Wang (Bytedance), G. J. Sullivan (Microsoft)]

This contribution suggests to include the following VVC text changes into JVET-AA2005 (New level and systems-related supplemental enhancement information for VVC (Draft 3), and the corresponding WG 5 output document, if any:

1. Not to specify the constraints on the values of BinCountsInPicNalUnits and BinCountsInSubpicNalUnits for the unlimited level (i.e., level 15.5).
2. A fix for [bug ticket #1556](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1556) for an asserted typo in clause 7.4.2.4.5 (Order of VCL NAL units and their association to coded pictures), wherein the phrase "subpicture level index" is suggested to be replaced with "subpicture index".

It is agreed that those changes are reasonable. To be resolved in the context of the DoC on CDAM1 (ballot comment in m59992 also refers to item 1).

The ballot comments were reviewed in session 15.

There is a ballot comment on converting to a new edition. It is agreed that this shall be done at final stage (January 2023), as the final text of the previous edition of ISO/IEC is not yet available.

There is a ballot comment on including code points for new SEI messages in VSEI. This can be done on the basis of the VSEI CDAM.

There is a ballot comment on payload type 205 for scalability dimension SEI (from version 2) which is in conflict with the number previously used for shutter interval in other standards. It was agreed to change the payload type of scalability dimension SEI to some other available value, to keep payload type numbers consistent across different standards.

Regarding progression in ITU, the next version of VVC could be finalized in October. Depending on the maturity of the VSEI extensions, the new payload types should appear in the next H.266 edition, but the 205 value should already be changed in any case.

The other comments are straightforward.

G. J. Sullivan and Y.-K. Wang will prepare the DoC.

[JVET-AA0222](https://jvet-experts.org/doc_end_user/current_document.php?id=11912) Motion Vector limits for AVC/H.264 levels 6.x [A. Tourapis (Apple)] [late]

The 24th version of the ITU-T Rec. H.264 | ISO/IEC 14496-10 document (11th edition of the H.264 specification) introduced, among other things, several additional levels for decoder capability, i.e., levels 6, 6.1, and 6.2, mostly to support larger picture sizes and higher frame rates. One of the introduced changes to support such functionality included the increase of the motion vector range limits that can be supported compared to lower levels. In particular, the vertical motion vector component limit MaxVmvR was increased from 512 to 8192, while the horizontal motion vector range for luma motion vectors was changed from [−2048, 2047.75] to [−8192, 8191.75]. This change was seen as a potentially desired change so as to better handle larger resolution images that may be characterized by higher range motion and to better align with more modern coding specifications such as HEVC and VVC. However, during the introduction of those levels, no concerns were raised about existing HW implementations.

It was recently drawn to our attention that there exist several HW AVC decoder implementations that although can support the resolutions that correspond to the 6.x levels, cannot unfortunately claim complete conformance to these levels because of this difference in motion vector range. This is because such implementations are mostly based on earlier AVC decoder designs that were limited to 5.x levels. Ideally, for properly supporting these new levels a low level change in the design of these decoders is necessary given the higher bitdepth precision needed for the new motion vector range. Such change is, unfortunately, seen as costly and undesirable. suporting. In some cases this would mean a complete replacement of existing HW devices that may have such capabilities.

This document requests an “amendment” to the capabilities of these levels so as to properly enable such decoders. A few possible solutions are suggested.

The following solutions to this issue were identified.

1. Update the motion vector range of the existing levels 6.x to reflect the desired change regardless of profile.   
   We understand that this solution may not be desirable especially if there are bitstreams already encoded using these levels.
2. Update the motion vector range of the existing levels 6.x of certain profiles to reflect the desired change. This would impact the following profiles:
   1. Constrained Baseline Profile
   2. Baseline Profile
   3. Extended Profile
   4. Main Profile
   5. High Profile
   6. Progressive High Profile
   7. Constrained High Profile
   8. Multiview High Profile

Given that this is limited around certain profiles, this likely limits the impact of invalidating some existing bitstreams, but does not eliminate this probability.

1. Specify that the profiles in (b) need to also indicate explicitly the log2\_max\_mv\_length\_horizontal and log2\_max\_mv\_length\_vertical values and be constrained with values 11 and 9, respectively. This does not necessarily invalidate any previously generated bitstreams although these are currently not exactly conforming to any particular level. However, one could argue that these bitstreams could be acknowledged by such devices and could be decoded in a best effort decoding mode.
2. Define new levels that would have the same capabilities as levels 6.x other than the motion vector range. This could be done by either introducing new level numbers or augmenting the existing level numbers, i.e., creating new levels 6b, 6.1b, and 6.2b, with a constraint set flag or by imposing the presence of the mv limits in VUI.

The proponents’ preference would be to consider option d, preferably with a constraint flag, and if that is not possible option c.

The problem comes due to the larger bit depth (16 bit) required for MV handling only in levels 6.x, and some implementations not using such bit depth for lower levels, and no redesign was made for 6.x.

Two other hardware manufacturers expressed that they would not see a problem. If they would implement those levels, they would implement it as specified, as some redesign would be necessary anyway.

The levels were originally designed in 2015. There was some correction of MV deltas in 2017, but currently there is no problem with the spec itself.

It is commented that option d would break the hierarchy of levels.

It was commented that defining new levels for the benefit of a wrong hardware implementation would be inappropriate.

The suggested solutions were asserted to be inappropriate.

It was suggested as a possible alternative to add a sentence like "It is recommended to not use motion vectors bigger than X and to indicate this with log2\_max … equal to Y, as there have been reports of problems in some decoder implementations …"

Regardless of that, it is note that existing decoders which do not support the full MV ranges would not be conformant.

The following text to be included into JVET-AA1004 was reviewed and agreed in session 23. It is understood that by its wording does not introduce a precedence case that JVET would take action in supporting existing non-conformant decoder.

“For bitstreams of levels 6, 6.1, and 6.2 (see Annex A), it has been reported that some decoder implementations do not properly decode bitstreams that have horizontal motion vector component values that exceed the range from −2048 to 2047, inclusive, or that have vertical motion vector component values that exceed the range from −512 to 511, inclusive, in units of ¼ luma sample displacement. It is therefore suggested that bitstreams should not contain motion vector component values that exceed these ranges. This constraint can be indicated by using values of log2\_max\_mv\_length\_horizontal less than or equal to 11 and values of log2\_max\_mv\_length\_vertical less than or equal to 9.”

Another necessary change of replacing “absolute value” by “range” (2x) in clause E.2.1 was detected in that context and shall also be included in JVET-AA1004 (for some later edition or amendment of AVC).

## Test conditions (6)

Contributions in this area were discussed in session 15 at 0625–0705 UTC on Tuesday 19 July 2022 (chaired by JRO) unless noted otherwise.

[JVET-AA0046](https://jvet-experts.org/doc_end_user/current_document.php?id=11721) [AhG4] Report on AhG4 meeting on development of a gaming-type CTC class [M. Wien (AHG chair)]

Was presented and discussed in joint meeting with AG 5 and VCEG at 1400-1440 UTC on Monday 18 July 2022 (chaired by JRO, presented by MW)

The teleconference was organized by AhG4 and conducted on 2022-07-05 15:00h UTC. Interested parties were invited to prepare input for discussion at the meeting. No input documents to the AhG meeting registered. The group agreed to discuss criteria for the setup of a proposed new CTC class which could be named “Class G”. In terms of resolution, it was suggested that at least 3840×2160 and 1920×1080 should be included, subdividing the class into G1 and G2. A table with favorable contents features was created which the proponents of new test sequences are recommended to use in their input contributions proposing sequences for the new class.

Consideration of:

* Set of sequences under consideration
* Number of sequences, number of resolutions (could become sub-classes)
* Composition of sequence features and size of test set
* Consideration of additional aspects.

No input documents to the AhG meeting registered. The group agreed to discuss criteria for the setup of a proposed new CTC class.

Required steps for the implementation of the new class:

* CTC document
  + Definition of set of sequences, incl. md5sums (plus license on the ftp site)
  + Definition of coding conditions
  + Extension of the Excel template for CTC reporting
* FTP site
  + Upload of sequences to CTC directory structure
    - Naming convention:  
      <NameInCamelCase
* Initial reporting on rate-distortion results, including encoder / decoder run times

Composition of the new class:

* Included resolutions
  + 3840×2160 50/60fps 10bit, BT.709
    - Would it be good or better to have BT.2020?
  + 1920×1080 50/60fps 10bit, BT.709
  + Inclusion of upright orientation next to the conventional landscape?
    - Represents a common use case on mobile devices
    - Question: is this distinction needed or do the landscape sequence sufficiently cover the technical challenges?
  + The question was raised if other resolutions should be included as well. No specific comments were noted in response to this question. (TBD)
  + The sequences should be 10s of length
* Composition of scene
  + 1st-person perspective
  + 3rd-person perspective
  + Other perspectives might be of interest (e.g. like in the AOV sequence).
  + Inclusion of text and overlays should be included to some extent
    - Scores
    - Chat
    - Graphical elements like e.g. maps
  + The set of sequences should represent different rendering characteristics
* Content type
  + Occurrence of persons / characters
  + Occurrence of vehicles
  + Naturalistic rendering as well as intentionally synthetic sequences
  + …

It was requested to provide input on the raised questions at the upcoming meeting.

Indication of a possible structure of the new class “G”

* G1: 6 UHD, LDB, LDP.   
  RA/AI could be non-mandatory
* G2: 6 HD, LDB, LDP.   
  RA/AI could be non-mandatory
* G3: ??

It was noted that the low-delay configuration would be the most relevant use case. It was noted that some sequences might have significant encoder run-times.

Matrix of intended content features:

|  |  |  |
| --- | --- | --- |
| Feature | 3840×2160 | 1920×1080 |
| 1st-person perspective | Fontainebleau\_Cinematic\_3840x2160\_60p\_420\_10\_bits | Darktree\_1920x1080\_60p\_420\_10\_bits  Fontainebleau\_Cinematic\_1920x1080\_60p\_420\_10\_bits  Fontainebleau\_FPV\_1920x1080\_60p\_420\_10\_bits |
| 3rd-person perspective | 3D\_game\_kit\_3840x2160\_60p\_420\_10bits  Racing \_3840x2160\_60p\_420\_10bits | 3D\_game\_kit\_Level-1\_1920x1080\_60p\_420\_10bits  3D\_game\_kit\_Level-2\_1920x1080\_60p\_420\_10\_bits |
| Other perspectives (name them) | -/- | -/- |
| Scores | Racing \_3840x2160\_60p\_420\_10bits (Speed) | -/- |
| Chat | -/- | -/- |
| Graphical elements | -/- | -/- |
| Persons / characters | 3D\_game\_kit\_3840x2160\_60p\_420\_10bits | 3D\_game\_kit\_Level-1\_1920x1080\_60p\_420\_10bits  3D\_game\_kit\_Level-2\_1920x1080\_60p\_420\_10\_bits  Darktree\_1920x1080\_60p\_420\_10\_bits |
| Vehicles | Racing \_3840x2160\_60p\_420\_10bits | Darktree\_1920x1080\_60p\_420\_10\_bits |
| Naturalistic rendering | Fontainebleau\_Cinematic\_3840x2160\_60p\_420\_10\_bits | Darktree\_1920x1080\_60p\_420\_10\_bits  Fontainebleau\_Cinematic\_1920x1080\_60p\_420\_10\_bits  Fontainebleau\_FPV\_1920x1080\_60p\_420\_10\_bits |
| Synthetic rendering | 3D\_game\_kit\_3840x2160\_60p\_420\_10bits  Racing \_3840x2160\_60p\_420\_10bits | 3D\_game\_kit\_Level-1\_1920x1080\_60p\_420\_10bits  3D\_game\_kit\_Level-2\_1920x1080\_60p\_420\_10\_bits |

Request to proponents of test sequences:

* Provide input contribution with proposed sequences, characterized by filling the above matrix
* Provide rate-distortion results
* Provide raw sequences and (ideally) example bitstreams

A proposal for the suggested class is intended to be composed as an activity at the meeting based on the input received.

It was commented that some sequences have high motion which is typical for the application, but not comfortable for subjective evaluation.

Would testing 4K on low delay be practical? Is it expected to happen in gaming video applications? It can happen in some applications such as airplay, but currently is not mainstream yet.

It was pointed out that JVET-AA0241 could provide additional criteria to classify sequences.

It was agreed to define a new optional class “G” for LB/LD/AI, but keep it simple in the beginning: Small number of sequences (4? HD) which represent the variety of content characteristics.

It was asked if requirements for cloud gaming are clear, in particular regarding the end-to-end delay? One expert reports that 1 frame delay would be OK (see JVET-Y0043).

[JVET-AA0098](https://jvet-experts.org/doc_end_user/current_document.php?id=11774) AHG12: Encoder configuration proposal to reduce worst case encoding time [F. Le Léannec, P. Andrivon, M. Radosavljević (Xiaomi)]

This contribution proposes to modify EE2 coding test conditions, to significantly reduce the worst-case encoding runtime for RA, with negligible impact on overall BD-rate results.

The CTC proposal consists in modifying the encoder configuration for class A sequences, QP 22, which is the worst case. The maximum MTT hierarchy depth is lowered to 2 from the 4th temporal layer (test 1) or from the 3rd temporal layer (test 2).

Test 1 reduces the worst-case encoding time by 7 days of overall computation (sequences DaylightRoad2 and ParkRunning3). Overall impact on BD-rate is +0.01% loss.

Test 2 reduces the worst-case encoding time by about 11 days of overall computation (sequences DaylightRoad2 and ParkRunning3). Overall impact on BD-rate is +0.02% loss.

Encoding time is decreased to 99% by test 1, 97% by test 2. The measurement in “days reduction” relates to specific sequences.

The change is only done for smallest QP value. As this appears inconsistent, it was suggested to report results when the change is made over all QP values. It was also asked how the change would affect results in other classes (though the problem is primarily for classes A).

Was reviewed in session 24 after results were available.

Loss with only modifying QP22is 0.03% in both A classes.

If used for all QPs in the A classes, the loss goes up to 0.13% (average over all classes, loss is 0.28%/0.38% in class A1/A2).

If used for all classes, the loss goes up to 0.52%.

Opinion was expressed by several experts that the reduction of worst-case encoding time is desirable

Decision(CTC): Adopt JVET-A0098 (original proposal, only use for QP22 and only in classes A)

It was suggested to apply the change also to the VTM anchors (not VTM CTC).

[JVET-AA0170](https://jvet-experts.org/doc_end_user/current_document.php?id=11858) Crosscheck of JVET-AA0098 (AHG 12: encoder configuration proposal to reduce worst case encoding time) [W. Lim, S.-C Lim (ETRI)] [late]

[JVET-AA0130](https://jvet-experts.org/doc_end_user/current_document.php?id=11806) AHG8: Draft VTM and HM common test conditions for high bit depth and high bit rate video coding [A. Browne, T. Ikai, D. Rusanovskyy, X. Xiu, Y. Yu (AHG chairs)]

This document proposes common test conditions (CTC) and software reference configurations to be used in the context of high bit depth and high bit rate video coding experiments after the 27nd JVET meeting. These common test conditions are recommended for use in technical contributions to the 28nd and following JVET meetings, as applicable.

Configurations mostly carried over from the previous CE conducted for the v2 developments. GOP16 was used there both for VTM and HM in the low QP range. It was suggested to investigate which gain might be possible if GOP32 would be used for VTM, would be desirable using GOP32 over all QPs.

It was agreed to convert this contribution into an output document JVET-AA2018

[JVET-AA0191](https://jvet-experts.org/doc_end_user/current_document.php?id=11880) [AHG12] On CTC for Low Complexity ECM [K. Naser, A. Robert, T. Poirier, P Le Guyadec, S. Puri, F. Galpin (InterDigital)] [late]

Discussed in session 19 0650-0715.

This contribution proposes an additional CTC for low complexity ECM software under CTC by removing tools that inherently cause delays and complexity in the processing pipeline of the decoder and encoder, namely template tools. It is argued that such disadvantages may jeopardize their implementation in practice.

It is proposed that all the tools that use template or require decoder analysis are disabled. Simulations are performed on the latest ECM-5.0 test model software under the CTC, and results are reported in the table below:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Test | Cfg | Y | U | V | EncT | DecT |
| Proposed vs ECM-5.0 | AI  RA  LDB | 1.74%  5.44%  5.28% | 1.90%  5.40%  5.19% | 1.96%  5.62%  5.48% | 62%  70%  61% | 78%  82%  66% |

It is proposed consider the new CTC for low complexity the ECM, while keeping the original ECM CTC. This enables better study of potential coding gain beyond VVC.

It was commented that at the current stage of exploration the highest priority is getting more compression while staying practical in terms of computing time.

It was also commented that in the VVC development extensive tool-on/off tests had been performed, which however would be a high additional workload for AHG12.

It was commented that by defining two different CTCs, this might end up in a situation that proponents would need to test their newly proposed ideas in two environments, and further these two environments might deviate more and more from each other. In the presentation of the proposal, an example was given where another tool for a certain purpose would be used in the “light” version. This would contradict the traditional way of “onion-style design”, where a light version is a subset of an extended version.

“Study the performance and complexity tradeoff of video coding tools” is already under the mandates of AHG12, and contributions like this are very welcome for getting more information about the complexity/gain tradeoff of individual tools or groups of tools (e.g., those related to TM).

[JVET-AA0194](https://jvet-experts.org/doc_end_user/current_document.php?id=11883) On VTM results for HDR content [A. Browne, S. Keating, K. Sharman (Sony)] [late]

This contribution reports results for VTM 17.0 vs VTM 16.0 for both the HDR CTC (JVET-Z2011) and the high bit depth CTC (JVET-U2018). It is reported that these results show differences between BD-rate changes for the PSNR and wPSNR metrics. Further, it is reported that the results show that there are also different results in terms of these metrics on a per sequence basis, with some sequences performing much better than others. It is suggested that if HDR is considered an important future use case of VVC that it should be strongly encouraged that any proposal is backed with results for both HDR and SDR CTCs.

After software bug fix (MR 2265) the initially reported problem was resolved. Gains are now more consistent for the different metrics.

It was suggested that regularly the effect of VTM changes on HDR content should be investigated. This did not happen after the HDR AHG was closed. Should be added as an action item to AHG3 (add co-chair).

[JVET-AA0241](https://jvet-experts.org/doc_end_user/current_document.php?id=11931) JVET CTC Content Characterization [P. Singh, A. Tourapis, Y. Zheng, A. Nalci (Apple)] [late]

Was presented in joint meeting with AG 5 and VCEG at 1450-1615 UTC on Monday 18 July 2022 (chaired by JRO and MW)

This document presents a preliminary analysis of some of the characteristics of the content currently specified in the Joint Video Experts Team (JVET) Common Test Conditions (CTC). This analysis is primarily based on the use of the Video Complexity Analyzer (VCA) tool. Certain recommendations on the complexity of the CTC content are also provided.

VCA determines energy function (DCT domain), and based on that, spatial complexity (SC) and temporal complexity (TC) as criteria. All based on luma.

Alternatively, SI/TI as specified in P.910. Main advantage of SC/TC is that they can be computed faster than SI/TI.

27 JVET CTC sequences are analyzed. It was found that most sequences are in the domain of relatively low SC/TC domains, but no sequence has both.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sequence** | **Average SC** | **Average TC** | **(Max SC, TC)** | **(SC, Max TC)** | **(Min SC, TC)** | **(SC, Min TC)** |
| Tango2 | 8.99 | 8.84 | (9, 8.66) | (9, 9.85) | (8, 8.61) | (9, 8.15) |
| FoodMarket4 | 9.49 | 10.68 | (14, 16.33) | (11, 19.72) | (3, 4.39) | (3, 3.21) |
| Campfire | 11.02 | 14.97 | (13, 25.35) | (12, 25.69) | (9, 14.51) | (11, 12.3) |
| BQSquare | 54.82 | 18.52 | (64, 23.78) | (64, 25.86) | (45, 17.03) | (59, 9.97) |
| BasketballPass | 19.46 | 20.07 | (27, 18.19) | (22, 36.55) | (7, 17.52) | (21, 6.42) |
| BlowingBubbles | 41.06 | 12.97 | (50, 14.8) | (28, 25.54) | (27, 23.49) | (38, 8.46) |
| RaceHorses | 24.82 | 37.03 | (33, 34.24) | (30, 56.54) | (16, 31.06) | (23, 16.27) |
| BQMall | 21.76 | 22.45 | (30, 28.17) | (20, 35.29) | (15, 17.57) | (28, 11.7) |
| RaceHorsesDS | 22.25 | 43.87 | (44, 55.3) | (39, 69.77) | (12, 31.53) | (21, 24.74) |
| PartyScene | 43.35 | 17.57 | (52, 20.51) | (44, 24.79) | (34, 16.93) | (51, 0.0) |
| BasketballDrill | 21.29 | 14 | (23, 13.02) | (20, 18.56) | (19, 13.52) | (20, 9.8) |
| MarketPlace | 13.41 | 12.67 | (27, 97.79) | (27, 97.79) | (8, 4.93) | (8, 2.74) |
| Cactus | 19.16 | 11.85 | (20, 0.0) | (20, 12.86) | (18, 11.99) | (18, 9.59) |
| RitualDance | 8.07 | 13.17 | (12, 11.83) | (8, 33.41) | (4, 7.78) | (4, 5.3) |
| BasketballDrive | 13.48 | 17.14 | (17, 16.52) | (14, 20.61) | (10, 17.31) | (16, 13.55) |
| BQTerrace | 24.33 | 21.28 | (28, 0.0) | (21, 30.17) | (16, 27.66) | (28, 12.6) |
| KristenAndSara | 10.94 | 4.17 | (12, 4.9) | (11, 5.48) | (10, 4.61) | (11, 3.57) |
| FourPeople | 13.08 | 4.4 | (14, 4.31) | (13, 5.31) | (13, 4.06) | (14, 3.7) |
| Johnny | 9.96 | 4.18 | (10, 4.9) | (10, 5.41) | (9, 4.69) | (10, 3.57) |
| ParkRunning3 | 10.74 | 19.19 | (12, 17.87) | (12, 23.67) | (9, 20.12) | (10, 15.83) |
| DaylightRoad2 | 14.98 | 11.66 | (16, 12.97) | (15, 14.52) | (14, 11.64) | (15, 10.26) |
| CatRobot1 | 11.85 | 8.89 | (12, 8.11) | (12, 10.2) | (11, 8.06) | (12, 7.86) |
| SlideEditing | 64.53 | 5.72 | (69, 0.93) | (67, 99.84) | (59, 99.16) | (67, 0.0) |
| SlideShow | 10.44 | 13.53 | (28, 0.0) | (0, 142.74) | (0, 0.54) | (9, 0.0) |
| BasketballDrillText | 25.93 | 16.94 | (30, 19.75) | (29, 23.21) | (21, 9.72) | (21, 9.72) |
| ArenaOfValor | 15.99 | 13.83 | (17, 5.46) | (16, 39.01) | (15, 14.92) | (16, 0.0) |

Comments:

Was comparison with SI/TI done? No, but could be done.

Analysis of the new sequences (gaming) should be coordinated with Mathias

Would be interesting to match the metrics with “outlier” behaviour of sequences regarding compression results (e.g., extraordinarily high gain compared to entire set.

Further study on this is welcome, also in coordination with AG 5. The topic is already in the mandates of AHG4.

(joint meeting with AG 5 closed 1520)

## Verification testing (0)

Section kept as a template for future use.

## Test material (1)

Contributions in this area were discussed in the joint meeting with AG 5 and VCEG at 1440–1450 UTC on Monday 18 July 2022 (chaired by JRO and MW).

[JVET-AA0123](https://jvet-experts.org/doc_end_user/current_document.php?id=11799) [AHG-7] Update on gaming sequences from InterDigital [T. Poirier, S. Puri, G. Martin-Cocher, E. Faivre d'Arcier (InterDigital)]

Was presented in joint meeting with AG 5

This contribution proposes an update on the new gaming sequences initially proposed in JVET-Y0041. The proposed sequences were created with Unity game engine and include depth and optical flow information. During former visual sessions, some issues were identified for some clips. This contribution completes JVET-Y0041 with updated versions of some content and more detailed information about the content, based on classification recommendations from JVET-AA0046. The compression results along with example bitstreams using VTM-17.0 and ECM-5.0 are not available but will be provided in the next meeting.

The contribution proposes to create a new class (class G) of content for gaming sequences, corresponding to cloud gaming and game casting use cases.

Some changes in content (modified rendering), to make them more suitable in particular for viewing.

The Excel sheet of v2 upload is defect.

Side activity (coordinated by M. Wien) was conducted to propose a selection of sequences for class “G”. It was reported in session 25 that rendering artifacts were observed in various sequences from the new set. It was suggested to defer the introduction of the new class until the issue is resolved.

It was further reported that an analysis using the method from JVET-AA0241 indicated both high spatial and temporal complexity characteristics in the sequences. It was however unclear whether this was caused by the rendering artifacts or is a property of the content itself.

Further study requested. It is understood that having a new class G set up is highly desirable.

## Quality assessment (0)

Section kept as a template for future use.

## Conformance test development (1)

Contributions in this area were discussed in session 20 at 0740–0755 UTC on Wednesday 20 July 2022 (chaired by JRO).

[JVET-AA0109](https://jvet-experts.org/doc_end_user/current_document.php?id=11785) Editors' update on conformance testing for VVC operation range extensions [D. Rusanovskyy (Qualcomm), T. Ikai (Sharp), H.-J. Jhu (Kwai), I. Moccagatta (Intel), Y. Yu (OPPO)]

This document is draft 4 of the conformance testing specification for VVC operation range extensions, for testing the new coding tool features being added in the second edition of VVC (Rec. ITU-T H.266 | ISO/IEC 23090-3).

Due to the size of the electronic attachments, the associated conformance test set data files are provided separately, at the following location:

<ftp://ftp3.itu.int/jvet-site/bitstream_exchange/VVCv2/under_test/VTM-14.0/>

(user id: avguest, password: Avguest201007)

For ftp access, it is suggested to use FileZilla, for which the site manager feature should be used and ftp encryption should be set to “Require implicit ftp over TLS” (see Annex A of this document).

Alternatively, the same files are available by http at:

<https://www.itu.int/wftp3/av-arch/jvet-site/bitstream_exchange/VVCv2/under_test/VTM-14.0/>

Annex A of this document contains JVET guidelines for contributing to the conformance bitstream test set. It is not expected to be published as part of the text approved by the parent bodies.

Purely editorial changes. As the DIS ballot is still open, there should not be an output according to ISO rules. It was also pointed out that some of the modifcations may not be needed for the ITU version (where it is planned to be integrated into a new edition of H.266.1). No need for a new output document JVET-AA2026.

It was further decided that the update/correction of v1 streams that was dicussed in the context of the AHG report should be integrated as a part of JVET-AA1004 (I. Moccagatta to send the text part to Y.-K. Wang).

## Software development (1)

Contributions in this area were discussed in session 20 at 0755–0810 UTC on Wednesday 20 July 2022 (chaired by JRO).

[JVET-AA0086](https://jvet-experts.org/doc_end_user/current_document.php?id=11762) AHG11: Small Ad-hoc Deep-Learning Library (SADL) update [F. Galpin, T. Dumas, P. Bordes, E. François (InterDigital)]

This contribution presents updates in the Small AdHoc Deep Learning (SADL) library already described in JVET-W0181, JVET-Y0110 and JVET-Z0161.

Changes:

* new sample programs:
  + count\_mac: count number of MAC/samples, overflow, number of operations
  + debug\_model: print information and issues on model (layers parameters, SIMD implementation etc.)
  + sample\_generic, sample\_simd256, sample\_simd512: test a model at various level of optimization
  + naïve\_quantization: convert a model from float to integer model knowing the layers quantizers
* added clang-format file
* change the Add layer quantizer policy: use the smallest quantizer instead of the 2nd input one. BiasAdd layer keeps the same policy.

Improvements:

* improve SIMD support for float (AVX512, conv1x1 etc.) (EE1.6, EE1.2 etc.)
* go back to HWCD loop order. Some gains on big models.
* New nbOutputs() method in Model
* Sanity check on quantizer values for some layers
* better inference debug info
* SSE42 for convolution in mod8 case (speedup JVET-Z0082)
* Update documentation

Corrections:

* init function is now re-entrant: it is possible to do inference with multiple input sizes by calling init with the new size
* remove all MAC counters in SIMD paths + add warning message
* carry over quantizer values for Expand and transpose layers (JVET-Z-EE-1.6)
* add conditions for special tensors multiplication case (JVET-Z0082)
* added leakyRelu case in naive\_quantization sample code (JVET-Z-EE-1.5)
* Cleanups:  
  clean the naming for conv2d (simd, dispatch etc.)
* clean messages for SIMD support
* clean macro for debug
* clean sample\_test.sh script
* clang format everything
* copyright year
* README update
* other minor cleanings

Various recommendations are given to implementers regarding normalization, input/output, receptive field size, dynamic quantization.

It was reported that various modifications on SADL were done by EE participants, but no merge request for integration was made.

Revisit – It was commented that the following items need to be discussed after deciding about a base software (see further notes under BoG report JVET-AA0247):

* Official status of SADL?
* Coordinators of software
* Repository/maintenance
* Separate AHG?
* Software guidelines?

## Implementation studies and complexity analysis (1)

Contributions in this area were discussed in session 20 at 0810–0830 UTC on Wednesday 20 July 2022 (chaired by JRO).

[JVET-AA0154](https://jvet-experts.org/doc_end_user/current_document.php?id=11842) 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 1.5.0 of the open VVC software encoder VVenC released in July 2022 and version 1.5.0 of the open VVC software decoder VVdeC released in April 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.

Main changes for VVenC v1.5.0 since version 1.3.1 include:

* GOP-based scene-cut detection (v1.4)
* Changed license to unmodified Clear BSD and added the authors as copyright holders (v1.4)
* Added features (v1.4):
  + Header reading functionality
  + Bit rate can be specified using common suffixes (Mbps, M, Kbps, K, and bps)
  + Integrated Y4M input support
* Improved presets:
  + Speedups: ~10% for faster, fast, medium, ~7% for slow and slower
  + BD-rate efficiency improvement: −1.3% for fast and faster
* Improved rate-control stability for noisy input and scene changes
* Arbitrary intra periods (not required to be multiples of GOP size)
* Added low-decoding-energy preset (~medium efficiency)
* Various fixes and improvements

VVEnc Changes: Scene cut detection, speedups and compression improvements, improved rate control, arbitrary intra periods (not multiples of GOP size), low-decoding-energy mode.

VVDec support for ARM and M1, allowing 1080p 10 bit 3 Mbit/s realtime playback on tablets, 5h nonstop with low energy mode.

How long would playback time be with one of the “normal” modes? Not known. However, it was found that the “encoder fast” modes increase energy consumption at decoder, whereas the energy saving mode requires more energy at the encoder (medium efficiency, but encoder speed more towards “slow”).

## AHG7: Low latency and constrained complexity (3)

Contributions in this area were discussed in session 20 at 0835–0930 UTC on Wednesday 20 July 2022 (chaired by JRO).

[JVET-AA0046](https://jvet-experts.org/doc_end_user/current_document.php?id=11721) [AhG4] Report on AhG4 meeting on development of a gaming-type CTC class [M. Wien (AHG chair)]

See section 4.3

[JVET-AA0117](https://jvet-experts.org/doc_end_user/current_document.php?id=11793) AHG-7: refining low delay configuration for cloud gaming [S. Puri, T. Poirier, P. Le Guyadec, A. Robert, K. Naser, G. Martin-Cocher, E. François (InterDigital)]

This contribution proposes an update of the low delay configuration that was agreed in the last meeting ([JVET-Z0114](https://jvet-experts.org/doc_end_user/current_document.php?id=11561)) by further removing tools that inherently cause delays and complexity in the processing pipeline on the decoder and encoder, namely template tools. It is argued that such disadvantages may jeopardize their implementation in practice particularly for the low delay applications.

Two sets of refinements of the LLCC baseline configuration are proposed. For the first proposed set, tools that use template matching are disabled. For the second proposed set, all the tools that use template are disabled including the template matching tools. Test 1 corresponds to the first set (set 1) and test 2 corresponds to the second set (set 2).

Simulations are performed on the latest ECM-5.0 test model software under the LDB-1Ref-asymmetric configuration (ECM-5.0 LLCC) based on [JVET-Z0114](https://jvet-experts.org/doc_end_user/current_document.php?id=11561), and average results on Class F and TGM are reported in the table below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Test (F + TGM) | BDR PSNRY | BDR PSNRU | BDR PSNRV | EncTime | DecTime |
| Test 1: Set 1 vs anchor | 4.23% | 4.71% | 4.66% | 76% | 77% |
| Test 2: Set 2 vs anchor | 6.06% | 6.37% | 6.43% | 65% | 66% |

Additionally, Test 2 results (under ECM-5.0 LLCC) are compared against VTM-17.0 LLCC for Class F and TGM and are reported in the table below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Test (F + TGM) | BDR PSNRY | BDR PSNRU | BDR PSNRV | EncTime | DecTime |
| Test 3: Set 2 vs VTM-17.0 LLCC | -15.61% | -24.10% | -23.16% | 328% | 235% |

It is asserted that the proposed configuration provides a better operating point due to significant reduction in both encoder and decoder time as well as reduction in decoder latency, however at an expense of lower coding efficiency.

It is proposed to adopt Set 2 as the default low delay configuration for at least class F, TGM and the newly proposed gaming class of sequences.

Does low delay configuration for cloud gaming include GDR? There was no agreement so far.

Main arguments are about low complexity. Is TM causing a low delay problem?

It was commented that also LDP might be interesting to consider (crosscheck contains results on that).

It was commented (from experience with VVEnc and other commercial deployments) that disabling tools with decoder side search is not highly beneficial for reduced encoding time. There could be decrease in decoder complexity, but it may partially depend on implementation). In reference implementations such as VTM and ECM, there are many other aspects of encoder search that can be optimized without losing as much performance as would be caused by disabling decoder side search tools. It is difficult to conclude from reference implementations which tools are preventing from low-complexity and real-time encoding.

Does intra template matching (other than inter template matching) really cause a decoder complexity problem? It is proposed to disable in set 2. For example, is there analysis of cycles needed, etc.? Not so far.

What would be a typical setting for cloud gaming applications in VVC?

* Would it use GDR for achieving low delay?
* Would it use only P or also B pictures?
* Would an encoder disable certain tools for being able to encode real-time? Probably yes, but the choice would likely be different depending on individual encoder implementations.
* Is decoder complexity an issue (probably not for VVC, and may be difficult to answer for future developments)

Several experts commented that disabling certain tools is not a good approach, as it may depend on specific application requirements (which are even different in the cloud gaming domain). Same may apply to the second bullet point.

It was suggested that a CTC using GDR should be developed from the background of cloud gaming application requirements (first starting for VTM, but easily extensible to ECM). Frequency of refresh points, clarify if it is better LDP or LDB, or both.

It was pointed out that GDR can also be differently implemented, but JVET has its own at current time.

Otherwise, it is good to define the new class G, but that can be used with any of the existing CTCs (as said in context of another discussion), or with a specific CTC for cloud gaming.

Further study on these aspects in AHG7.

[JVET-AA0227](https://jvet-experts.org/doc_end_user/current_document.php?id=11917) Crosscheck of JVET-AA0117 (AHG-7: refining low delay configuration for cloud gaming) [G. Li (Tencent)] [late]

Results on LDB confirmed. Also contains LDP results.

[JVET-AA0123](https://jvet-experts.org/doc_end_user/current_document.php?id=11799) [AHG-7] Update on gaming sequences from InterDigital [T. Poirier, S. Puri, G. Martin-Cocher, E. Faivre d'Arcier (InterDigital)]

See section 4.5

## AHG10: Encoding algorithm optimization (1)

Contributions in this area were discussed in session X at XXXX–XXXX UTC on XXday XX July 2022 (chaired by JRO).

[JVET-AA0129](https://jvet-experts.org/doc_end_user/current_document.php?id=11805) Non-EE2/AHG10: Improved inter hash RDO considering OBMC off in ECM5.0 [X. Xiu, C.-W. Kuo, H.-J. Jhu, W. Chen, N. Yan, X. Wang (Kwai)]

See discussion in 5.3.4

## Profile/tier/level specification (1)

Contributions in this area were discussed in session 21 at 1300–1330 UTC on Wednesday 20 July 2022 (chaired by JRO).

[JVET-AA0239](https://jvet-experts.org/doc_end_user/current_document.php?id=11929) Multiview profiles in the HEVC standard [A. Tourapis, D. Podborksi (Apple)] [late]

This contribution proposes the creation of several new multiview profiles for the HEVC standard with 10 bit support.

It is suggested that this could be useful for V-PCC.

Several experts supported this proposal. It is straightforward to be implemented, and draft text is attached in the contribution.

It was suggested to issue a preliminary WD for a new amendment of HEVC. Hypothetically, this might later be included in the ongoing progression of the new HEVC edition.

Proponents are asked to generate and propose related conformance bitstreams until the next meeting.

It is noted that currently multiview profiles are not included in the HM software, only in HTM, and some checking might be necessary if that works properly with 10 bit.

It is also mentioned in the contribution that, whereas the proposal includes a monochrome 10 bit profile for multiview, an equivalent scalable monochrome 10 bit profile has never been specified (but there is a scalable monochrome 12 bit).

Decision: Adopt JVET-AA0239, generate a draft 1 of specification text (JVET-AA1011)

## Proposed modification of system interface (0)

Section kept as a template for future use.

# Low-level tool technology proposals

## AHG8: High bit rate and high bit depth coding for VVC (1)

Contributions in this area were discussed in session 2 at 0850–0930 UTC on Wednesday 13 July 2022, and session 7 at 1300–XXXX UTC on Thursday 14 July 2022 (chaired by JRO).

[JVET-AA0130](https://jvet-experts.org/doc_end_user/current_document.php?id=11806) AHG8: Draft VTM and HM common test conditions for high bit depth and high bit rate video coding [A. Browne, T. Ikai, D. Rusanovskyy, X. Xiu, Y. Yu (AHG chairs)]

## AHG11: Neural network-based video coding (24)

### Summary and BoG reports

Contributions in this area were discussed in session 2 at 0850–0920 UTC on Wednesday 13 July 2022 and in session 8 at 1520–1625 UTC on Thursday 14 July 2022 (chaired by JRO).

[JVET-AA0023](https://jvet-experts.org/doc_end_user/current_document.php?id=11879) EE1: Summary of Exploration Experiments on Neural Network-based Video Coding [E. Alshina, W. Chen, F. Galpin, Y. Li, Z. Ma, H. Wang, L. Wang]

This document summarizes Exploration Experiment 1 (EE1) tests performed between the JVET-Z and JVET-AA meetings to evaluate **Neural Network-based Video Coding (**NNVC) technologies, analyze their performance, and analyze their complexity aspects. In NN-based in-loop filter category BD-rate gain over AhG11 anchor goes up to **9~10%** (depending on complexity). By content adaptive post-filter gain of **5%** over AhG11 anchor was demonstrated. In Super Resolution category **1.5** % BD-rate average gain over AhG11 anchor can be achieved (among them more than 3% for 4K content). End-to-End AI video codec in this round of EE1 demonstrated almost 5% improvement, but still cannot reach VTM performance.

**Introduction**

Group continues evaluation of new promising NN-based video coding technologies, answering questions and addressing suggestions from JVET members made during presentation NN-based technologies at JVET-Z meeting. Additionally in this round of EE the group focused on investigating platform-independent reproducibility (which is desirable feature for video codec [1]), drift-free loop operation by integerization of NNs, and usage of a software package which supports that.

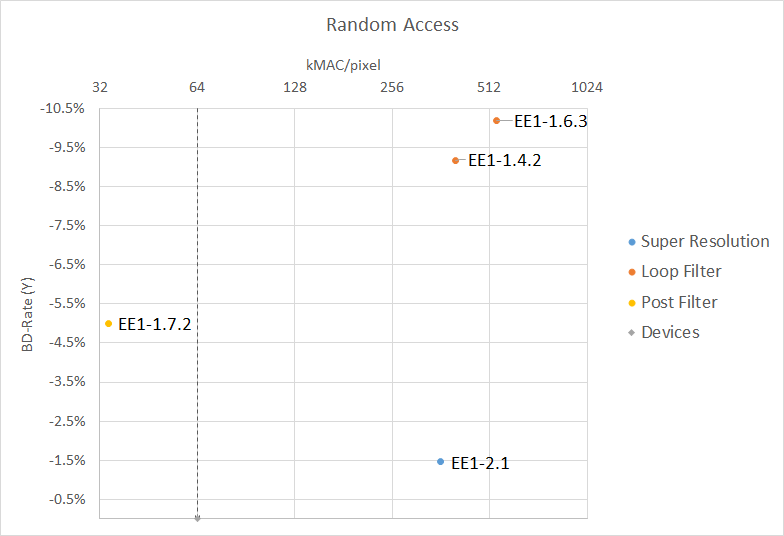
The ultimate goal of the NNVC Exploration Experiments is the creation of AhG11 anchor based SW with examples of NN-tools implementation for further collaborative development of NN-based video coding tools. It was expected that 2 NN-based tools recommended by JVET to be added to common SW basis at JVET-Z meeting **undergo procedure of cross-check for the training**.

Tests were conducted in three categories: enhancement filters (in-loop or post), super-resolution and E2E NN-based video coding. One test was performed both on top of VTM and ECM anchors.

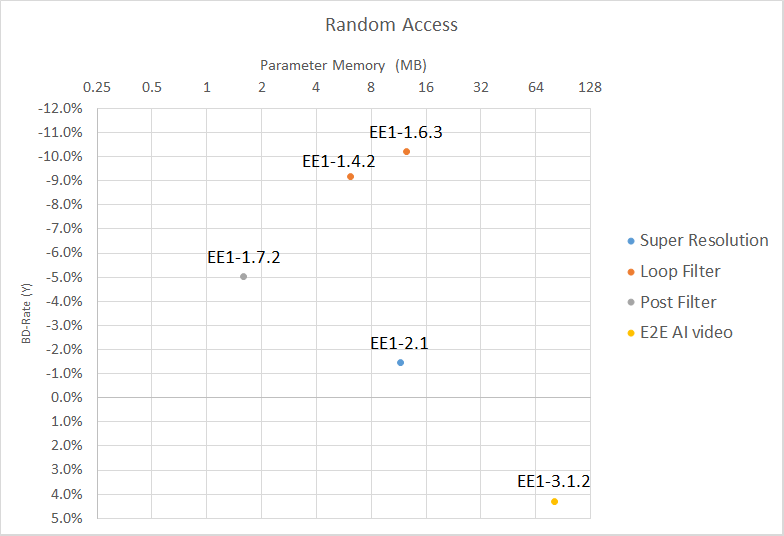
All proponents **used** AhG11 anchor [2] (VTM-11.0 + “newMCTF” patch [3], QP=22, 27, 32, 37, 42) and the reported template recommended by AhG11.

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

Around 5% gain can be obtained with NN-based post-filter. Twice higher gain (~10%) can be achieved with NN-based in-loop filter. NN-based super-resolution provides gain primarily for 4K content, so in average performance is quite modest (1.5%). End-to-end video compression tested in this round of EE1 still cannot reach VTM performance.



*Figure 1 Gain vs computational complexity for ttypical representatives of tests in EE1: loop-filters (EE1-1.4 “Tencent” architecture and EE1-1.6 “Bytedance” architecture), post filter (EE1-1.7) and NN-based super-resolution (EE1-2.1.)*



*Figure 2****.*** *Gain vs memory for typical representatives of tests in EE1: loop-filters (EE1-1.4 “Tencent” architecture and EE1-1.6 “Bytedance” architecture), post filter (EE1-1.7), NN-based super-resolution (EE1-2.1) and End-to-End AI video codec (EE1-3.1.2)*

**Exploration experiments on Enhancement filters**

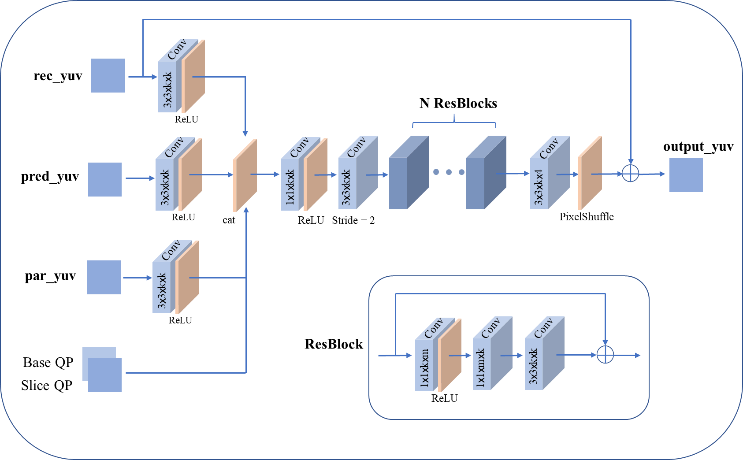
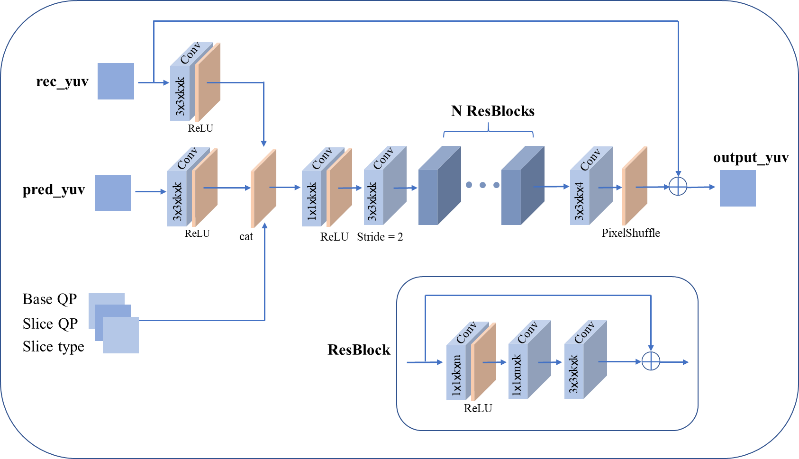
Tests status is as follows:

* 7 tests have been completed (EE1-1.1...1.7);
* One test (EE1-1.8) was withdrawn
* 15 cross-checks registered (5 still not available), 3 cross-checks for training have been completed;
* For 1 test (EE1-1.2, [JVET-AA0081](https://jvet-experts.org/doc_end_user/current_document.php?id=11757)) inference was successfully cross-checked; minor deviation (0.00004% BD-rate) in performance due to float point arithmetic was reported;
* Cross-checker of training for EE1-1.5 ([JVET-AA0178](https://jvet-experts.org/doc_end_user/current_document.php?id=11866)) was able to complete training, performance difference with model from proponent is 0.04% (RA configuration).
* Cross-checker of training for EE1-1.5 ([JVET-AA0](https://jvet-experts.org/doc_end_user/current_document.php?id=11866)217) was able to complete training, performance difference with model from proponent is ~0.02% (RA configuration).
* Cross-checker of training for EE1-1.6 ([JVET-AA0181](https://jvet-experts.org/doc_end_user/current_document.php?id=11869)) was able to complete training, but performance difference with model from proponent is 0.39% (RA configuration).
* Cross-checker of training for EE1-1.5 and EE1-1.6 was not able to complete training ([JVET-AA0174](https://jvet-experts.org/doc_end_user/current_document.php?id=11862)), but came up with practical suggestions for training procedure unification
* EE1-1.1 was partially cross-checked (only class D RA). Minor deviations {0.001%, 0.005%, 0.003%} in BD-rate Y for EE1-1.1.1, EE1-1.1.2 and EE1-1.1.3, respectively, reported due to floating point arithmetic.

Among ***in-loop filters*** two architectures were recommended during last meeting to be considered for adoption to common SW base for NN-based video coding tools study. One architecture was originated by Tencent (later Interdigital joined design and OPPO working on further improvements for this architecture). This family of in-loop fillers is summarized in section 2.1. Another promising architecture was initiated by Bytedance further improved by Qualcomm, Ericsson and Interdigital. Test based on this architecture are summarized in section 2.2. Post-filter architecture was originated by Qualcomm, tested in EE1 as content adaptive post-filter in this EE1 and described in section 2.3.

***Family of in-loop filters (originated by Tencent).***

There are 2 major variants: with one (model size 1.9 M parameters, Figure 3 right) and two models (total models sizes is 3.1 millions parameters, Figure 3 left). Prediction, Partitioning, QP and Slice Type (only for variant with 1 model) information are taken as extra inputs to this NN-based filters. All three color planes processed jointly.

*Figure 3 NN-filter architecture from Tencent: 2 models (left) and one model (right).*

Reference number for both those filters as tested during previous EE1 round are shown in grey font in Table 1. Filter parameters were partially fine-tuned, to be better with SADL implementation, so 0.2~0.3% performance improvement in RA configuration was achieved (if NN is quantized to Int 16).

Cross-check for training stage (major goal of this EE1 round) put into the box: EE1-1.5.0 (model trained and tested by proponent), EE1-1.5.1 , 2, 3 (model trained by cross-checker and tested by proponent). Different cross-checkers (OPPO and Qualcomm) and two slightly different training strategies were used, but in all cases difference in within 0.04% BD-rate (in RA cfg).

*Table 1 In-loop filters based in “Tencent” architecture vs VTM anchor.*

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | |  | Random Access (CTC) | | | All Intra (CTC) | | | |
| test | Source | Precision | Total Parameters (Millions) | | kMAC/pixel | Y | Cb | Cr | Y | Cb | | Cr |
| EE1-1.4.1.0 | JVET-Z0094 | Float 32 | **3.1** | 401.0 | | **-8.8%** | -19.9% | -19.4% | **-7.4%** | -16.8% | -17.3% | |
| EE1-1.4.2.0 | JVET-Z0094 | Float (Int 16) | **3.1** | 401.0 | | **-8.8%** | -19.9% | -19.4% | **-7.4%** | -16.8% | -17.3% | |
| EE1-1.4.1 | [JVET-AA0087](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0087-v1.zip) | Float 32 | **3.1** | 401.0 | | **-9.2%** | -19.1% | -19.6% | **-7.4%** | -16.8% | -17.3% | |
| EE1-1.4.2 | [JVET-AA0087](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0087-v1.zip) | Float (Int 16) | **3.1** | 401.0 | | **-9.2%** | -19.9% | -19.8% | **-7.5%** | -16.6% | -17.6% | |
| EE1-1.3.1 | [JVET-AA0122](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0122-v1.zip) | Float 32 | 3.1 | 401.0 | | **-9.0%** | -20.0% | -19.5% | **-7.4%** | -16.8% | -17.3% | |
| EE1-1.3.2 | [JVET-AA0122](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0122-v1.zip) | Float 32 | 1.9 | 485.0 | | **-9.0%** | -19.1% | -19.3% | **-6.5%** | -14.9% | -16.0% | |
| EE1-1.5.0 | [JVET-Z0091](https://jvet-experts.org/doc_end_user/current_document.php?id=11526) | Float 32 | **1.9** | 485.0 | | **-8.68%** | -18.6% | -19.0% | **-6.50%** | -14.9% | -16.0% | |
| EE1-1.5.1.1 | [JVET-AA0088](https://jvet-experts.org/doc_end_user/current_document.php?id=11764) | Float 32 | **1.9** | 485.0 | | **-8.78%** | -19.4% | -19.5% | **-6.47%** | -15.1% | -15.8% | |
| EE1-1.5.1.2 | [JVET-AA0088](https://jvet-experts.org/doc_end_user/current_document.php?id=11764) | Float 32 | **1.9** | 485.0 | | **-8.64%** | -18.0% | -19.3% | **-6.35%** | -14.4% | -15.8% | |
| EE1-1.5.1.3 | [JVET-AA0088](https://jvet-experts.org/doc_end_user/current_document.php?id=11764) | Float 32 | **1.9** | 485.0 | | **-8.64%** | -19.6% | -19.6% | **-6.42%** | -14.3% | -15.4% | |
| EE1-1.5.2 | [JVET-AA0088](https://jvet-experts.org/doc_end_user/current_document.php?id=11764) | Float 32 | **1.9** | 485.0 | | **-8.8%** | -18.9% | -19.0% | **-6.5%** | -15.4% | -16.5% | |
| EE1-1.5.3 | [JVET-AA0088](https://jvet-experts.org/doc_end_user/current_document.php?id=11764) | Float (Int 16) | **1.9** | 485.0 | | **-9.0%** | -18.5% | -19.2% | **-6.5%** | -15.1% | -16.4% | |

In test EE1-1.3 (by OPPO) ~0.3% performance improvement was achieved by a slice-level adjustment for QP which used as input parameter for filtration. Instead of BaseQP (original design) selected by encoder and signalled per slice finalBaseQP will be used in filtration process.

It looks promising to combine two potential performance improvements for this architecture in the next round of EE1.

*Table 2 In-loop filters based in “Tencent” architecture vs ECM anchor*

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | |  | Random Access (CTC) | | | All Intra (CTC) | | | |
| test | Source | Precision | Total Parameters (Millions) | | kMAC/pixel | Y | Cb | Cr | Y | Cb | | Cr |
| vs VTM | [JVET-Z0091](https://jvet-experts.org/doc_end_user/current_document.php?id=11526) | Float 32 | **1.9** | 485.0 | | **-8.7%** | -18.6% | -19.0% | **-6.5%** | -14.9% | -16.0% | |
| EE1-1.1.1 | [JVET-AA0085](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0085-v1.zip) | Float 32 | 1.9 | 485.0 | | **-5.5%** | -13.2% | -13.7% | **-3.9%** | -8.5% | -10.4% | |
| EE1-1.1.2 | [JVET-AA0085](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0085-v1.zip) | Float 32 | 1.9 | 485.0 | | **-5.6%** | -13.0% | -13.0% | **-3.6%** | -7.4% | -8.9% | |
| EE1-1.1.3 | [JVET-AA0085](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0085-v1.zip) | Float 32 | 1.9 | 485.0 | | **-5.9%** | -13.7% | -13.7% | **-4.0%** | -8.9% | -10.3% | |

This architecture (original design, w/o improvements) was tested on top of ECM-4.0. For the reference performance vs VTM anchor is shown in grey font in Table 2. Overlap between ECM and BNN-based filter gain is about 3%. If VTM trained filter is directly applied to ECM (Test EE1-1.1.1) the performance is almost the same as if filter specifically trained for ECM artifacts (Test EE1-1.1.2). Interestingly the fine tuning of VTM trained filter for ECM generated artifacts (Test EE1-1.1.3) provides the best performance (almost 0.3% extra gain). This might indicate that longer training for “Tencent” filter might bring additional performance improvement.

***Family of in-loop filters (originated by Bytedance)***

NN-filter architecture used for tests in this sub-category is shown on Figure 4. Prediction, Boundary Strength, Partitioning and QP information are taken as extra inputs to this NN-based filters. Also attention Residual Block is used as basis element.

**

*Figure 4 NN-filter architecture from Bytedance.*

*Table 3 In-loop filters based in “Bytedance” architecture vs VTM anchor.*

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | |  | Random Access (CTC) | | | All Intra (CTC) | | | |
| test | Source | Precision | Total Parameters (Millions) | | kMAC/pixel | Y | Cb | Cr | Y | Cb | | Cr |
| EE1-1.2.0 | [JVET-Z0106](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0106-v1.zip) | Float 32 | **6.24** | 649.0 | |  |  |  | **-7.6%** | -18.5% | -19.1% | |
| EE1-1.2.1 | [JVET-AA0081](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0106-v1.zip) | Float 32 | **6.21** | 635.0 | |  |  |  | **-7.6%** | -18.5% | -19.1% | |
| EE1-1.2.2 | [JVET-AA0081](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0106-v1.zip) | Float 32 | **6.20** | 625.0 | |  |  |  | **-7.5%** | -18.6% | -19.2% | |
| EE1-1.6.1.0 | JVET-Z0113 | Float 32 | **6.24** | 649.0 | | **-10.18%** | -22.2% | -22.1% | **-7.24%** | -19.7% | -20.0% | |
| EE1-1.6.1 | [JVET-AA0111](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0111-v1.zip) | Float 32 | **6.24** | 649.0 | | **-9.79%** | 22.3% | -22.8% | **-7.34%** | -20.4% | -21.0% | |
| EE1-1.6.2 | [JVET-AA0111](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0111-v1.zip) | Float 32 | **6.24** | 539.0 | | **-10.2%** | -22.1% | -21.8% | **-7.3%** | -20.2% | -20.5% | |
| EE1-1.6.3 | [JVET-AA0111](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0111-v1.zip) | Int 16 | **6.24** | 539.0 | | **-10.2%** | -22.1% | -21.7% | **-7.3%** | -20.1% | -20.6% | |

Reference numbers for performance complexity as technology was tested at the last meeting are shown in grey font. Training stage cross-check is in the box. Model was trained by cross-checker (PKU) and tested by proponent. Difference to the original proposal performance is 0.39% (in RA cfg).

The major goal of Test EE1-1.2 was to search for the better complexity performance trade-off. In Test EE1-1.2.1 partitioning and attention are removed from design. In Test EE-1.2.2 additionally Boundary Strength removes from the design. Unfortunately tests was performed only for Intra model (in all-intra configuration). But at least for Intra coding those elements of this NN-filter design (partitioning, attention and BS) can be removed providing 5% saving in computation complexity, almost w/o performance degradation.

Test EE1-1.6.1 is supposed to be verification of training procedure. Model was trained by cross-checker and tested by original proponent. It should be noticed that model trained by proponent provides 0.4% lower gain than original model by proponent. Unfortunately by time this report was prepared, cross-check report has not been uploaded. Training cross-check results and procedure should be discussed during JVET-AA meeting.

In Tests EE1-1.6.2 and 3 Chroma filters was forced to be the same for both Chroma components. This resulted in impressive reduction of computational complexity: 649 kMAC/pxl to 539 kMAC/pxl. This was possibility noticed last meeting and confirmed during this EE1 round.

Quantizing NN-filter to 16 bits was performed in test EE1-1.6.3. Similar to other attempts for NN-filter parameters quantizer to 16 bits integer, almost no performance variation compared to 32bit float precision was found.

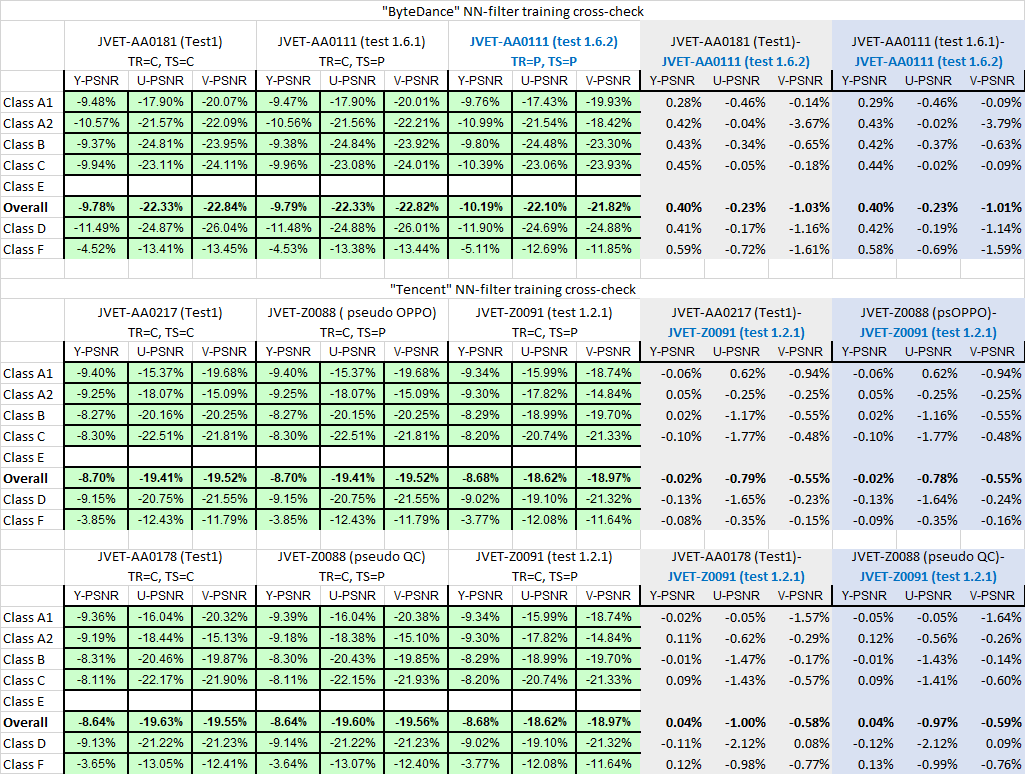
***Status of training cross-check***

Several attempts for the training process cross-check were taken. Status of this cross-check is summarized in Table 4 . Training was conducted by proponent and cross-checker and “cross-tested” (proponent tested model generated by cross-checker). “TR” – training, “TS” – test, “C” – cross-checker, “P” –proponent.

It was easier to match performance of “Tencent” NN-filter. Several cross-checkers reported the problem with data loader for “Bytedance” NN-filter training scripts. Eventually one of the cross-checkers (PKU, JVET-AA0181) modified data loader and was able to complete training.

Procedure for the training cross-check requires discussion during the meeting. It is desirable to have training procedure unified for different NN-based technologies (at least for enhancement filters). This can be part of discussion about potential common SW for NNVC tools.

Table 4. Cross-check of training, Random Access test results are show.



***Family of post-filter***

NN-based post-filter architecture is shown on Figure 5. Bias parameters are chosen content adaptively and transferred together (using SEI message syntax).

*Figure 5 NN-based Post-filter architecture.*

Quantizing to 16 and 32 Int precision and implementation with SADL and comparison with C++ tensor flow implementation was the purpose of the study. Results are shown in Table 4. Again it is shown that quantizing to 16 bits integer doesn’t have noticeable effect on the performance. SADL appears to be slower than Tensor Flow, likely because it was not optimized for Float 32 and Int 32. Also filter architecture operates with basic blocks 72 or 24, which are not multiple of 16, so SADL was not optimized for those sizes (SIMD used instead).

*Table 4 Complexity and performance on Content adaptive PNN-based Post-filter vs AhG11 anchor (different implementations).*

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | |  | Random Access (CTC) | | |  | | |
| test | Source | Precision | Total Parameters (Millions) | | kMAC/pixel | Y | Cb | Cr | Dec. Time | Comments |  |
| EE1-1.7.0 | [JVET-AA0066](https://jvet-experts.org/doc_end_user/current_document.php?id=11742) | Float 32 | 0.4 | 34.0 | | **-5.0%** | -18.9% | -17.3% | ×46 | C++ TF | |
| EE1-1.7.1 | [JVET-AA0066](https://jvet-experts.org/doc_end_user/current_document.php?id=11742) | Float 32 | 0.4 | 34.0 | | **-5.0%** | -18.9% | -17.3% | ×158 | SADL, Float 32 | |
| EE1-1.7.2 | [JVET-AA0066](https://jvet-experts.org/doc_end_user/current_document.php?id=11742) | Int 32 | 0.4 | 34.0 | | **-5.0%** | -18.9% | -17.3% | ×336 | SADL, Int 32 | |
| EE1-1.7.3(\*) | [JVET-AA0066](https://jvet-experts.org/doc_end_user/current_document.php?id=11742) | Int 16 | 0.4 | 34.0 | | -4.3% | -18.6% | -16.0% | ×304 | SADL, Int 16 | |

(\*) classes A1, A2 are missed

***Open question on kMAC/pxl computation***

For all NN-based filters studied in this EE1 extended block is fed to the NN-filter. And block w/o extension is output of the filter. So multiplication operations are performed on higher amount of samples than actual output. Assuming whole picture is processed the effect of block extension is negligible. Under this assumption kMAC/pxl was computed in all tests and reported in the summary (including this EE summary). But in reality NN-filter will be implemented to process signal on a block basis. Two proponents reported kMAC/pxl both assuming picture level and block level processing (~27% higher). It is recommended to clarify which kMAC/pxl shall be computed and update NNVC CTC (or at least EE1 description) accordingly. Block based processing seems to be a reasonable assumption.

**NN-based super-resolution**

In this round of EE1 only one test was performed. Adaptive coded frame size and QP selection was combined with NN-based up-sampling filter on a clean way.



*(a)*

*(b)*

*Figure 6 Architecture of NN-based filter for super-resolution (a) with Advanced Residual Block (b).*

Compared to previous round of EE1 (shown in Table 5 in grey font) performance improvement of ~0.2% in average is demonstrated, also computational complexity and memory size for the model size have been reduced (854 🡪 361 kMAC/pxl, 7 🡪 2.9 millions). This is definitely a good move, but still NN-based super-resolution is relatively heavy in terms of computational complexity. Likely architecture of NN-based super-resolution filter (Shown on Figure 6, 16 Advanced residual blocks(!)) can be simplified.

*Table 5 BD-rate performance vs AhG11 anchor for NN-based super-resolution filter.*

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | |  | Random Access (CTC) | | | All Intra (CTC) | | | |
| test | Source | Precision | Total Parameters (Millions) | | kMAC/pixel | Y | Cb | Cr | Y | Cb | | Cr |
| [EE1-2.1.1](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0065-v1.zip) | [JVET-Z0065](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0065-v1.zip) | Int 16 | 0.0 | 0.0 | | **-1.0%** | 2.6% | 2.5% | **-0.6%** | 3.1% | 2.7% | |
| EE1-2.1.0 | JVET-Z0098 | Float 32 | 7.0 | 854.0 | | **-1.3%** | -0.6% | -0.8% | **-1.1%** | -0.4% | -0.7% | |
| EE1-2.1 | [JVET-AA0071](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0071-v1.zip) | Float 32 | 2.9 | 361.0 | | **-1.5%** | 1.0% | 1.1% | **-1.2%** | 1.4% | 1.3% | |

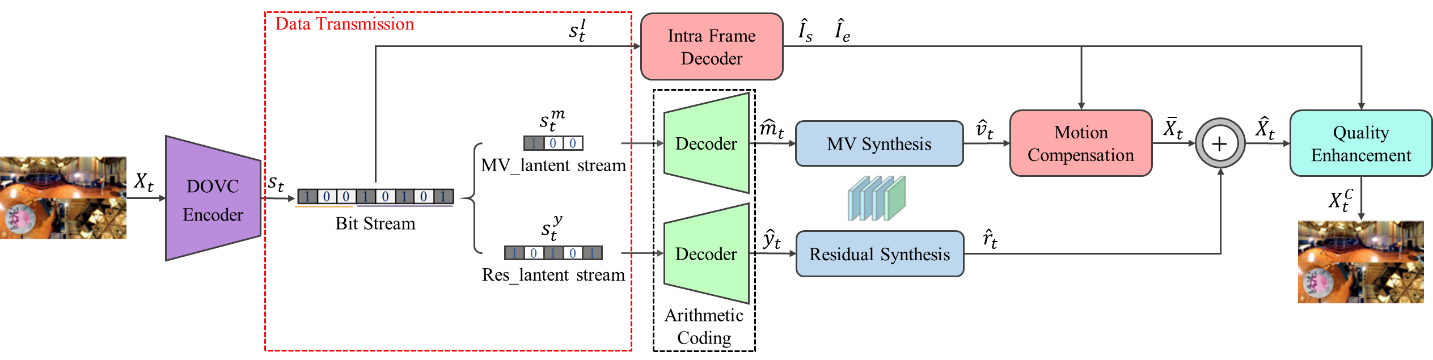
Gain is reported in average across all classes. It should be noticed that major gain is observed for 4K classes.

Note that results of JVET-Z0065 reflect a method of adaptively using RPR upsampling. JVET-AA0071 uses a similar approach. Gain is coming from only those cases where reduced resolution coding is used, whereas the numbers are averaging over all sequences. It would be interesting to analyze how large the gain compared to RPR upsampling is when only those cases where it is used are considered.

It was later reported verbally that the gain over RPR goes up to more than 6% for Food Market and Tango. It was commented that these were also sequences where subjective benefit was identified.

**End-to-End AI video coding**

In this round of EE1 one technology was tested. NN architecture which was used for End-to-End AI video codec is shown on Figure 7.



*Figure 7 End-to-End Ai video codec architecture.*

The purpose of experiment was to replace first (Intra) frame coding by BPG (original design, Test EE1-3.1.1) by VTM11.0 Intra coding (Test EE1-3.1.2). Results were provided w/o class A2. Computational complexity information is not found in the submission.

RD-curves for anchor and tests cross each other, difference to target rates in 20-30% (which makes BD-rate measurement not reliable). Similar problems appeared and have been solved in NN-based super-resolution. Hopefully they will be resolved for End-to-End video coded in a future.

*Table 6. Performance (w/o A2 class) and complexity for EE1 tests in End-to-End AI video codec category.*

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | |  | Random Access (CTC) | | | All Intra (CTC) | | | |
| test | Source | Precision | Total Parameters (Millions) | | kMAC/pixel | Y | Cb | Cr | Y | Cb | | Cr |
| EE1-3.1.1(\*\*) | [JVET-AA0059](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0059-v2.zip) | Float 32 | 81.0 | ? | | 9.8% | 9.1% | 8.0% |  |  |  | |
| EE1-3.1.2(\*\*) | [JVET-AA0059](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0059-v2.zip) | Float 32 | 81.0 | ? | | 4.3% | 4.0% | 3.7% |  |  |  | |

(\*\*) class A2 is missed

Test results in Table 1 confirm that replacement of BPG by VTM I-frame coding improves performance in random access configuration by more than 5%. It should be noticed that Intra frames in E2E Ai video coding solution are inserted more often compared to anchor (intra period 64 for almost all test videos).

It also should be noticed that, for example, in CVPR 2022 CLIC challenge in video coding section, none of End-to-End AI video coding solution was able to compete with VTM or ECM based solutions. Considering this only 4.3% drop relatively to AhG11 anchor (VVC) is not that bad.

Likely distinguishing features of this E2E AI based video coding bi-directional motion fusion. Motion field and residual are coded are picture.

In the discussion, it is pointed out that the performance could probably be improved when less frequent I frames are used (same period as in VTM anchor), and the other key frames would be encoded as B pictures with VVC.

Conclusions

* For the training cross-check in two cases training was completed by non-proponent. Performance difference in RA cfg between model trained by cross-checker and proponent is 0.04% (for “Tencent” architecture) and 0.39% (for “Bytedance” architecture).
* SADL allows transparent and reproducible implementation of NN-based tools, but so far it is slower than libtorch or/and TensorFlow.
* Performance improvement or/and complexity reduction for both two major architectures of NN-based in-loop filter have been demonstrated in this EE1 round.
* Content adaptive NN-based post-filter provides 5% gain with moderate complexity (but content dependent parameters).
* End-to-End AI based video coding still cannot reach VVC performance, but progressing quickly.
* The performance NN-based super-resolution was improved by 0.2%, also design was significantly simplified.

Recommendations

* Discuss cross-check for training stage and re-fine procedure during the JVET-AA meeting. Encourage further efforts on training stage verification.
* Discuss the adoption and conditional adoption of tools into the common SW base.
* Encourage modifications of SADL library for missing accelerated path and features.
* Computational complexity measurement (kMAC/pxl) shall be clarified and aligned with actual implementation (assumption for the block-base implementation is recommended).
* More fruitful ideas are expected from EE1-related proposals, which shall be reviewed in details.
* Select promising technologies among EE1 to undergo training stage cross-check during next EE1 round.

Follow-up in session 8

* New version of the summary report, more detailed analysis of training cross-check included.
* The Bytedance method has relatively larger deviation (0.4%) in luma than the other two methods, but all methods have also deviations in chroma (>1% sometimes)
* It was pointed out that the quantization/integerization was applied differently in the different proposals. A common approach as supported by SADL (also including option using integer in training) should be used in future experiments
* The possible reasons for deviations, and also difficulties that popped up in the training cross-check should be tried to be identified when reviewing the proposals.
* It was reported that block extension (for avoiding block boiundary artifacts in das of block-wise processing) was not considered in the kMAC/pixel computations of some proposals
* Recommendations above are agreed.

[JVET-AA0047](https://jvet-experts.org/doc_end_user/current_document.php?id=11722) [AHG 11] Brief information about JPEG AI CfP status [E. Alshina, J. Ascenso, T. Ebrahimi, F. Pereira, T. Richter]

This informative contribution shares the status of JPEG AI CfP: goal of the standard, testing conditions, three-phase of the CfP submission; objective performance results and examples of visual quality. In total 10 teams responded to the CfP in standard reconstruction category; five successfully passed cross-check. Among them, three submissions show substantial (20~30%) BD-rate gain over VVC anchor with ×30 slower decoder (CPU run time).

Presented in session 15, 0905-0935 UTC (chaired by JRO)

BD rate based on 7 different metrics, averaged. Tendency deviations between different metrics can be detected, but are not excluded in the averaging. As one example, BD rate between the VVC and HEVC anchors is roughly 13% based on that, whereas it is known to be close to 25% based on PSNR BD rate (note that none of these is targeting optimization for still image coding).

Some examples from the visual test are included in the presentation deck, JPEG intends making all images from the subjective test available.

Reconstruction on different devices limited to 0.5% BD-rate (tentatively).

Final standard targeted for 24/01.

Contributors are thanked for the valuable summary.

[JVET-AA0247](https://jvet-experts.org/doc_end_user/current_document.php?id=11938) BoG on Neural Network Video Coding (NNVC) [A. Segall, E. Alshina]

This is a report of activities from the BoG on Neural Network Video Coding (NNVC). The BoG held the following meetings during the 27th JVET meeting:

* July 19 – 05:00-7:00 UTC
* July 19 – 15:35-17:35 UTC
* July 20 – 15:20-17:00 UTC

The BoG activity is summarized as:

* The BoG reviewed the training crosschecks, including:
  + Summarizing key takeaways as:
    - The crosschecks took on the order of weeks for training.
    - The crosschecks took on the order of weeks for data generation. In some cases, this was further increased due to data transfer.
    - The training scripts provided by proponents did not work in all compute environments
    - The crosschecking effort would be improved by defining the naming convention for common training data, such as the number of coded frames and QPs.
    - Re-training of the model leads to a worst-case performance deviation of 0.7% in LDB between Option 1 and Option 2 training. The worst-case performance deviation for RA was 0.4%.
  + Recommendation: Consider the proposals sufficiently crosschecked
* The BoG developed recommendations for executing and reporting crosscheck information, including:
  + Developed three options for conducting a training crosscheck
    - Option 1: Generate data from scratch and train model from scratch
    - Option 2: Generate data using an existing model and train model from scratch
    - Option 3: Generate data using an existing model and initialize training using an existing model.
  + Recommendation: Allow the use of Option 1, Option 2 and/or Option 3 for future crosschecks. A crosschecker may select the most appropriate method but must report what method was used in the crosscheck report.
* The BoG developed recommendations for creating common software and methods, including:
  + Recommended that the common software have the following requirements:
    - Encoder software
      * Support for inference
    - Decoder software
      * Common data dumper with support for VTM11
      * Common data dumper with support for VTM11+reference models
      * Support for inference
    - Training scripts
      * Support for Option1, Option 2 and Option 3 training
  + Created a software development plan. (See table in notes below.)
  + Recommendation: Develop the common software using the software development plan.
  + Recommendation: Include design choices that (i) data dumper should output data in a frame based and planar format, (ii) data loader should use a JSON format for its configuration file, and (iii) software should be drift free.
  + Recommendation: Franck Galpin, Yue Li, Hongtao Wang, and Liqiang Wang be the software coordinators of the common software project.
  + Recommendation: Add a mandate to the NNVC AHG to develop the common software base, prepare and release anchor data for the common software base, and prepare and release verification data of the training scripts included in the common software base.
* As part of the common software discussion, the BoG also discussed the use of the common software in EE1, including:
  + Recommendation: EE should use the existing software anchor but give participants the option to use the common software when its available.
  + Recommendation: Proposals in the EE that include improvements on JVET-AA0088 and/or JVET-AA0111 be required to report results relative to these networks in addition to reporting results relative to the anchor described in the CTC.

The BoG was established with the following mandates:

* Summarize the takeaways from the training crosschecks
* Determine if common software base and/or methods can be defined
* Develop and/or refine recommendations for executing and reporting crosscheck information

|  |  |  |
| --- | --- | --- |
| **Common Software Development Plan** | | |
| Task | Description | Draft Timeline |
| Build Environment | Include support for the SADL library in the common software and build environment |  |
| Data dumper | Create a generic data dumper and integrate into the VTM decoder.  Data dumper able to store:   * Reconstructed sample value array * Prediction sample value array * Partition information array * Boundary strength information array * Base QP * Slice QP * Slice type     Data dumper enabled on the command line.  Data dumper easily configured to store a subset of the data listed above. |  |
| Data Loader | Create a generic data loader that is able to load data from:   * DIV2K * BVI-DVC * TVD * Data dumper output   Data loader easily configured to load a subset of the data provided in the data dumper output. |  |
| Incorporate JVET-AA0088 and JVET-AA0111 | Incorporate JVET-AA0088 and JVET-AA0111 into VTM encoder and decoder.  Networks are based on SADL.    Networks should use the same API that takes the following as input:   * Reconstructed sample value array * Prediction sample value array * Partition information array * Boundary strength information array * Base QP * Slice QP * Slice type     Networks should provide API that output the following:   * Filtered sample value array     Networks enabled through the encoder config file and command line.  A flag in the SPS enables the neural network loop filter, and a second flag in the SPS selects between JVET-AA0088 and JVET-A0111.  These bits indicate the use of the networks in the bit-stream for the decoder. |  |
| Training scripts | Create training script that supports:   * Using the generic data loader * Training of networks in JVET-AA0088 and JVET-AA0111 |  |
| Documentation | Create documentation describing the use of the software. This should include. Description of the command line options, configuration file options, data dumper format, and API. |  |
| Release of common software (including training scripts) rc1 | Release the rc1 version of the software and training scripts for people to study | Aug. 12 2022 (3 weeks after meeting)\* |
| Tag common software and training scripts version 1 | Address potential issues such as bug fix in the rc1 versions of common software and training scripts, release version 1 of common software and training scripts | Aug. 26 2022 (5 weeks after meeting)\* |

*\*Note: The timeline is only advisory. The common software development process will prioritize software quality and completeness over meeting the draft timeline.*

Notes from JVET plenary:

It was commented that “option 3” should be considered carefully in case of crosschecks. Might initialization of a model from the proposal lead to overfitting?

It was clarified that option 3 should only be used for the training crosscheck that would verify a modification of a model that existed before, and was previously verified with option 1 or option 2.

It was clarified that also in case of option 2 the proponent will provide the crosschecker with the software that allows extraction of the data, and that the crosschecker should verify this piece of software at least with a part of the data.

A precise description of these procedures to be included in the CTC document. Review an initial version on Friday.

The recommendations of the BoG were confirmed by the JVET plenary.

It was discussed whether a separate AHG for the software should be established. It was agreed to better perform the SW development in the context of the existing AHG for the current cycle.

The possible name of the software package was discussed, but it was agreed deciding that later.

It was suggested that the AHG should also be given a mandate to discuss the structure of repository(ies). It was also reported that it is planned to develop software guidelines (input doc for next meeting ideally).

### EE1 contributions: Neural network-based video coding (9)

Contributions in this area were discussed in session 8 at 1630–1730 UTC on Thursday 14 July 2022, and in session 12 at 1520–1735 UTC on Friday 15 July 2022 (chaired by JRO).

[JVET-AA0059](https://jvet-experts.org/doc_end_user/current_document.php?id=11735) EE1-3.1: Supplementary experiments based on JVET-Z0077 [Q. Qin, C. Jung (Xidian Univ.), D. Zou, M. Li (OPPO)]

This contribution reports an extension of the deep omnidirectional video compression framework (DOVC) to regular 2D videos. Specifically, we provide supplementary experiments based on JVET-Z0077. For I frames, we use VTM intra coding instead of BPG for image compression. For B frames, we generate five DOVC models for both luma and chroma channels. Compared with the VTM-11.0 NNVC-1.0 anchor, DOVC achieves average BD-rate reductions of {3.64% (Y), 4.53% (U) and 3.41% (V)} and average BD-PSNR gains of {-0.1495dB (Y), -0.2141dB (U), and -0.1411dB (V)} on the all 2D video dataset under RA configuration. Furthermore, DOVC takes advantage of GPU parallel processing and thus the average encoding time of DOVC is only 0.025 times that of the VTM-11.0 NNVC-1.0 anchor.

Better performance than anchor at lower bit rates.

Was GOP=16 used in anchor as said in the slides? Regularly, GOP=32 is used. Also the results tables are consistent with other anchor results.

Quality of keyframes and of the NN coded frames is aligned via a lambda parameter.

It was suggested that in the next round of EE, it should be tried to use the same intra period (64) as in the anchor, and code the keyframes between as B pictures

It was also suggested to try using an end-to-end still image coding method for the keyframes

Further study in EE.

[JVET-AA0066](https://jvet-experts.org/doc_end_user/current_document.php?id=11742) EE1-1.7: Content-adaptive post-filter based on SADL inference [R. Yang, M. Santamaria, F. Cricri, H. Zhang, J. Lainema, R. G. Youvalari, M. M. Hannuksela (Nokia)]

This contribution reports the results of tests EE1-1.7.1 and EE1-1.7.2, related to the content-adaptive CNN post-processing filter presented in JVET-Z0082. In these tests, the inference was done using SADL library. In EE1-1.7.1, the models used float32 precision. In EE1-1.7.2, the models are quantised to int16 and int32 precision. The evaluation was done against VTM 11.0 NNVC 1.0 in the RA configuration. It is reported that, on average, for EE1-1.7.1 the coding gains are 5.01% (Y), 18.95% (Cb) and 17.33% (Cr). In addition, for EE1-1.7.2 int16 the coding gains are 4.55% (Y), 16.95 (Cb) and 16.11 (Cr); for EE1-1.7.2 int32 the coding gains are 4.99% (Y), 18.90% (Cb) and 17.28 (Cr). It is further reported that float32 based inference with SADL is slower than with TensorFlow (C++). Finally, it is reported that both int32 and int16 based inferences with SADL are slower than float32 based inference with SADL.

Quantization of parameters is performed post training. Quantizer step size is applied layer-wise, depending on the maximum range of values in that layer (input and weights).

It was pointed out that using max value is very conservative. It may be better to allow overflow, and rather use smaller quantization step size. One expert mentions that percentages of overflow of up to 5% can be tolerable. Clipping in certain cases could even improve performance.

Cross-check of training to be investigated in EE.

[JVET-AA0220](https://jvet-experts.org/doc_end_user/current_document.php?id=11910) Crosscheck of JVET-AA0066 (EE1-1.7: Content-adaptive post-filter based on SADL inference) [T. Shao (Dolby)] [late]

The cross-checker mentions that one difficulty was that the current implementation does not allow segment-wise computation of NNR bits (which would be necessary for the segment-wise parallel simulations).

[JVET-AA0071](https://jvet-experts.org/doc_end_user/current_document.php?id=11747) EE1-2.1: A CNN-based Super Resolution Method Combined with GOP Level Adaptive Resolution [S. Peng, C. Fang, D. Jiang, J. Lin, X. Zhang (Dahua), J. Nam, S. Yoo, J. Lim, S. Kim (LGE)]

This contribution reports the EE1-2.1 test results, which is a combination of JVET-Z0088 and JVET-Z0065 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 in the latter case. Compared with VTM-11.0-NNVC, the experimental results show 3.06% and 3.27% BD-rate gains on average (A1 and A2) for luma, under AI and RA configurations, respectively.

Why only for luma? Extension to chroma is still under investigation.

Encoding and decoding time reported in the contribution are not reliable.

No plans so far on cross-checking the training.

[JVET-AA0081](https://jvet-experts.org/doc_end_user/current_document.php?id=11757) EE1-1.2: NN intra model without attention, partitioning and boundary strength [J. Ström, D. Liu, M. Damghanian, K. Andersson, Y. Li, P. Wennersten, R. Yu (Ericsson)]

This contribution presents additional results to the ablation study in JVET-Z0106. In that study, the intra model from JVET-Y0143 was retrained several times with some aspect of the neural network changed. In this contribution, further combinations are tested. Test EE1-1.2.1 involves removing the partitioning input as well as the attention branch from the intra model. Test EE1-1.2.2 involves removing both the partitioning and boundary strength inputs as well as the attention branch from the intra model. The reported BD-rate results over the VTM-11.0 + newMCTF anchor are:

EE1-1.2.1: -7.56% (Y), -18.48% (U) -19.09% (V) for AI, worst case complexity 635 kMAC/s.

EE1-1.2.2: -7.48% (Y), -18.64% (U) -19.17% (V) for AI, worst case complexity 625 kMAC/s.

It is proposed to adopt one of these two models as the intra luma model.

The original proponent of the method comments that the partitioning input is important according to their own investigation. This might however relate to the number of training epochs.

Is there a corresponding approach for inter? Yes, EE-related JVET-AA0090

Cross-check of training was suggested, also for better understanding the influence on the importance of partitioning input.

Cross-check of training to be investigated in EE.

[JVET-AA0172](https://jvet-experts.org/doc_end_user/current_document.php?id=11860) Cross-check of JVET-AA0081 (EE1-1.2: NN intra model without attention, partitioning and boundary strength) [M. Santamaria, F. Cricri (Nokia)] [late]

Small deviation was found. Only inference was tested.

[JVET-AA0085](https://jvet-experts.org/doc_end_user/current_document.php?id=11761) EE1-1.1: The Performance of Single-Model Filter Trained on the VTM and ECM Reconstruction [R. Chang, L. Wang, X. Xu, S. Liu (Tencent)]

(include abstract)

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

[JVET-AA0179](https://jvet-experts.org/doc_end_user/current_document.php?id=11867) Crosscheck of JVET-AA0085 (EE1-1.1: The Performance of Single-Model Filter Trained on the VTM and ECM Reconstruction) [C. Zhou (vivo)] late]

Cross-check was successful, exact match in AI (only inference).

[JVET-AA0223](https://jvet-experts.org/doc_end_user/current_document.php?id=11913) Cross-check of JVET-AA0085 (EE1-1.1: The Performance of Single-Model Filter Trained on the VTM and ECM Reconstruction) [M. Santamaria, F. Cricri (Nokia)] [late]

Cross-check was successful, very minor deviations (only inference).

[JVET-AA0087](https://jvet-experts.org/doc_end_user/current_document.php?id=11763) EE1-1.4: Neural network based in-loop filter with 2 models [L. Wang, S. Lin, X. Xu, S. Liu (Tencent), F. Galpin (InterDigital)]

This contribution reports the EE results of JVET-Z0094 [[[1]](#endnote-2)]. Compared with the previous method, the models are fine-tuned and the SADL library is updated. Based on the EE1 anchor, the related results are shown in order of RA and AI configurations as follows.

Test 1.4.1 (flt32, Libtorch implementation):

{-9.15%, -19.13%, -19.64%}, {-7.41%, -16.80%, -17.33%}

Test 1.4.2 (int16, SADL implementation):

{-9.17%, -19.92%, -19.80%}, {-7.46%, -16.63%, -17.57%}

How were quantizers for the integer version determined? Identical quantizers used for weights, inputs and outputs. Runtime is speeded up in the integer version.

It was commented that the usual template (without YUV PSNR) should be used.

[JVET-AA0216](https://jvet-experts.org/doc_end_user/current_document.php?id=11906) Crosscheck of JVET-AA0087(EE1-1.4: Neural network based in-loop filter with 2 models) [Z. Xie (OPPO)] [late]

Exact match after the integer implementation (only inference)

[JVET-AA0088](https://jvet-experts.org/doc_end_user/current_document.php?id=11764) EE1-1.5: Neural network based in-loop filter with a single model [L. Wang, S. Lin, X. Xu, S. Liu (Tencent), F. Galpin (InterDigital)]

This contribution reports the EE results of JVET-Z0091. Compared with the previous method, the models are fine-tuned and the SADL library is updated. Based on the EE1 anchor, the related results are shown in order of RA and AI configurations as follows.

Test 1.5.1 (flt32, Libtorch implementation):

{-8.67%, -18.63%, -18.97%}, {-6.50%, -14.88%, -15.97%}

Test 1.5.2 (flt32, SADL implementation):

{-8.84%, -18.86%, -18.99%}, {-6.50%, -15.35%, -16.54%}

Test 1.5.3 (int16, SADL implementation):

{-8.95%, -18.52%, -19.21%}, {-6.52%, -15.07%, -16.40%}

Static quantization was used. It was not checked for overflows (SADL by default performs clipping)

1.5.2 used one more training stage compared to 1.5.1, which improved the performance

Training time reported may not be accurate

It was suggested that some approach should be developed to measure training time (or training effort in general) more accurately across different proposals. Would be important to use very similar parameters in the crosschecks of training.

Combination with JVET-AA0122 to be considered.

[JVET-AA0178](https://jvet-experts.org/doc_end_user/current_document.php?id=11866) Crosscheck of JVET-AA0088 (EE1-1.5: Neural network based in-loop filter with a single model) [H. Wang (Qualcomm)] [late]

[JVET-AA0217](https://jvet-experts.org/doc_end_user/current_document.php?id=11907) Crosscheck of JVET-AA0088(EE1-1.5: neural network based in-loop filter with a single model) [Z. Xie (OPPO)] [late]

Re-training crosscheck gave similar gains (0.04% difference for luma, however but up to 1% for chroma) More investigation necessary why in the re-training crosschecks (not only for this proposal) chroma deviations are larger. One reason might be the chroma BD computation relative to total rate, but that may not explain such large deviation. It was commented by the proponents that the current training procedure gives more weight to luma in the loss function (10:1:1), which could explain such behaviour.

It is further reported that the results reported here (as well as in JVET-AA0178) are based on pseudo training, initializing the second stage (which uses data from decoded outputs of the NN codec) from the results of the proponents. The full training from scratch (including the first stage which starts from VVC decoded data, and then codes with the NN codec) is not finished yet. After results became available, the BoG JVET-AA0247 and analysed how close pseudo training (which is much faster) came to the real training. Also training times were analysed (see BoG report).

Initial results were shown in session 12 which indicated good match both between pseudo training and the real training crosscheck for RA and AI, but some larger deviation for LB (where the results were still incomplete). More analysis to be done after completion.

[JVET-AA0174](https://jvet-experts.org/doc_end_user/current_document.php?id=11862) [EE1] Crosscheck of training stage for EE1-1.5 and EE1-1.6 tests [J. Sauer, B. Wang, E. Alshina (Huawei)] [late]

This document reports difficulties encountered during the attempted crosschecks of EE1-1.5 and EE1-1.6 and suggests to align the training methods for both experiments. Training data preparation is the most time consuming part, it takes thousands of hours. Without unification at least naming convention, number of coded frames, QPs, training cross-check efforts are double.

Suggestions for faster crosschecks

It is expected that a successful crosscheck will produce a model which is close in performance to the one of the proponent. If we can assume that this implies the performance of the generated model for a given epoch during the training process should also be similar for different training runs we can speed-up the crosscheck procedure as follows:

1. The proponent runs his training. Proponent should keep checkpoints for each epoch and the log files for training and validation error
2. The crosschecker requests checkpoints for certain epochs, decided (possibly randomly) by the crosschecker  
   Question/Note: First few epochs might have to be excluded. After how many epochs can similar model performance be expected?
3. Proponent provides the requested checkpoints
4. The crosschecker continues/resumes training from the provided checkpoints for some small number of epochs. Note that if resources are available, training resuming from those checkpoints can be done in parallel. It is expected that training and validation loss should be close to the one the proponent observed.
5. Performance of final model should still be checked in VTM/ECM.

BoG (E. Alshina, A. Segall) on test and training conditions, in particular summarize the takeaways from the training crosschecks and develop/refine recommendations for executing and reporting. Based on that, determine if common software base methods can be defined.

[JVET-AA0111](https://jvet-experts.org/doc_end_user/current_document.php?id=11787) EE1-1.6: Deep In-Loop Filter With Fixed Point Implementation [Y. Li, K. Zhang, J. Li, L. Zhang (Bytedance), H. Wang, M. Coban, A.M. Kotra, M. Karczewicz (Qualcomm), F. Galpin (InterDigital), K. Andersson, J. Ström, D. Liu, R. Sjöberg (Ericsson)]

This contribution presents the EE test results of JVET-Z0113. In EE1-1.6.1, design and performance are identical to JVET-Y-EE1-1.7. The models are the re-trained ones from the cross-checker. In EE1-1.6.2 and EE1-1.6.3, convolutional neural network-based in-loop filters are implemented with SADL using floating point-based and fixed point-based calculations, respectively. BD-rate changes of {Y, Cb, Cr} on top of NNVC anchor are reportedly summarized as below:

EE1-1.6.1:

RA: {-9.79%, -22.33%, -22.82%}, AI: {-7.34%, -20.39%, -20.98%}

EE1-1.6.2 (floating point, SADL):

RA: {-10.19%, -22.10%, -21.82%}, AI: {-7.25%, -20.15%, -20.52%}

EE1-1.6.3 (fixed point, SADL):

RA: {-10.20%, -22.11%, -21.71%}, AI: {-7.26%, -20.14%, -20.56%}

Was encoder optimization enabled by macro? Yes, gives approx. 0.4% gain.

[JVET-AA0181](https://jvet-experts.org/doc_end_user/current_document.php?id=11869) Cross-check of JVET-AA0111 (EE1-1.6: Deep In-Loop Filter With Fixed Point Implementation) [K. Lin, C. Jia, S. Wang (PKU)] late]

Training crosscheck gives deviations 0.4% for RA, 0.1% for AI (luma), slightly higher for chroma. Perfect match for inference in integer. Training cross-check was done from scratch. No results for LB.

[JVET-AA0122](https://jvet-experts.org/doc_end_user/current_document.php?id=11798) EE1-1.3: On BaseQP adjustment in CNNLF [Z. Xie, Y. Yu, H. Yu, D. Wang (OPPO)]

This contribution reports the EE results of JVET-Z0128. A QP-adjustment method was proposed to improve the coding performance of JVET-Z0091 and JVET-Z0094. In those two contributions, the QP parameters were part of the input to the network at the inference stage and these QP parameters include BaseQP and SliceQP. JVET-Z0128 proposed to add an offset to the BaseQP at the slice-level and use the final adjusted BaseQP as the input of the network. On top of those two proposals, it is reported that the proposed method can further improve the coding efficiency, {Y, U, V, EncT, DecT}, shown as follows:

EE1-1.3.1:

RA: Y -0.31%, U -0.14%, V -0.14%, EncT 105%, DecT 100%;

LDB: Y -0.43%, U -0.38%, V -0.21%, EncT 100%, DecT 96%.

EE1-1.3.2:

RA: Y -0.32%, U -0.60%, V -0.43%, EncT 105%, DecT 101%;

LDB: Y -0.44%, U -0.15%, V 0.55%, EncT 107%, DecT 111%

JVET-Z0091 is equivalent with JVET-AA0088 of the current meeting. Seems straightforward to combine (to be integrated in software).

[JVET-AA0185](https://jvet-experts.org/doc_end_user/current_document.php?id=11873) Crosscheck of JVET-AA0122 (EE1-1.3: On BaseQP adjustment in CNNLF) [L. Wang (Tencent)] [late]

### EE1 related contributions: Neural network-based video coding (8)

Contributions in this area were discussed in session 13 at 0750–0930 UTC on Monday 18 July 2022, and in session 16 at 0720–0900 UTC on Tuesday 19 July 2022 (chaired by JRO).

[JVET-AA0074](https://jvet-experts.org/doc_end_user/current_document.php?id=11750) [EE1-related] Lighter WCDANN: CNN Based In-Loop Filters [H. Zhang, C. Jung (Xidian Univ.), D. Zou, M. Li (OPPO)]

This contribution presents convolutional neural network-based in-loop filters. In this contribution, we modify the network structure and training strategy of WCDANN proposed by JVET-Y0052 [1], make WCDANN lightweight, and provide two filters based on the QP value, namely Filter1 and Filter2. Compared with VTM-11.0\_NNVC-1.0, the proposed Filter1 achieves average {4.36%, 12.59%, 11.95%} BD-rate reductions for {Y, Cb, Cr} under AI configuration, and the proposed Filter2 achieves average {4.58%, 12.67%, 13.21%} BD-rate reductions for {Y, Cb, Cr} under AI configuration.

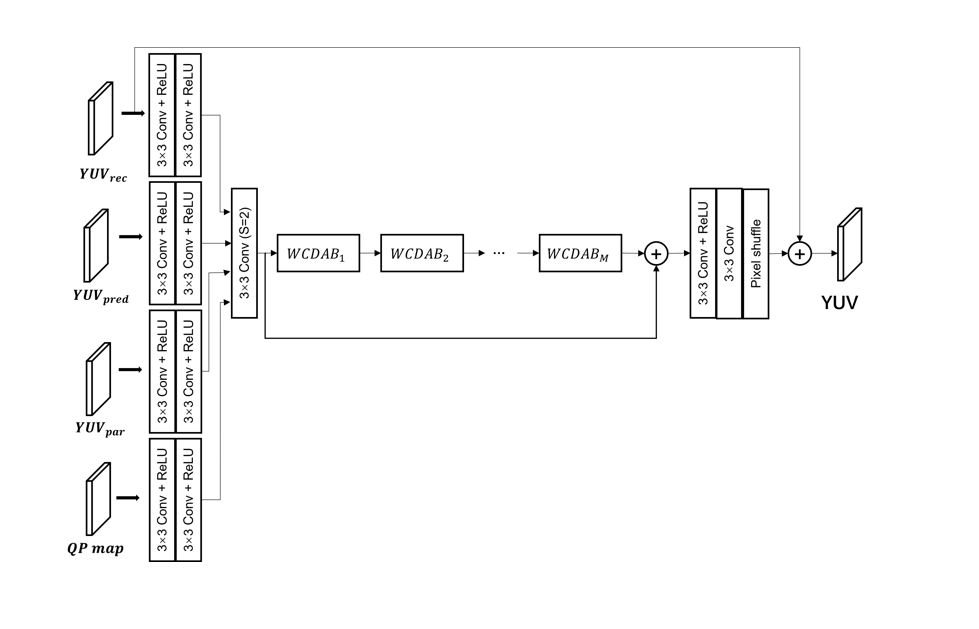


Fig. 1 Network architecture of the proposed Lighter WCDANN. The parameter M is set to 4.

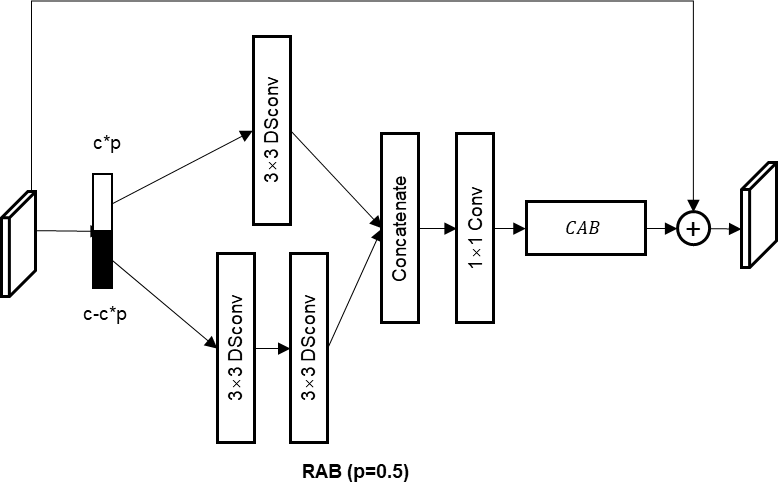


Fig. 2 Network architecture of modified RAB. The parameter p is set to 0.5.

Filter 1 one network for all QPs, filter 2 five different networks.

What is the complexity of the filter? How much was complexity reduced compared to JVET-Y0052? From 650 KMAC/pix to 434.

Subjective performance of both filters also similar, according to proponents’ opinion.

Not competitive compared to methods currently in EE, for example JVET-AA0088 has roughly kMAC/pixel using 1 model, but gain around 6.5% (for AI)

Results for other configurations than AI? Not yet.

Further study recommended to improve the complexity/performance tradeoff. For example, other weights (higher weight for luminance) could be used in the loss function. Investigate this aspect in EE.

[JVET-AA0089](https://jvet-experts.org/doc_end_user/current_document.php?id=11765) EE1-related: More refinements on EE1-1.4 and EE1-1.5 [L. Wang, S. Lin, X. Xu, S. Liu (Tencent), Z. Xie, Y. Yu, H. Yu, D. Wang (OPPO)]

This contribution presents more refinements, including updating the models, revising the encoder configuration, as well as the improvement used in EE1-1.3, for EE1-1.4 and EE1-1.5 tests. In addition, the model with the int16 precision is also included in the tests. Based on the EE1 anchor, the related results are shown in order of RA and AI configurations as follows.

Filter 1 (flt32, two models, new models):

{-9.30%, -19.42%, -20.02%}, {-7.41%, -16.80%, -17.33%}

Filter 1 (flt32, two models, new models, EncDbOpt):

{-9.79%, -20.43%, -20.97%}, {-7.36%, -17.34%, -18.00%}

Filter 1 (int16, two models, new models, EncDbOpt):

{-9.83%, -21.30%, -21.16%}, {-7.46%, -17.21%, -18.29%}

Filter 2 (flt32, one model, new model):

{-9.06%, -19.15%, -19.38%}, {-6.50%, -15.34%, -16.54%}

Filter 2 (flt32, one model, new model, EncDbOpt):

{-9.61%, -20.12%, -20.41%}, {-6.54%, -15.96%, -17.15%}

Filter 2 (int16, one model, new model, EncDbOpt):

{-9.64%, -19.89%, -20.60%}, {-6.48%, -15.62%, -16.87%}

Model refinements, and encoder optimization for deblocking. The latter is not enabled in the anchor. However, some other proposals are enabling it. For a fair comparison against the anchor, it should also be enabled for the anchor. It is noted that this might lead to less gain over VVC compared to the situation where it was disabled.

It was reported that ECM and VTM CTC also use this tool in the anchors.

Decision(CTC): Enable EncDbOpt in anchor, and allow proposals using this or equivalent as well.

See also notes under JVET-AA0088 and JVET-AA0122 which recommend the combination.

For the new models, only inference was cross-checked, not the training.

Training of the new models to be investigated in EE.

[JVET-AA0229](https://jvet-experts.org/doc_end_user/current_document.php?id=11919) Crosscheck of JVET-AA0089 (EE1-related: More refinements on EE1-1.4 and EE1-1.5) [D. Liu (Ericsson)] [late]

[JVET-AA0090](https://jvet-experts.org/doc_end_user/current_document.php?id=11766) EE1-related: One luma model with IPB and skip for filtering intra and inter luma slices [D. Liu, J. Ström, M. Damghanian, P. Wennersten, Y. Li (Ericsson)]

This contribution proposes to use one luma model with block type information and block skip information for NN-filtering of intra luma and inter luma slices. It is claimed that the combined model has a worst-case complexity in terms of kMAC/pixel that is 1.7% higher than JVET-Z0112, but that this model makes it possible to reduce the number of models from four to three, and to reduce the total number of parameters from 6.24M to 4.70M. The proposed model uses block type information IPB (intra/uni-predicted/bi-predicted) and block skip information (bypassed or not) as two additional inputs on top of the inter luma model from JVET-Z0112. It is reported that the IPB+skip model gives 0.35% (= 9.99% − 9.64%) higher luma gain for RA than the model without IPB or skip in the tests carried out in the contribution. It is further reported that the proposed IPB+skip model has a loss of 0.1% comparing to using separate intra luma and inter luma models. The IPB+skip model tested on top of JVET-AA0111 EE1-1.6.2 with a combination of deblocking filter and SADL gives a luma gain of 10.27% for RA.

* Test 1: Inter luma model from JVET-Z0112 retrained with only RA data and and with separate intra luma and inter luma models:

BDR-Y -10.09% RA, -7.39% AI, and -8.48% LDB. Worst case kMAC/pixel 649.

* Test 2: Inter luma model from JVET-Z0112 retrained with AI and RA data:

BDR-Y -9.64% RA, -7.48% AI, and -8.47% LDB. Worst case kMAC/pixel 638.

* Test 3: IPB+skip model trained with AI and RA data:

BDR-Y -9.99% RA, -7.44% AI, and -8.47% LDB. Worst case kMAC/pixel 660.

* Test 4: IPB+skip model tested on top of JVET-AA0111 (EE1-1.6.2, float32, combination with deblocking and SADL)

BDR-Y -10.27% RA, -7.32% AI, and xxx LDB. Worst case kMAC/pixel 660

Only inter models are retrained compared to JVET-Z0112. Intra models were retrained in EE (JVET-A0111).

Partitioning information not used (was also not used in Z0112).

In test 1, a larger number of epochs was used as the data set was smaller (such that training time was approximately same as in tests 2 and 3).

Attention block is unchanged relative to original proposals, except that IPB and skip information is added as input.

It was suggested that this could also be used with combined luma/chroma models.

Investigate in EE.

[JVET-AA0238](https://jvet-experts.org/doc_end_user/current_document.php?id=11928) Crosscheck of JVET-AA0090 (EE1-related: One luma model with IPB and skip for filtering intra and inter luma slices) [L. Wang (Tencent)] [late]

Results match.

[JVET-AA0094](https://jvet-experts.org/doc_end_user/current_document.php?id=11770) EE1-related: Deep In-Loop Filter in EE1-1.6 with Adaptive Input Samples [C. Zhou, Z. Lv, J. Zhang (vivo), W. Chen, J. Guo, B. Ai (BJTU)] [late]

Initial version rejected as “placeholder”

This contribution proposes a NN filter method with adaptive choice of the input samples on top of EE1-1.6. In EE1-1.6, the reconstruction samples before deblocking filtering is used as part of the inputs to the network. It is proposed to determine the type of reconstruction samples which will be processed by the NN filter based on the RD cost. In this contribution, a choice is made between the samples before deblocking filtering and the samples after deblocking filtering.

Due to the limited time, only the simulation results based on the model after training phase II were obtained so far. Compared to EE1-1.6 using the phase II trained model, the BD-rate figures are as follows:

RA: Y -0.08%, U 0.00 %, V 0.05 %;

LDB: Y -x.xx %, U -x.xx%, V -x.xx%.

Training uses samples both from before and after deblocking. It is assumed by the proponents that the output of the deblocking gives additional benefit

Would subjective improvement be visible in video?

Other experts found the idea interesting, but results are still preliminary. Further study recommended, would be premature to include in EE.

[JVET-AA0200](https://jvet-experts.org/doc_end_user/current_document.php?id=11889) Crosscheck of JVET-AA0094 (EE1-related: Deep In-Loop Filter in EE1-1.6 with Adaptive Input Samples) [Y. Li (Bytedance)] [late] [miss]

Floating point, reasonably close match.

[JVET-AA0112](https://jvet-experts.org/doc_end_user/current_document.php?id=11788) EE1-1.6-related: Deep In-Loop Filter with Additional Input Information [Y. Li, K. Zhang, L. Zhang (Bytedance)]

In EE1-1.6, a convolutional neural network-based in-loop filtering method is proposed and implemented with SADL using both floating point-based and fixed point-based calculations. In this contribution, two aspects are proposed to improve the CNN models presented in EE1-1.6 with additional input information:

Aspect #1: Forward/backward collocated blocks can be input to the in-loop filtering network.

Aspect #2: The input samples of CNN can be flipped and the output samples of CNN can be flipped back.

The proposed CNN model is implemented using SADL int16 precision. Compared with VTM-11.0 + NewMCTF, the proposed method reportedly show on average {10.67%, 22.05%, 21.67%} BD-rate reductions for {Y, Cb, Cr} under RA configuration, respectively.

Results only for RA. It was suggested that a similar approach would be applicable in LDB and LDP

Aspect 1 adds temporal filtering. Collocated blocks (not motion compensated) from other pictures have same size as current block. There is an on/off control.

kMAC/pixel roughly the same.

There is also some encoder optimization.

Aspect 2 (flipping) has less benefit than aspect 1. Separate results for the two aspects are not available.

It was commented that the memory size is increased by approx. 25%

Was training data the same as in original proposal? Set is the same, but the input of training is different in terms of selecting temporal layers and generating input from the collocated blocks.

Was visual quality inspected? Could the usage of collocated blocks cause blurring effects in case of motion?

It was commented that potentially motion vectors could be used as additional input.

Some interest was expressed from other experts in cross-checking the training.

Investigate in EE

[JVET-AA0214](https://jvet-experts.org/doc_end_user/current_document.php?id=11904) Crosscheck of JVET-AA0112 (EE1-1.6-related: Deep In-Loop Filter with Additional Input Information) [T. Shao (Dolby)] [late] [miss]

Cross-check only for inference. Class D has perfect match so far.

[JVET-AA0113](https://jvet-experts.org/doc_end_user/current_document.php?id=11789) EE1-1.6-related: RDO Considering Deep In-Loop Filter with SADL [J. Li, Y. Li, K. Zhang, L. Zhang (Bytedance)]

This contribution presents an encoder optimization technique on top of the deep learning-based in-loop filtering method combined with deblocking in EE1-1.6. The proposed method is asserted to improve the compression efficiency at encoder by introducing CNN-based filtering and deblocking into the rate-distortion optimization (RDO) process. In the RDO process, CNN-based filtering is implemented with SADL using fixed point-based calculation. Compared with VTM11.0\_nnvc, the proposed method reportedly shows on average {7.56%, 20.30%, 20.70%}, {10.54%, 22.05%, 21.67%}, and {x%, x%, x%} BD-rate reductions for {Y, Cb, Cr} components, under AI, RA, and LDB configurations, respectively.

Luma improvement over EE1-1.6 0.38% for RA, 0.32% for AI. Encoding time increase 8% for RA, 31% for AI. Less gain (or sometimes even loss) for chroma, probably due to the fact that RDO only considers luma. Only one model considered in RDO, model specifically trained for the encoder optimization.

Visual quality? Was not considered.

It was commented that an equivalent method could be implemented with any NN based loop filter.

Investigate in EE (also cross-check training).

[JVET-AA0218](https://jvet-experts.org/doc_end_user/current_document.php?id=11908) Crosscheck of JVET-AA0113 (EE1-1.6-related: RDO Considering Deep In-Loop Filter with SADL) [K. Lin, C. Jia, S. Wang (PKU)] [late]

Only inference. Partial results (class D) match was found.

[JVET-AA0115](https://jvet-experts.org/doc_end_user/current_document.php?id=11791) EE1-1.6-related: ALF with Samples before Deep In-Loop Filter [J. Li, K. Zhang, Y. Li, L. Zhang (Bytedance)]

This contribution presents a loop filter (ALF) technique on top of the deep learning-based in-loop filtering method in EE1-1.6. ALF is proposed to take luma samples before CNN-based filtering as the additional input where filters with 3x3 and 5x5 diamond shapes are utilized to derive the filtered samples. Compared with VTM11.0\_nnvc, the proposed method with 5x5 diamond shapes reportedly shows on average {x%, x%, x%}, {x%, x%, x%}, and {x%, x%, x%} BD-rate reductions for {Y, Cb, Cr} components, under AI, RA, and LDB configurations, respectively. With 3x3 diamond shapes, the proposed method reportedly bring on average {x%, x%, x%}, { x%, x%, x%}, and {x%, x%, x%} BD-rate reductions for {Y, Cb, Cr} components, under AI, RA, and LDB configurations, respectively.

The slide deck has complete results for AI and RA. The additional luma gain compared to EE1-1.6 is roughly 0.2% for RA, 0.1% for AI.

The additional ALF used to generate the input is different from VTM in terms of filter shape (only 3x3 and 5x5 are used), with input of samples from before the NN filtering. Classification part is identical to VTM’s ALF. The combination is done similarly as in the EE2 proposal JVET-Z0146, which feeds samples from before deblocking into ALF.

No new model was trained.The contribution rather targets improving ALF rather than improving NN technology, and the gain is relatively low. It is asserted that this is less relevant for the EE1 exploration, might be more interesting in combination with EE2.

[JVET-AA0219](https://jvet-experts.org/doc_end_user/current_document.php?id=11909) Crosscheck of JVET-AA0115 (EE1-1.6-related: ALF with Samples before Deep In-Loop Filter) [K. Lin, C. Jia, S. Wang (PKU)] [late]

Partial results (class D) match was found.

[JVET-AA0131](https://jvet-experts.org/doc_end_user/current_document.php?id=11807) EE1-related: CNN based in-loop filtering with large activation layer [H. Wang, S. Eadie, M. Coban, M. Karczewicz (Qualcomm)]

In this contribution, an alternative network structure which is derived based on EE1-1.5 and EE1-1.6 is studied. The network studied in EE1-1.6 is used as baseline and the network backbone is modified to use residue blocks with large activation layers. Experimental results reportedly show 11.25 %, 23.82 % and 25.22 % BD rate saving for RA and 7.78%, 20.87 %, 22.09 % BD rate saving for AI, for Y, Cb and Cr components, respectively.

4 models, 1.9 M parameters/model.

kMAC/pixel almost identical (slightly reduced) with EE1-1.6.1 (which also has four models, but less parameters 1.55 M parameters).

32 residual blocks are used, no attention blocks.

Different training procedure than in EE, but similar to EE1-1.6

Could the large number of res block increase encoding time?

Investigate in EE, quantization implementation in SADL, varying number of res blocks, and other complexity aspects.

[JVET-AA0186](https://jvet-experts.org/doc_end_user/current_document.php?id=11874) Crosscheck of JVET-AA0131 (EE1-related: CNN based in-loop filtering with large activation layer) [L. Wang (Tencent)] [late]

No exact match initially. 32bit FP is used, and there may be some difference depending on compiler version settings for SADL. Using identical settings gave matching partial results.

It was commented that a minor deviation dependent on compiler version settings was also observed in the anchor generation recently.

### Other NN technology related contributions (7)

Contributions in this area were discussed in session 21 at 1330–1500 UTC on Wednesday 20 July 2022 (chaired by JRO).

[JVET-AA0063](https://jvet-experts.org/doc_end_user/current_document.php?id=11739) AHG11: A hybrid codec using E2E image coding combined with VVC video coding [Y. He, B. Wang, E. Alshina, J. Sauer (Huawei)]

This informative contribution presents a hybrid codec, where I frames are coded using an E2E (End-2-End) image codec while all remaining frames (P and B frames) are coded with VVC. As the chosen end-to-end neural-network based image codec provides a better than VVC coding performance on I frame, the coding performance of P and B frames are also improved due to inter-prediction propagation. Intra coding inside P and B slices remains the same as VVC. Compared to VVC 14.0 under random access configuration, experiments reportedly show that the proposed hybrid codec provides 4.1% BD-rate gain on Y component in random access configuration (3.7%, 5.6%, 3.4% BD-rate gain for Class A1, Class A2, and Class B, respectively). It should be noted that gain is demonstrated in random access configuration, while only 1 out of 64 frames coding is modified (for most of the sequences Intra period is 64). In demonstrated codec design the switching between “classical” (CPU) and “end-to-end neural-network based” coding (both CPU and GPU coding is possible) happens at picture level.

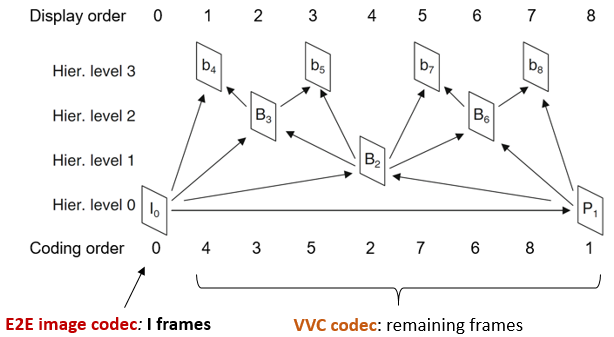


Figure 1: coding with a hybrid approach, the GOP setup might differ

Input is YUV 420. Luma is encoded directly, and also downsampled and put as additional input into chroma channel (both chroma components coded jointly. Autoencoder used at the core for generating the binary output.

The image codec is also submitted as proposal in the JPEG AI CfP. Also there, the YUV conversion was used. Original contribution is available from WG 1 doc repository.

For class F, significantly higher bit rate (50% increase) than for VVC.

Are there data for gain in AI CTC? No

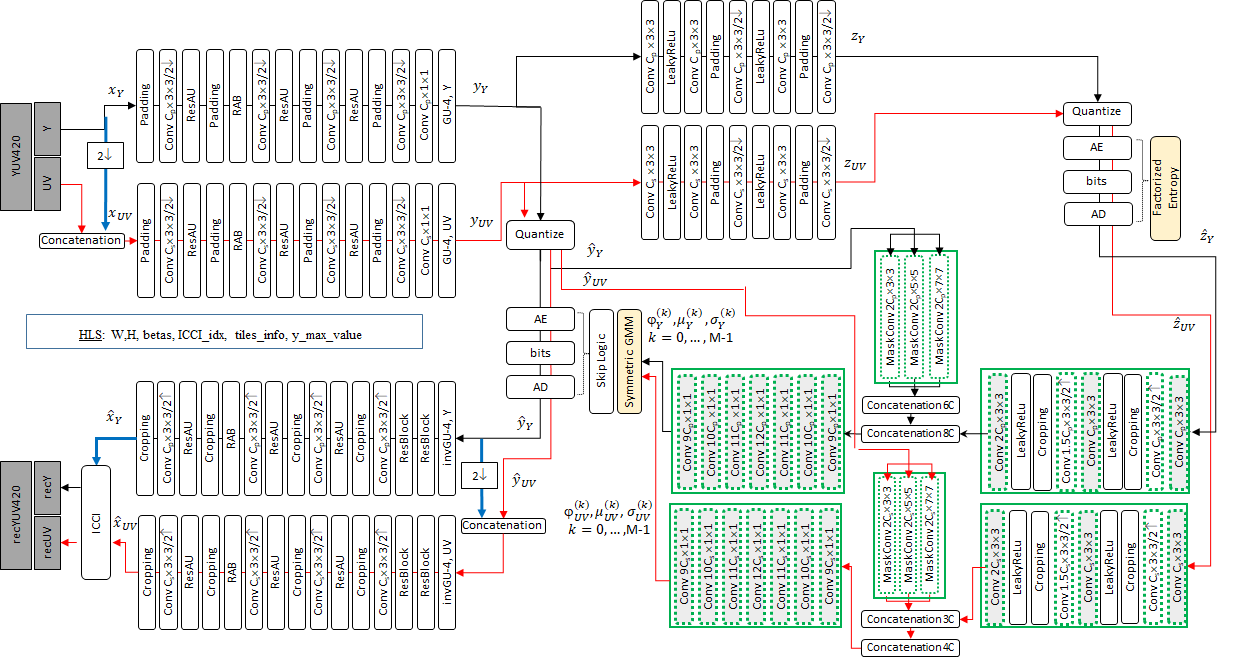


Fig. 2. End-2-end image codec architecture.

For the experimentation, separate bitstreams were used the E2E codec and VVC coded frames. In the latter, only HLS and block skip bits were written for I frames.

Decoding run time is approx. 34x of VVC I frame coding on CPU. Decoding complexity is approx. 700-900 kMAC/pixel.

[JVET-AA0065](https://jvet-experts.org/doc_end_user/current_document.php?id=11741) AHG11: CNN Filter for Super-Resolution with RPR functionality in VVC [S. Huang, C. Jung, Y. Liu, M. Li]

This contribution presents a super-resolution network that combines CNN with existing RPR functionality in VVC, called MMSDANet. In MMSDANet, we propose a new basic block, multi-mixed scale and depth information with attention block (MMSDAB) to extract multi-scale and convolutional layer depth information, and shared convolution is used to reduce network parameters. Compared with VTM-11.0 RPR anchor, MMSDANet achieves {-6.72%, -26.89%, -28.19%} and {-8.16%, -25.32%, -26.30%} BD-rate gains on average for {Y, Cb, Cr}, under RA and AI configurations, respectively. Compared with VTM-11.0 NNVC-1.0 anchor, MMSDANet achieves {-4.21%, 4.53%, -9.55%} and {-8.5%, 18.78%, -12.61%} BD-rate gains on average for {Y, Cb, Cr}, under RA and AI configurations, respectively.

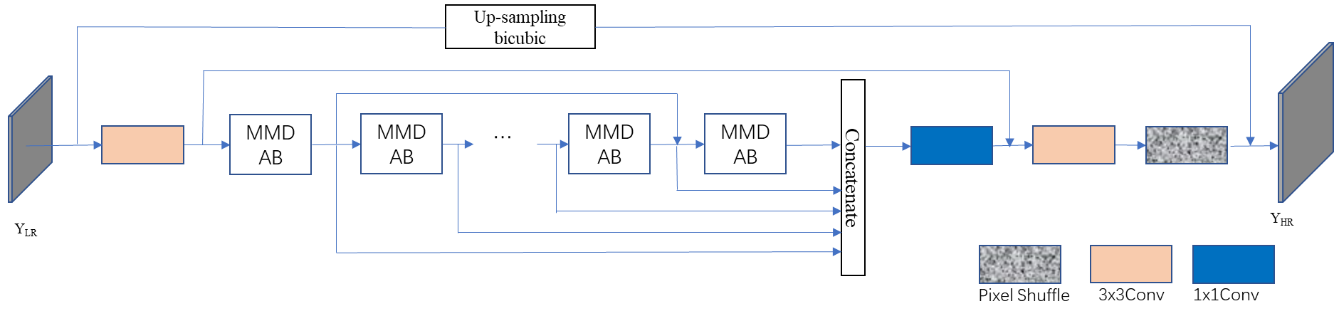


Illustration of the proposed MMSDANet for luma channel.

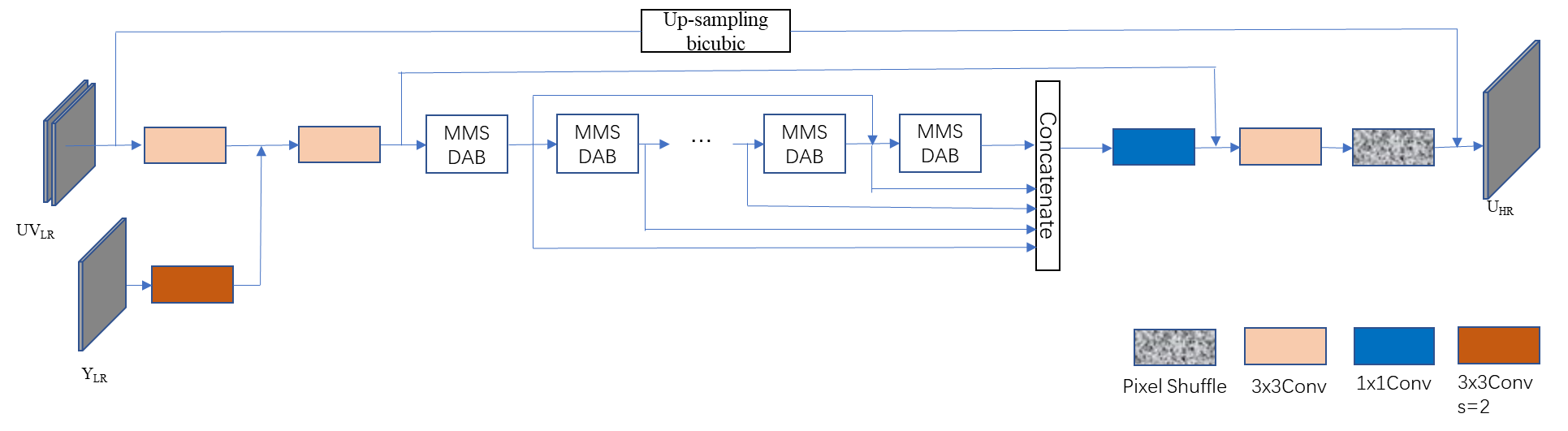


Illustration of the proposed MMSDANet for chroma channels.

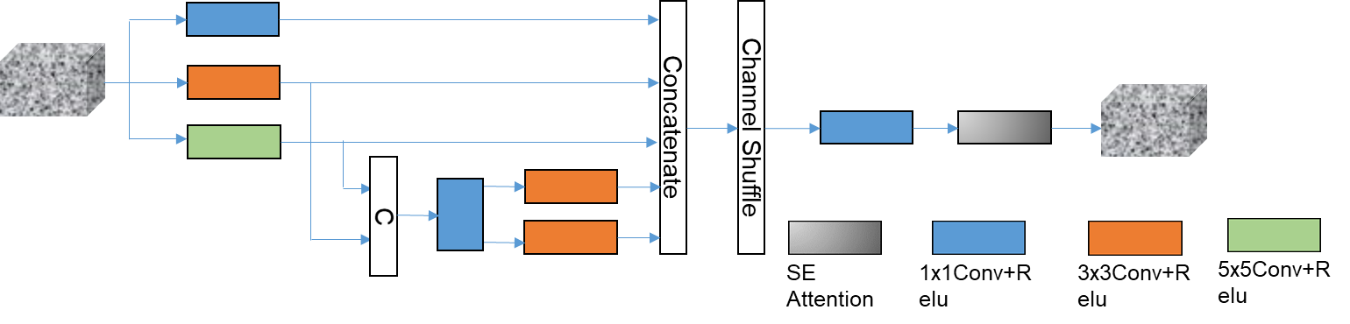


Illustration of the proposed MMSDAB.

|  |  |  |
| --- | --- | --- |
| **Network Information in Inference Stage** | | |
| Mandatory | HW environment: | Intel Core i9 12900k @3.9GHz |
| Framework: | LibTorch v1.9 |
| Number of GPUs per Task | 0 |
| Number of Parameters (Each Model) | luma up-sampling model: 2.0M/model  chroma up-sampling model: 2.1M/model |
| Total Parameter Number | 31M |
| Parameter Precision (Bits) | 32 (F) |
| Memory Parameter (MB) | 15 models in total: 118 MB |
| Multiply Accumulate (MAC) | 2044kMAC/pixel |
| Optional | Total Conv. Layers | 75 for up-sampling the luma, 77 for up-sampling the chroma |
| Total FC Layers | 0 |
| Batch size: | 1 |
| Patch size | Whole frame |

Loss on some sequences. The coding is always performed at low resolution. It is commented that with an adaptation as used in the EE there should be no loss.

The method gives gain also for some other sequences than the previous EE proposals.

It is commented that the complexity is extremely high. It should be investigated if the high number of convolutional layers and the high number of models is necessary, or if the gain would be retained when reducing complexity.

Investigate in EE.

[JVET-AA0076](https://jvet-experts.org/doc_end_user/current_document.php?id=11752) AHG11: RPR-Based Super-Resolution Guided by Partition Information [Q. Han, C. Jung (Xidian Univ.), Y. Liu, M. Li (OPPO)]

In this contribution, we propose a CNN filter for RPR-based super-resolution guided by partition information. Since the blocking artifacts are closely related to block partitioning, the coding tree unit (CTU) partition map is helpful for predicting the blocking artifacts. Thus, we propose a reference spatial attention block (RSAB) for blocking artifact removal based on the CTU partition map. To make full use of correlation between luma and chroma, we use a U-Net backbone that extracts and fuses multi-scales features from an image. Through the U-Net backbone, we design a dilated convolution-based dense block with channel attention. The proposed CNN filter achieves {-9.25% (Y), 8.82% (Cb), -16.39% (Cr)}, {-4.67% (Y), -1.75% (Cb), -11.70% (Cr)} and {-7.75% (Y), -3.89% (Cb), -13.44% (Cr)} BD-rate changes over official AHG11 anchor (VTM-11.0 + new MCTF) under all intra (AI), random access (RA) and low delay P (LDP) configurations, respectively.

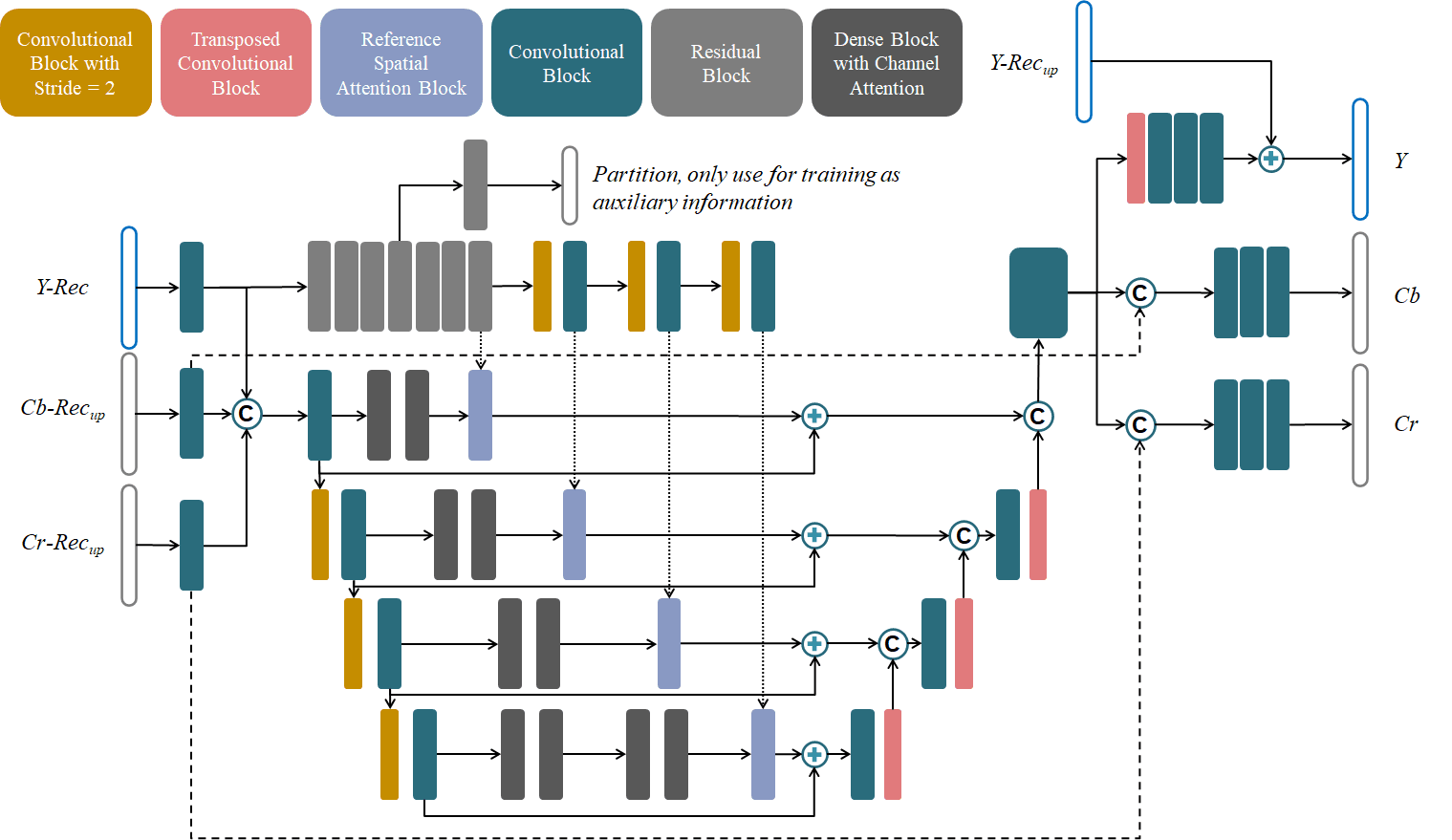


Illustration of the proposed network architecture for RPR-based SR guided by partition information

|  |  |  |
| --- | --- | --- |
| **Network Information in Inference Stage** | | |
| Mandatory | HW environment: | Intel(R) Core i9-12900KF CPU @ 2.40GHz |
| Framework: | LibTorch v1.7.1 |
| Number of GPUs per Task | 0 |
| Number of Parameters (Each Model) | 2,493,351/model |
| Total Parameter Number | 12,466,755 |
| Parameter Precision (Bits) | 32 (F) |
| Memory Parameter (MB) | 47.55689621 |
| Multiply Accumulate (MAC) |  |
| Optional | Total Conv. Layers | 85 |
| Total FC Layers | 0 |
| Batch size: | 1 |
| Patch size | Whole frame |

kMAC/pixel is about 820.

The method is using attention blocks on low resolution input. It was asked why partitioning information is relevant for superresolution. Could it be that the network implicitly performs some kind of enhancement of the reduced resolution input.

It is noted that this proposal also has benefit on some sequences where the proposal JVET-AA0065 did not perform as good. Would be interesting to investigate the influence of the attention mechanisms.

Has it been tried to use the quantization information (QP) from low resolution layer as input? This could be interesting to reduce the number of models (currently, different models for different QP)

It would also be interesting to investigate the reduction of the number of convolutional layers.

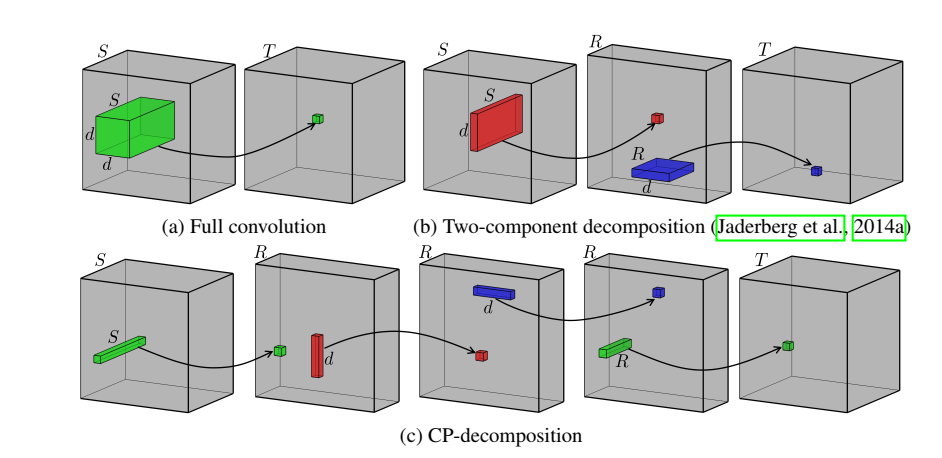
Investigate in EE.

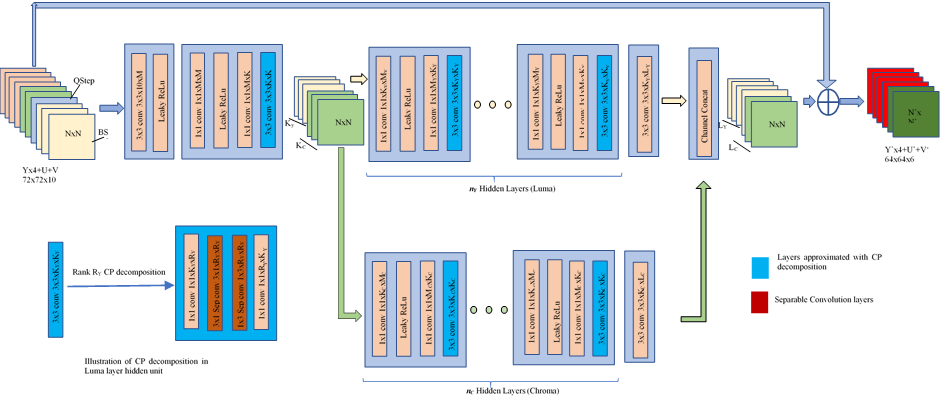
[JVET-AA0080](https://jvet-experts.org/doc_end_user/current_document.php?id=11756) AHG11: Complexity reduction on neural-network loop filter [J. N. Shingala, S. Kadaramandalgi, A. Shyam (Ittiam), T. Shao, A. Arora, P. Yin, F. Pu, T. Lu, S. McCarthy (Dolby)]

This contribution proposes to use CP decomposition to reduce the inference computational complexity of neural-network-based loop filtering. This contribution also proposes to fuse adjacent 1x1 convolutions to achieve additional reduction in computational complexity. The baseline model is taken from JVET-X0140 low complexity model (EE1-1.4.1), which shows BD-rate saving of 5.16%, 5.93%, 5.19% (Y, Cb, Cr, respectively) for AI compared to EE1 VTM anchor. Using block-level KMAC/Pixel as the complexity metric, the baseline model has worst case complexity of 33.6 KMAC/Pixel. The proposed CP decomposition with rank 24 is about 20.1 KMAC/Pixel. The complexity is further reduced to 16.2 KMAC/Pixel by fusing adjacent 1x1 convolutions, i.e., approximately half the complexity of the baseline model. Compared to the baseline model (JVET-X0140 EE1-1.4.1), the simulation shows that CP decomposition with rank 24 plus fusing adjacent 1x1 convolutions has the BD-Rate loss of 0.50%, 0.21%, 0.39% (Y, Cb, Cr, respectively) for AI. Compared to EE1 VTM anchor, the BD-rate saving is 4.68%, 5.72%, 4.81% (Y, Cb, Cr, respectively) for AI.

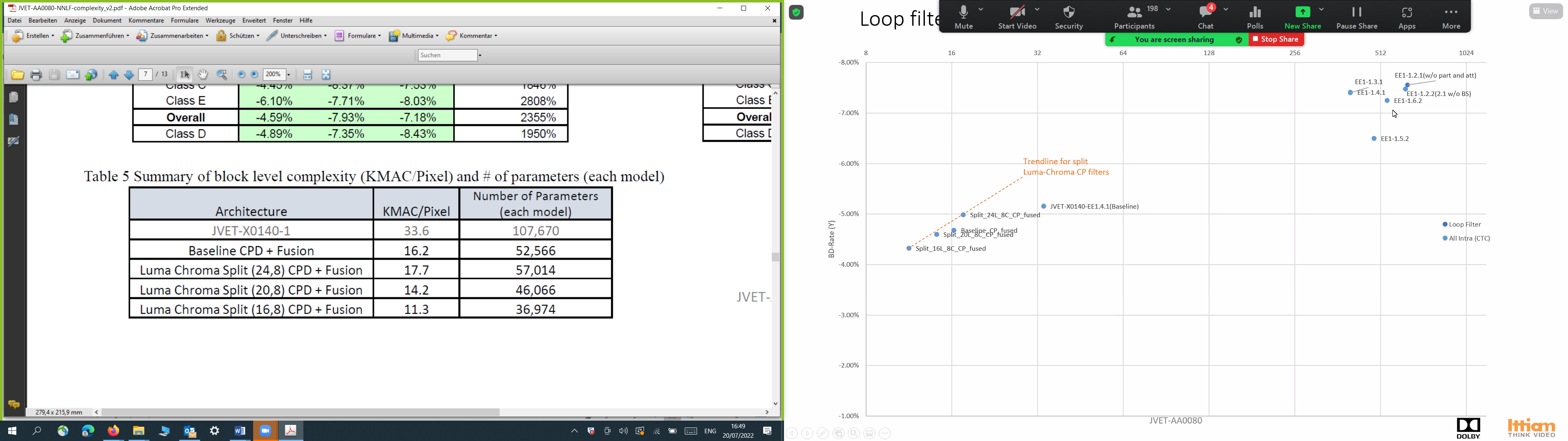
In v2, one additional method is added to further reduce the complexity of NNLF model, which retains the majority of the luma gains and achieves better coding efficiency for chroma. The method is to split architecture for luma and chroma components between the input network and output network, therefore different number of features can be used for luma and chroma components, indicated by (xL, yC). Combined with CP decomposition and fusion adjacent 1x1 convolutions, when compared against EE1 VTM anchor, the BD-Rate saving for AI (Y, Cb Cr, respectively) is:

* (24L, 8C): 4.99%, 7.92%, 7.59%, with worst case block level complexity of 17.7 KMAC/Pixel
* (20L, 8C): 4.59%, 7.93%, 7.18%, with worst case block level complexity of 14.2 KMAC/Pixel
* (16L, 8C): 4.32%, 8.28%, 7.66%, with worst case block level complexity of 11.3 KMAC/Pixel





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



Decoding time (CPU) is 20-27x that of VTM (for AI), depending on the different models.

CP decomposition uses more layers, but reduces kMAC/pixel.

It was suggested that potentially the number of channels in chroma might be further reduced after the luma chroma split, without experiencing too much loss.

Several experts expressed it would be interesting for EE, but Proponents would prefer to first also investigate RA.

Further study recommended.

[JVET-AA0082](https://jvet-experts.org/doc_end_user/current_document.php?id=11758) AHG11: Deep Reference Frame Generation for Inter Prediction Enhancement [Z. Liu, X. Xu, S. Liu (Tencent), J. Jia, Z. Chen (Wuhan Univ.)]

This contribution presents a deep-neural-network-based reference frame generation (DRF) method for VVC inter prediction enhancement. During the encoding and decoding process, the deep reference frame generation 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 for the current encoding picture as an alternative reference to enhance the inter prediction. Different from the previous work in which only the frames with bi-directional reference are processed, this contribution extends the reference frame generation method to uni-directional prediction cases and apply it to LD configuration. It is reported that this method can bring {LDB: 3.41%/13.59%/14.17%, RA: 2.03%/7.02%/6.45%} bitrate savings for Y/U/V components respectively, following the JVET-NNVC CTC.

Was presented in session 23 0550 UTC

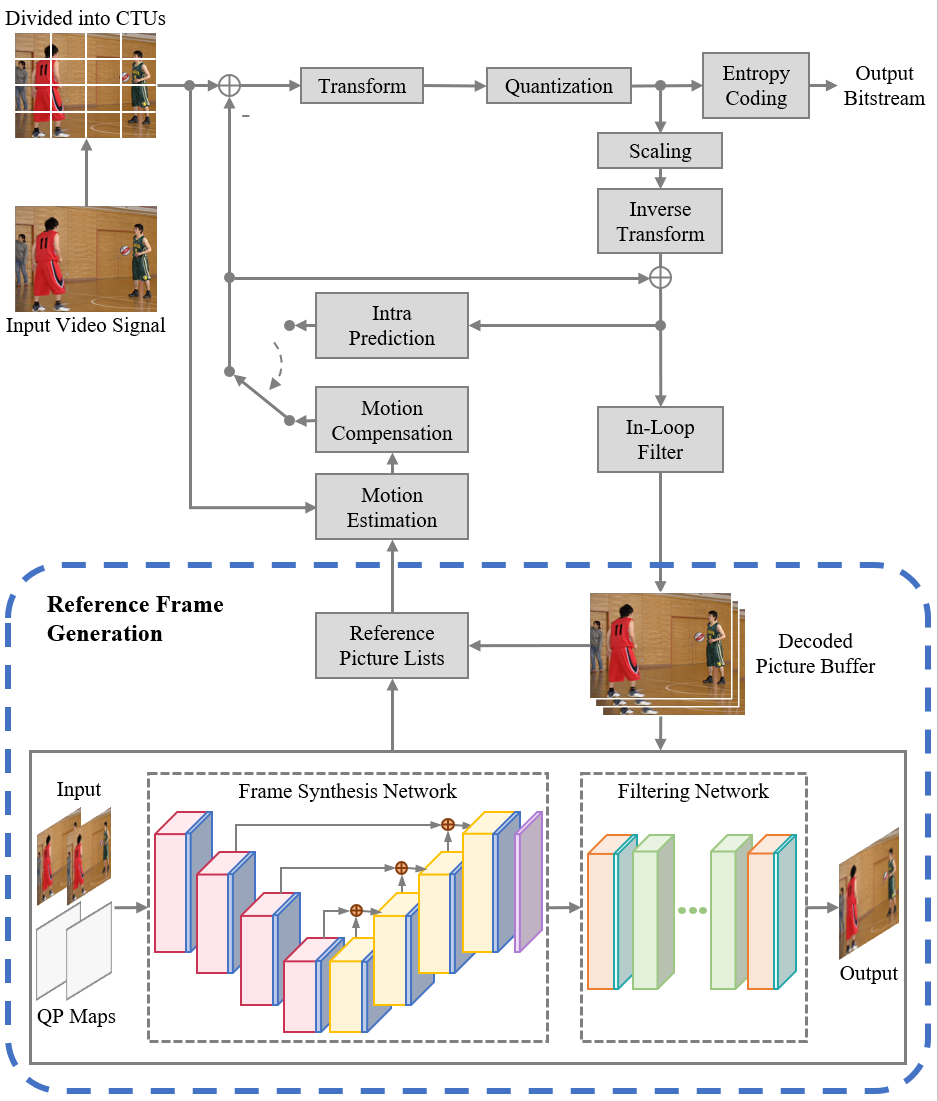


Fig. 1 The frame work of the proposed deep refrence frame generation (DRF) method

图示

描述已自动生成

Fig. 2 the architecure of the frame synthesis network

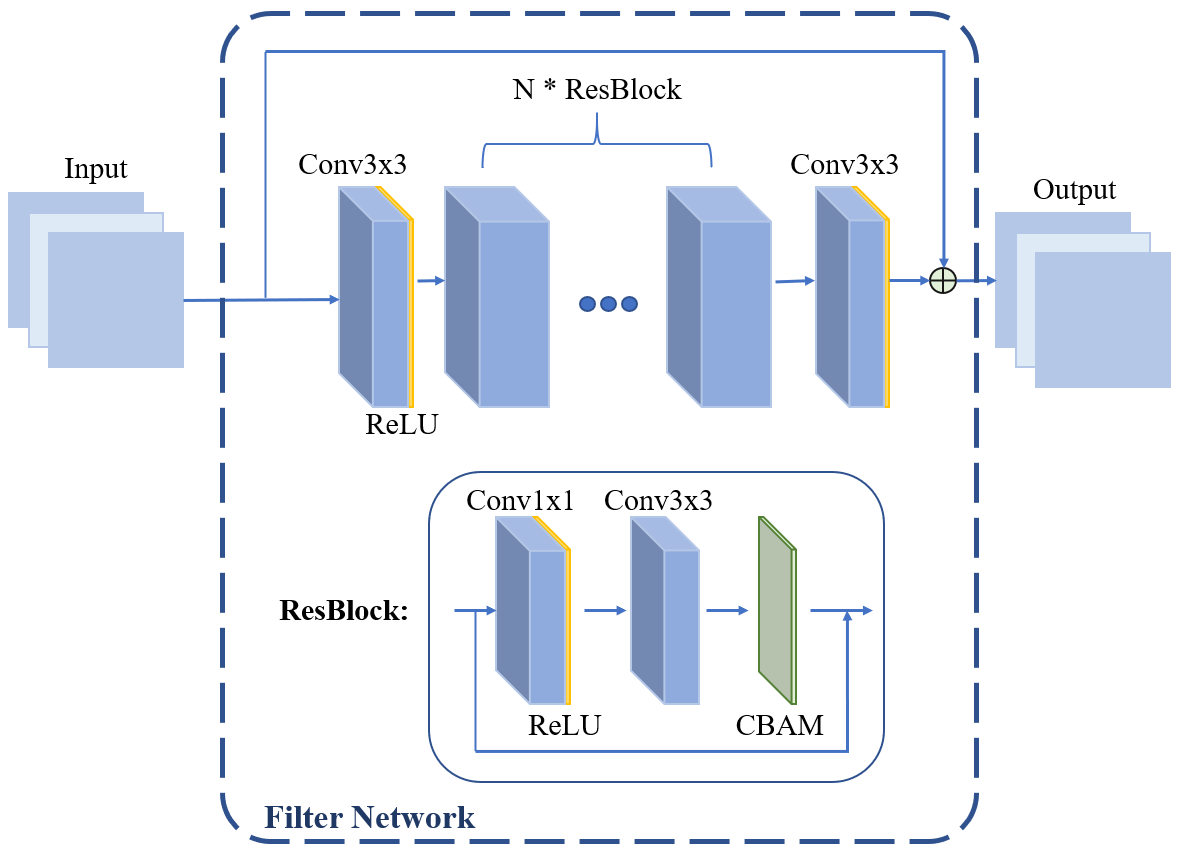


Fig. 3 the architecure of the filter network

|  |  |  |
| --- | --- | --- |
| **Network Information in Inference Stage** | | |
| Mandatory | HW environment: | |
| GPU Type | CPU only |
| Framework: | PyTorch v1.9 |
| Number of GPUs per Task | 0 |
|  |  |
| Number of Parameters (Each Model) | 44769K |
| Total Number of Parameters (All Models) | 44769K x 2 |
| Parameter Precision (Bits) | 32 |
| Memory Parameter (MB) | 170.78MB x 2 |
| Multiply Accumulate (kMAC/pixel) | 1410 |
| Calculation Method | On a block basis |
| Optional |  |  |
| Total Conv. Layers |  |
| Total FC Layers |  |
| Total Memory (MB) |  |
| Batch size: |  |
| Patch size | 256 x 256 |
| Changes to network configuration or weights required to generate rate points |  |
| Peak Memory Usage (Total) |  |
| Peak Memory Usage (per Model) |  |
| Border handling |  |
| Other information: |  |
|  |  |

The model for RA is unchanged from JVET-Y0096.

It was suggested to investigate whether the same gain is still achieved if a NN based loop filter is enabled. This could be done on the basis of the new NN software package.

The proponent do not consider joining the EE at this moment.

Further study.

[JVET-AA0084](https://jvet-experts.org/doc_end_user/current_document.php?id=11760) AHG11: Neural Network based Super Resolution for Video Coding Using Multiple Side Information [R. Chang, L. Wang, X. Xu, S. Liu (Tencent)]

This contribution presents a neural network-based super resolution method using multiple side information for video coding. During the encoding process, the input video is first downsampled by RPR and encoded. The proposed super resolution network is then used to generate the upsampled video during the decoding process. In addition to the reconstruction, multiple side information including prediction, slice QP, base QP and slice type is also fed into the network. Besides, considering the different characteristics between luma and chroma component, different models are designed for luma and chroma component.

Compared with VTM-11.0 + NewMCTF, the related results with libtorch implementation are shown on average {-6.22%, 18.57%, -4.94%} and {-12.05%, 32.44%, -7.44%} BD-rate savings for {Y, Cb, Cr} under RA and AI configurations, while the related results with SADL implementation are shown on average {-6.21%, 18.74%, -4.86%} and {-12.03%, 32.44%, -7.44%} BD-rate savings for {Y, Cb, Cr} under RA and AI configurations.

Was presented in session 23 0550 UTC

Superresolution is always enabled for both luma and chroma. It was suggested using an adaptation mechanism that codes in original resolution when reduced resolution coding does not provide gain.

The method apparently gives higher gain than the methods investigated in EE so far, and also gives gaini chroma for many sequences.

Why is the performance of the two chroma components so different? Could be that PSNR curves are crossing.

Investigate in EE

It was suggested to report in the EE category of superresolution also results with RPR (in adaptive mode) as a second anchor. RPR can use downsampling 1x, 1.5x, 2x, as this is part of its functionality (as per JVET-Y0068).

NN approaches do not necessarily need to use the same downsampling filters as RPR.

[JVET-AA0086](https://jvet-experts.org/doc_end_user/current_document.php?id=11762) AHG11: Small Ad-hoc Deep-Learning Library (SADL) update [F. Galpin, T. Dumas, P. Bordes, E. François (InterDigital)]

See under section 4.8

## AHG12: Enhanced compression beyond VVC capability (59)

### Summary and BoG reports

Contributions in this area were discussed in sessions 3 and 4 at 1300–1500 UTC and 1520–1730 UTC on Wednesday 13 July 2022, and in session 7 at 1300–1330 UTC on Thursday 14 July 2022 (all chaired by JRO).

[JVET-AA0024](https://jvet-experts.org/doc_end_user/current_document.php?id=11897) EE2: Summary Report on Enhanced Compression beyond VVC capability [V. Seregin, J. Chen, G. Li, K. Naser, J. Ström, M. Winken, X. Xiu, K. Zhang]

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-5.0, released at <https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tags/ECM-5.0>. ECM-5.0 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-z-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-z-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** | | | |
| 1.1a | Convolutional cross-component intra prediction model | Nokia  P. Astola  [JVET-AA0057](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0057-v1.zip) | Qualcomm Y.-J. Chang  [JVET-AA0156](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0156-v1.zip) |
| 1.1b | Filter-based linear model | Kwai  C.-W. Kuo  [JVET-AA0125](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0125-v1.zip) | Alibaba  X. Li  [JVET-AA0162](https://jvet-experts.org/doc_end_user/current_document.php?id=11850) |
| 1.1c | Test 1.1a + Test 1.1b | Nokia  P. Astola  Kwai  C.-W. Kuo  [JVET-AA0126](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0126-v1.zip) | Bytedance B. Vishwanath  [JVET-AA0193](https://jvet-experts.org/doc_end_user/current_document.php?id=11882)  Tencent  G. Li  [JVET-AA0209](https://jvet-experts.org/doc_end_user/current_document.php?id=11899) |
| 1.2 | Gradient linear model | Kwai  C.-W. Kuo  [JVET-AA0125](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0125-v1.zip) | Alibaba  X. Li |
| 1.3a | Test 1.1a + Test 1.2 | Nokia  P. Astola  Kwai  C.-W. Kuo  [JVET-AA0126](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0126-v1.zip) | Bytedance B. Vishwanath  [JVET-AA0160](https://jvet-experts.org/doc_end_user/current_document.php?id=11848) |
| 1.3b | Test 1.1a + Test 1.1b + Test 1.2 | Nokia  P. Astola  Kwai  C.-W. Kuo  [JVET-AA0126](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0126-v1.zip) | Bytedance B. Vishwanath  Tencent G. Li  [JVET-AA0209](https://jvet-experts.org/doc_end_user/current_document.php?id=11899) |
| 1.4a | Spatial GPM | OPPO  F.Wang  [JVET-AA0118](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0118-v1.zip) | InterDigital  K. Naser |
| 1.4b | Spatial GPM with restriction in inter-coded slices | OPPO  F.Wang  [JVET-AA0118](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0118-v1.zip) | InterDigital  K. Naser  [JVET-AA0151](https://jvet-experts.org/doc_end_user/current_document.php?id=11839) |
| 1.5 | Chroma intra modes derived from collocated luma and neighboring chroma blocks | Qualcomm  Y.-J. Chang  [JVET-AA0135](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0135-v1.zip) | InterDigital  K. Naser  [JVET-AA0152](https://jvet-experts.org/doc_end_user/current_document.php?id=11840) |
| 1.6a | Weighted chroma prediction | Xidian Univ.  J. Huo  [JVET-AA0078](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0078-v1.zip) | Alibaba  X. Li  [JVET-AA0164](https://jvet-experts.org/doc_end_user/current_document.php?id=11852) |
| 1.6b | Test 1.6a + Test 1.1a | Xidian Univ.  J. Huo  Nokia  P. Astola  [JVET-AA0153](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0153-v1.zip) | Qualcomm Y.-J. Chang  [JVET-AA0157](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0157-v1.zip) |
| 1.7 | IBC adaptation for camera-captured content | Bytedance J. Xu  [JVET-AA0106](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0106-v1.zip) | Ofinno  A. Filippov  [JVET-AA0205](https://jvet-experts.org/doc_end_user/current_document.php?id=11894) |
| **2 Inter prediction** | | | |
| 2.1a | Regression based affine candidate derivation | Qualcomm  Y. Zhang  [JVET-AA0107](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0107-v1.zip) | Tencent  G. Li  [JVET-AA0210](https://jvet-experts.org/doc_end_user/current_document.php?id=11900) |
| 2.1b | Regression based affine candidate derivation with increased affine candidate list size | Qualcomm  Y. Zhang  [JVET-AA0107](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0107-v1.zip) | Tencent  G. Li  [JVET-AA0210](https://jvet-experts.org/doc_end_user/current_document.php?id=11900) |
| 2.1c | Test 2.1b + diversity criterion for ARMC from Test 2.5 | Qualcomm  Y. Zhang  Canon  G. Laroche  [JVET-AA0107](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0107-v1.zip) | Tencent  G. Li  [JVET-AA0210](https://jvet-experts.org/doc_end_user/current_document.php?id=11900) |
| 2.1d | Test 2.1b + Test 2.5 | Qualcomm  Y. Zhang  Canon  G. Laroche  [JVET-AA0176](https://jvet-experts.org/doc_end_user/current_document.php?id=11864) | Tencent  G. Li  [JVET-AA0211](https://jvet-experts.org/doc_end_user/current_document.php?id=11901) |
| 2.2a | Motion compensation boundary padding | Xiaomi  F. Le Léannec  Qualcomm  Z. Zhang  [JVET-AA0096](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0096-v1.zip) | Ofinno  V. Rufitskiy  [JVET-AA0206](https://jvet-experts.org/doc_end_user/current_document.php?id=11895) |
| 2.2b | Motion compensation boundary padding with reduced memory usage | Xiaomi  F. Le Léannec  Qualcomm  Z. Zhang  [JVET-AA0096](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0096-v1.zip) | Ofinno  V. Rufitskiy  [JVET-AA0206](https://jvet-experts.org/doc_end_user/current_document.php?id=11895) |
| 2.3 | MMVD and affine MMVD extension | Bytedance  M. Salehifar  [JVET-AA0116](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0116-v1.zip) | Tencent  G. Li  [JVET-AA0212](https://jvet-experts.org/doc_end_user/current_document.php?id=11902) |
| 2.4 | Refined motion for ARMC | Bytedance  Y. Wang  [JVET-AA0072](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0072-v1.zip) | Canon  G. Laroche |
| 2.5 | Diversity criterion for ARMC with merge candidate modifications | Canon  G. Laroche  [JVET-AA0092](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0092-v1.zip) | Bytedance  Y. Wang  [JVET-AA0158](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0158-v1.zip) |
| 2.6a | Test 2.4 + Test 2.5 | Canon  G. Laroche  Bytedance  Y. Wang  [JVET-AA0093](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0093-v2.zip) | Tencent  G. Li  [JVET-AA0213](https://jvet-experts.org/doc_end_user/current_document.php?id=11903) |
| 2.6b | Test 2.3 + Test 2.4+ Test 2.5 | Canon  G. Laroche  Bytedance  Y. Wang  [JVET-AA0093](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0093-v2.zip) | Tencent  G. Li  [JVET-AA0213](https://jvet-experts.org/doc_end_user/current_document.php?id=11903) |
| 2.7a | GPM adaptive blending | KDDI  Y. Kidani  Kwai  N. Yan  [JVET-AA0058](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0058-v1.zip) | Bytedance  Z. Deng  [JVET-AA0159](https://jvet-experts.org/doc_end_user/current_document.php?id=11847) |
| 2.7b | Test 2.7a + block-size dependent signalling | KDDI  Y. Kidani  Kwai  N. Yan  [JVET-AA0058](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0058-v1.zip) | Bytedance  Z. Deng |
| 2.8 | Longer chroma filters for RPR | Ericsson  K. Andersson  [JVET-AA0042](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0042-v1.zip) | Qualcomm C. S. Coban  [JVET-AA0192](https://jvet-experts.org/doc_end_user/current_document.php?id=11881) |
| 2.9a | RPR luma filters 12-tap | Ericsson  R. Yu  [JVET-AA0042](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0042-v1.zip) | Qualcomm C. S. Coban  [JVET-AA0192](https://jvet-experts.org/doc_end_user/current_document.php?id=11881) |
| 2.9b | RPR luma filters 12-tap for non-affine blocks and RPR luma filters 10-tap for affine blocks | Ericsson  R. Yu  [JVET-AA0042](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0042-v1.zip) | Qualcomm C. S. Coban  [JVET-AA0192](https://jvet-experts.org/doc_end_user/current_document.php?id=11881) |
| 2.10a | Test 2.8 + Test 2.9a | Ericsson  R. Yu  [JVET-AA0042](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0042-v1.zip) | Qualcomm C. S. Coban H. Golestani  [JVET-AA0192](https://jvet-experts.org/doc_end_user/current_document.php?id=11881) |
| 2.10b | Test 2.8 + Test 2.9b | Ericsson  R. Yu  [JVET-AA0042](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0042-v1.zip) | Qualcomm C. S. Coban H. Golestani  [JVET-AA0192](https://jvet-experts.org/doc_end_user/current_document.php?id=11881) |
| **3 Screen content coding** | | | |
| 3.1 | IBC merge mode with block vector differences | Bytedance  N. Zhang  [JVET-AA0061](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0061-v2.zip) | Ofinno  D. Ruiz Coll  [JVET-AA0201](https://jvet-experts.org/doc_end_user/current_document.php?id=11890) |
| 3.2 | IBC with reconstruction reordering | Bytedance  Z. Deng  [JVET-AA0070](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0070-v1.zip) | Ofinno  D. Ruiz Coll  [JVET-AA0202](https://jvet-experts.org/doc_end_user/current_document.php?id=11891) |
| 3.3 | Test 3.1 + Test 3.2 | Bytedance  N. Zhang  [JVET-AA0062](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0062-v1.zip) | Ofinno  D. Ruiz Coll  [JVET-AA0203](https://jvet-experts.org/doc_end_user/current_document.php?id=11892) |
| **4 Transform** | | | |
| 4.1a | Inter-MTS optimization | Qualcomm  B. Ray  [JVET-AA0133](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0133-v1.zip) | Sharp  Z. Fan  [JVET-AA0182](https://jvet-experts.org/doc_end_user/current_document.php?id=11870) |
| 4.1b | Inter-MTS optimization without KLT | Qualcomm  B. Ray  [JVET-AA0133](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0133-v1.zip) | Sharp  Z. Fan  [JVET-AA0182](https://jvet-experts.org/doc_end_user/current_document.php?id=11870) |
| **5 In-loop filtering** | | | |
| 5.1a | Adaptive filter shape switch for ALF | Bytedance  W. Yin  [JVET-AA0095](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0095-v1.zip) | Qualcomm  N. Hu |
| 5.1b | Longer filter length for ALF | Bytedance  W. Yin  [JVET-AA0095](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0095-v1.zip) | Qualcomm  N. Hu |
| 5.2 | Using samples before deblocking filter for ALF | Qualcomm  N. Hu  [JVET-AA0095](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0095-v1.zip) | Bytedance  W. Yin  [JVET-AA0168](https://jvet-experts.org/doc_end_user/current_document.php?id=11856) |
| 5.3a | Test 5.1a + Test 5.2 | Bytedance  W. Yin  Qualcomm  N. Hu  [JVET-AA0095](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0095-v1.zip) | Ericsson  K. Andersson |
| 5.3b | Test 5.1b + Test 5.2 | Bytedance  W. Yin  Qualcomm  N. Hu  [JVET-AA0095](https://jvet-experts.org/doc_end_user/documents/27_Teleconference/wg11/JVET-AA0095-v1.zip) | Ericsson  K. Andersson |

**Test 1.x Intra prediction**

**Test 1.1a: Convolutional cross-component intra prediction model**

In convolutional cross-component intra prediction model, chroma samples are predicted from collocated and adjacent reconstructed luma samples.

A convolutional 7-tap filter consists of a 5-tap spatial samples, a nonlinear term, and a bias term. The input to the spatial 5-tap component of the filter consists of a center (C) luma sample which is collocated with the chroma sample to be predicted and its above/north (N), below/south (S), left/west (W) and right/east (E) neighbors as illustrated in Figure 1.

A picture containing text, shoji, crossword puzzle

Description automatically generated

**Figure 1. Spatial part of the convolutional filterSpatial part of the convolutional filter**

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

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

For 10-bit content, it is calculated as:

P = ( C\*C + 512 ) >> 10

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

Output of the filter is calculated as a convolution between the filter coefficients ci and the input values and clipped to the range of valid chroma samples:

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

Filter coefficients are derived from the reconstructed luma and chroma neighbourhood shown in Figure 2 by performing MSE minimization.

Chart

Description automatically generated

**Figure 2. Reference area (with its paddings) used to derive the filter coefficients**

The mode flag is only signalled if intra prediction mode is LM\_CHROMA\_IDX (to enable single mode CCCM) or MMLM\_CHROMA\_IDX (to enable multi-model CCCM).

**Test 1.1b: Filter-based linear model**

Filter-based linear model extends CCLM model by including neighboring reconstructed luma samples as follows:

where is the to-be-predicted chroma sample; is the -th reconstructed downsampled luma sample value surrounding the chroma sample, is the coefficient associated with , is the offset, and is the number of the used downsampled luma samples.

To derive the model parameters, up to 6 rows and columns of chroma samples above and left to the current CU are used.

The mode is applied to chroma CUs with the area greater or equal to 128, and 4 filter shapes (with N equal to 2 or 6) are adaptively switched at CU level.

The mode is combined with multi-model linear model (MMLM) for CUs with the area greater or equal to 256, and it disabled with multi-directional linear model (MDLM).

**Test 1.2: Gradient linear model**

In the gradient linear model, luma sample gradients are used instead of the reconstructed sample values to derive a linear model.

,

where *G* is the sample gradient.

16 gradient patterns shown in Figure 3are supported in the mode and the pattern choice is signaled.

A picture containing crossword puzzle, text, mounted

Description automatically generated

**Figure 3. Gradient patterns**

When CCLM mode is enabled, two flags are signaled separately for Cb and Cr components to indicate the usage of the gradient linear model, if it is enabled, a syntax element is further signaled to select one of 16 gradient filters utilized for the gradient calculation.

**Test 1.3: Combination of linear model tests**

Convolutional cross-component intra prediction (Test 1.1a), filter-based linear (Test 1.1b), and gradient linear (Test 1.2) models are tested in combinations.

Test 1.1c combines Test 1.1a and Test 1.1b.

Test 1.3a combines Test 1.1a and Test 1.2.

Test 1.3b combines Test 1.1a, Test 1.1b, and Test 1.2.

**Test 1.4: Spatial GPM**

In ECM, GPM supports inter-inter and inter-intra partitions. In EE, GPM with intra-intra partitions is tested.

In this mode, a candidate list is built which includes partition split and two intra prediction modes shown in Figure 4. The selected candidate index is signalled.

**Figure 4. Spatial GPM candidates**

The list is reordered using template shown in Figure 5. GPM blending process is not used in the template, and SAD between the prediction and reconstruction of the template is used for ordering.



**Figure 5. GPM template**

In test 1.4b, spatial GPM is not applied in inter slices.

**Test 1.5: Chroma intra modes derived from collocated luma and neighboring chroma blocks**

In the test, chroma DIMD mode is replaced by using multiple DM modes and the non-CCLM chroma intra mode list is constructed in the following order, where the first 9 unique modes are added into the non-CCLM chroma mode list:

* 5 DM modes in the order of C, TL, TR, BL and BR as shown in Figure 6(left).
* 5 neighbouring chroma modes from the neighboring positions shown as blocks 0, 1, 2, 3, and 4 in Figure 6(right).
* The derived chroma modes by offsetting +1 or -1 to the first two intra modes in the mode list.
* Default modes: PLANAR\_IDX, VER\_IDX, HOR\_IDX, DC\_IDX, VDIA\_IDX, VER\_IDX - 4, VER\_IDX + 4, HOR\_IDX - 4, HOR\_IDX + 4.

A picture containing text

Description automatically generated

**Figure 6. DMs (left) and neighbouring chroma modes (right)**

**Test 1.6: Weighted chroma prediction**

In weighted chroma prediction, for each sample in the downsampled luma block, the luma absolute difference vector between and the neighboring luma vector is obtained. Weights are derived as a nonlinear mapping model based on .

Diagram

Description automatically generated

**Figure 7. Weighted chroma prediction**

Predicted chroma samples with x = 0…W − 1, y = 0…H − 1 are derived by multiplying the weights and neighboring chroma vector:

In Test 1.6b, the weighted chroma prediction is tested in combination with convolutional cross-component linear model (Test 1.1a)

**Test 1.7: IBC adaptation for camera-captured content**

In the test, IBC reference range is reduced from 2 CTU rows to 2x128 rows shown in Figure 8.

Shape, square

Description automatically generated

**Figure 8. IBC reference area**

At encoder side to reduce the complexity, the local search range is set to [–8,8] horizontally and [–8,8] vertically centered at the first block vector predictor of the current CU. This encoder modification is not applied to SCC sequences (class F and TGM).


Cross-component methods: For camera captured content and class F, 1.1a is giving highest gain, and 1.2 does not show additional benefit in the combination (1.3x). For class TGM, 1.2 has some additional benefit. It is also noted that 1.2 is more straightforward to implement than 1.1a.

Both are additional modes. Decision: Adopt Test 1.3a\* (lower encoder run time than 1.3), normative parts from JVET-AA0126/JVET-AA0057.

Spatial GPM (1.4x) has too large increase in encoding time, while the compression benefit is comparably low. No action.

Test 1.5 intends complexity reduction by replacing chroma DIMD, while still giving small gain. The cross-checker confirms that conceptually the method is a simplification, though not significantly observable in run time. The previous proponent of DIMD argues that the encoding of the new proposal appears more complex.

The tool has valuable benefit, but “small level” optimization of selected tools in terms of implementability is not of highest importance at this stage.

Test 1.6: No results on RA yet, and no run time reported (unreliable according to proponents). From the description, the method requires additional computations at the decoder. No action at this point.

Test 1.7: This proposal enables IBC for camera captured content, showing 0.9% luma gain with 30% encoder runtime increase for AI. The proposal also includes a normative change by restricting the reference range to 2x128 rows (instead of 2xmaxCTU as currently defined in ECM which would be 2x256, and this maxCTU size is used in camera content classes). For screen content CTC, CTU size 128 is used, such that such a change would not make a difference in current CTC.

When IBC is enabled in VTM for CC content, it gives 0.5%, while increasing encoding time by 60%.

Decision: Adopt the normative aspect JVET-AA0106, reducing the max search range to 2x128 rows. Regarding the relative large increase in encoder run time, it is not desirable to enable IBC for camera captured content in CTC.

It is suggested to perform further study if the encoder run time could be reduced, e.g. designing a fast algorithm for estimating the IBC displacement vectors rather than using the hash-based search which is known to be only effective for screen content.

**Test 2.x Inter prediction**

**Test 2.1: Regression based affine candidate derivation**

Regression based motion vector field method models the motion vectors of each block on a sub-block level based on the spatially neighboring motion vectors by a linear model. The motion vectors and center positions from the neighboring subblocks of the current CU, as illustrated in Figure 9, are used as the input to the linear regression process to derive a set of linear model parameters.

Diagram

Description automatically generated

**Figure 9. Regression based motion vector field**

The MSE minimization problem is solved for the following equations

where *a* and *b* are the linear model parameters. Regression based affine candidates are then derived by applying the linear model to the top-left, top-right, and bottom-left positions of a block resulting in three CPMVs.

In the test, the subblock motion field from a previously coded affine CUs and the motion vectors from the adjacent subblocks of a current CU are used as the input for the regression process.

In Test 2.1b, the number of affine candidates for ARMC is increased from 15 to 30, but the output list size is kept as 15.

In Test 2.1c, the diversity criterion for ARMC sorting from EE2-2.5 is applied on top of Test 2.1b.

In Test 2.1d, Test 2.1b is combined with Test 2.5.

**Test 2.2: Motion compensation boundary padding**

In the test, samples outside of the picture boundary are derived by motion compensation instead of using only repetitive padding as shown as red area in Figure 10. The total padded area size is increased by 64 (Test 2.2a) or 16 (Test 2.2b) compared to ECM.



**Figure 10. MC boundary padding method**

For motion compensation padding, MV of a 4×4 boundary block is utilized to derive a M×4 or 4×M padding block. The value M is derived as the distance of the reference block to the picture boundary as shown in Figure 11.



**Figure 11. An example of deriving a M×4 padding block with a left padding direction**

If boundary block is intra coded, then MV is not available, and M is set equal to 0. If M is less than 64 or 16, the rest of the padded area is filled with the repetitive padded samples. The pixels in MC padding block are corrected with an offset, which is equal to the difference between the DC values of the reconstructed boundary block and its corresponding reference block.

**Test 2.3: MMVD and affine MMVD extension**

For a bi-prediction MMVD base in VVC or ECM, the MVD is added to one list, but derived from the first list MVD with a possible scaling and mirroring depending on the POC differences for the other list, known as two-sided MMVD.

In the test, MMVD is modified with the following aspects:

* One-sided MMVD (or affine MMVD) is added, where individual MVD is used for each list independently, where a non-zero MVD is applied for list X, but a zero MVD is applied for list (1-X).
* The scaling process for the two-sided MMVD is removed.
* Candidates with similar costs are pruned.
* The number of the bases for MMVD is increased from 2 to 3. For affine MMVD, the number of bases is increased from 1, to up to 3, depending on the neighboring block affine flags.

**Test 2.4: ARMC with refined motion**

In the test, each merge candidate in the merge candidate list is refined using TM/multi-pass DMVR first and the refined motion will be used in ARMC to reorder the merge candidate list as follows:

* When multi-pass DMVR is used to derive the refined motion, only the first pass (i.e., PU level) of multi-pass DMVR is applied in reordering.
* When template matching is used to derive the refined motion, the template size is set equal to 1. Only the above or left template is used during the motion refinement of TM when the block is flat (w > 2×h) or narrow (h > 2×w). TM is extended to perform 1/16-pel MVD precision. The first four merge candidates are reordered with the refined motion in TM merge mode.
* When constructing the AMVP list, an MVP candidate with a TM cost larger than a threshold, which is equal to five times of the cost of the first MVP candidate, is skipped.

**Test 2.5: Diversity criterion for ARMC with merge candidate modifications**

In ARMC, MV candidates are reordered based on TM cost. In the test, additional consideration of the cost difference between a candidate pair is considered in the reordering process. A candidate is considered redundant if the cost difference between a candidate and its predecessor is less than a lambda value e.g. |D1-D2| < λ, where D1 and D2 are the costs obtained during the first ARMC ordering and λ is the Lagrangian parameter used in the RDO process at encoder side. The redundant candidate is placed after non-redundant candidates.

This algorithm is applied to the regular, TM, BM, affine merge modes, MMVD, and MVD sign prediction where ARMC is utilized.

A set of λ values corresponding to each signaled QP offset is provided in the SPS or in the slice header for the QP offsets which are not present in the SPS.

Additionally, the following modifications have been applied to regular merge list:

- The same MV threshold is used for temporal, non-adjacent, and other candidates.

- No ARMC reordering for sub-groups is performed

- Up to 4 pairwise candidates are added after the first ARMC reordering for regular, TM and BM merge modes.

**Test 2.6: Combination of Test 2.3, Test 2.4, and Test 2.5**

Test 2.6a is a combination of ARMC with refined motion (Test 2.4) and diversity criterion for ARMC with merge candidate modifications (Test 2.5).

Test 2.6b is a combination of MMVD and affine MMVD extension (Test 2.3), ARMC with refined motion (Test 2.4), and diversity criterion for ARMC with merge candidate modifications (Test 2.5).

**Test 2.7: GPM adaptive blending**

In GPM mode, the blending of the two partitions shown is Figure 12 is done using two integer blending matrices (*W0* and *W1*) for the fixed area of *τ =* 2 samples on each side of the GPM partition split boundary, and the weights are derived by a ramp function with the displacement (*d*) and the blending area size (*τ*) as shown in **Figure 13**.

Chart

Description automatically generated

**Figure 12. Blending for GPM partitions**

Chart, line chart

Description automatically generated

**Figure 13. GPM blending weights**

In this test, besides the existing blending area, extra blending area sizes of *τ*/4, *τ*/2, 2*τ*, and 4*τ* are added for the GPM mode as shown in Figure 14, the weighting precision is increased from the maximum weight value of 8 to 32 to accommodate the new blending sizes. The selected blending area size is signalled at CU-level.

A picture containing text, laser

Description automatically generated

**Figure 14. GPM lending weights for the additional blending sizes**

In Test 2.7b, block size dependent restriction on blending area sizes is introduced such that when the shorter side of the current block is larger than 16 (*τ*, 2*τ*, and 4*τ*)can be used*,* otherwise (*τ*/4, *τ*/2, and *τ*) can be selected.



Test 2.1: Still uses affine block size 8x8. Increased complexity due to regression, and also due to additional pruning, partially could be due to using more 4x4 subblocks.

Encoder run time increase is relatively low. Adoption of the proposal was supported by several experts.

Decision: Adopt test 2.1b, JVET-AA0107.

Test 2.2: It is pointed out that the boundary padding is requiring an additional processing stage after decoding (including loop filter) the entire picture, beyond additional memory that is needed. Results of 2.2.b indicate that a large padding area (64) is not necessary.

Taking into account that at current stage of exploration stud of implementability is not of highest importance, adopting the method was supported by several experts

Decision: Adopt Test 2.2b (16 samples padding area) from JVET-AA0096

Test 2.6b indicates that 2.3-2.5 have almost additive gains. All three tests can only be used when template matching is on. The test 2.1d indicates that the gains of at least that proposal are additive on top of 2.1b. It was expressed that all three elements of 2.6b have some individual benefit and the tradeoff of rate reduction versus complexity is appropriate.

Decision: Adopt Test 2.6b (JVET-AA0093)

Test 2.7 allows additional widths of the GPM blending area. Test 2.7a allows five options (including the existing one), whereas 2.7b allows three out of five options, but the specific width is dependent on block size. This also reduces encoder run time. Test 2.7a has higher gain (0.25% for camera content), for class TGM no difference.

It was asked if the proponents also tested the option of disabling the blending for screen content (as was proposed earlier for VVC, as well as JVET-Z0137). The proponents claim that their method performs better for class TGM than turning blending off (it is noted that the method in test 2.7 also allows blending with subsample width which might perform better than turning off).

It is also noted that the subsample width blending is not using interpolation, just smaller weight of the adjacent sample beyond the boundary.

Decision: Adopt Test2.7a from JVET-AA0058

**Tests 2.8, 2.9, 2.10: RPR filters**

In Test 2.8, chroma RPR filter size is increased from 4-tap to 6-tap.

In Test 2.9a, luma RPR filter size is increased from 8-tap to 12-tap.

In Test 2.9b, luma RPR filter size is 12-tap for non-affine blocks and 10-tap for affine blocks.

Test 2.10a is a combination of Test 2.8 and Test 2.9a.

Test 2.10b is a combination of Test 2.8 and Test 2.9b.

Test results in LB configuration following RPR CTC are summarized below.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | scaling ratio 1.5x | | | | | | |
|  | BD-rate PSNR1 | | | BD-rate PSNR2 | | |  |
| EE2 TEST # | Y | U | V | Y | U | V | DecT |
| 2.8 | -0.02% | -0.52% | -0.56% | -0.04% | -0.66% | -0.70% | 100% |
| 2.9a | -0.72% | -0.26% | -0.38% | -0.45% | -0.33% | -0.46% | 99% |
| 2.9b | -0.70% | -0.34% | -0.30% | -0.43% | -0.38% | -0.39% | 99% |
| 2.10a (2.8 + 2.9a) | -0.73% | -0.77% | -0.62% | -0.44% | -0.96% | -0.81% | 100% |
| 2.10b (2.8 + 2.9b) | -0.76% | -0.98% | -0.80% | -0.47% | -1.23% | -0.97% | 100% |
|  | scaling ratio 2.0x | | | | | | |
|  | BD-rate PSNR1 | | | BD-rate PSNR2 | | |  |
| EE2 TEST # | Y | U | V | Y | U | V | DecT |
| 2.8 | 0.01% | -0.75% | -1.14% | 0.02% | -1.18% | -1.28% | 100% |
| 2.9a | -0.45% | 0.12% | -0.12% | -0.18% | 0.01% | -0.17% | 99% |
| 2.9b | -0.40% | 0.23% | 0.15% | -0.12% | 0.10% | 0.11% | 99% |
| 2.10a (2.8 + 2.9a) | -0.47% | -0.55% | -0.91% | -0.19% | -1.00% | -1.09% | 100% |
| 2.10b (2.8 + 2.9b) | -0.43% | -0.78% | -1.02% | -0.13% | -1.21% | -1.17% | 100% |

It is noted that the RPR CTC may not be practically realistic, but the results provide evidence that the proposed filters (which are identical with the MC filters for the case of upsampling) provide gain.

One expert points out that he found that there seems to be a mismatch with the RA configuration of RPR (which is not part of the CTC). It should be checked if this is also the case with ECM5.0, after various bug fixes had been done in context of RPR since the last meeting.

Solution in Test 2.10b is conceptually similar to the approach of VTM (except for scaling up the length of all filters), by using shorter filters for luma in affine blocks. It also has slightly better compression performance than 2.10a which uses same length for affine and non-affine cases.

Decision: Adopt Test 2.10b from JVET-AA0042.

**Test 3.x Screen content tools**

**Test 3.1: IBC merge mode with block vector differences**

IBC merge mode with block vector differences is tested. The distance set is {1-pel, 2-pel, 4-pel, 8-pel, 12-pel, 16-pel, 24-pel, 32-pel, 40-pel, 48-pel, 56-pel, 64-pel, 72-pel, 80-pel, 88-pel, 96-pel, 104-pel, 112-pel, 120-pel, 128-pel}, and the BVD directions are two horizontal and two vertical directions.

The base candidates are selected from the first five candidates in the reordered IBC merge list. And based on the SAD cost between the template (one row above and one column left to the current block) and its reference for each refinement position, all the possible MBVD refinement positions (20×4) for each base candidate are reordered. Finally, the top 8 refinement positions with the lowest template SAD costs are kept as available positions, consequently for MBVD index coding.

**Test 3.2: IBC with reconstruction reordering**

In the test, horizontal and vertical flip modes are added to IBC. A flag is signalled for an IBC AMVP coded block, indicating whether the reconstruction is flipped, and if it is flipped, another flag is further signaled specifying the flip type. For IBC merge, the flip type is inherited from neighbouring blocks without syntax signalling.

When a horizontal flip is applied, the vertical component of the BV is not signalled and inferred to be equal to 0. Similarly, the horizontal component of the BV is not signaled and inferred to be equal to 0 when a vertical flip is applied.

BV adjustment considering the neighboring block flip mode is applied to refine the block vector candidate. As shown in Figure 15, (*xnbr*, *ynbr*) and (*xcur*, *ycur*) represent the coordinates of the center sample of the neighbouring block and the current block, respectively, *BVnbr* and *BVcur* denotes the BV of the neighbouring block and the current block, respectively.

Instead of directly inheriting the BV from a neighbouring block, the horizontal component of *BVcur* is calculated by adding a motion shift to the horizontal component of *BVnbr* (denoted as *BVnbrh*) in case that the neighbouring block is coded with a horizontal flip, i.e., *BVcurh* =2(*xnbr* -*xcur*) + *BVnbrh*.

Similarly, the vertical component of *BVcur* is calculated by adding a motion shift to the vertical component of *BVnbr* (denoted as *BVnbrv*) in case that the neighbouring block is coded with a vertical flip, i.e., *BVcurv* =2(*ynbr* -*ycur*) + *BVnbrv*.

Diagram

Description automatically generated Diagram

Description automatically generated

Figure 15. BV adjustment for horizontal flip (left) and vertical flip (right)

**Test 3.3: Combination of Test 3.1 and Test 3.2**

This test is a combination of IBC merge mode with block vector differences (Test 3.1) and IBC with reconstruction reordering (Test 3.2), where IBC MBVD coded block does not inherit flip type from a neighbour block.



Combination 3.3 indicates that gains are additive. The cross-checker and other experts supported the adoption of both proposals.

It is noted that the test 3.2 (flipping) might require block size restriction (such as using only for 16 or below) for a practical implementation. This most likely would not affect the performance, but is not urgent at the current stage of exploration

It is also noted that test 3.1 relies on TM.

Decision: Adopt Test 3.3 from JVET-AA0062.

**Test 4.x Transforms**

**Test 4.1: Inter MTS optimization**

In ECM-5.0, four candidates {(DST7, DST7), (DST7, DCT8), (DCT8, DST7), (DCT8, DCT8)} are used in inter MTS. In the test, the following modifications for inter MTS are applied:

* For the larger resolution sequences (width > 1080) maximum CU size for inter MTS usage is set to 32 (i.e., inter MTS is used for CU with width <=32 and height <=32), and for the remaining sequences (smaller resolution) it is set to 16.
* For 4-pt, 8-pt and 16-pt transforms, the current MTS transform cores, i.e., DST-7 and DCT-8, is replaced with separable KLTs.

In Test 4.1a, both aspects are applied, and in Test 4.1b only the first modification is utilized.



It is confirmed that the KLTs came from an old proposal in VVC development, trained outside the test set.

Both changes would be normative.

It is noted that currently MTS is disabled in CTC for inter, as the encoder run time increase was not asserted to be justified considering the gain it provides.

The proposal comes with an encoder optimization which avoids the run time increase (but also requires some normative restrictions, see first bullet point). It was asked what the performance would be with only encoder optimization without the normative restrictions. This had not been tested.

Test 4.1b is the same approach using the existing inter MTS, and test 4.1a indicates that replacing by KLT provides an additional 0.1% gain. This would add another transform kernel which is not used intra MTS, but that consideration is not overly important at this stage.

Several experts supported adoption of the proposal (including KLT).

Decision: Adopt Test 4.1a from JVET-AA0133. It was also decided that inter MTS should be enabled in the VTM anchor, and it should be studied if the runtime of VTM with MTS enabled could be reduced by using the encoder optimizations of test 4.1b but implementing the restrictions in a non-normative way (by encoder decision).

Continued from here in session 7

**Loop filters**

**Test 5.1a: Adaptive filter shape switch for ALF**

In the test, two candidate filter shapes: diamond shape used in ECM as shown in Figure 16 and introduced cross shape as shown in Figure 17 are adaptively selected in the luma ALF. The number of coefficients of a luma filter is 22 for both the filter shapes.

In each adaptation parameter set (APS), a shape index for the derived luma filters is signalled to decoder. Each APS contains luma filters that are associated with the filter shape index.

For each CTB, an APS index is signaled to indicate which luma filter shape is used to filter the current CTB. When filtering a luma sample, the coefficients and clip indices are also rearranged according to the corresponding filter shape.

图表

描述已自动生成

Figure 16. The diamond shape of ALF in ECM-5.0

图示

低可信度描述已自动生成

Figure 17. Cross filter shape for ALF

**Test 5.1b: Longer filter length for ALF**

In the test, the diamond ALF filter shape is replaced with the longer cross shape filter shown on Figure 17.

Test 5.2: Using samples before deblocking filter for ALF

In the test, samples before deblocking filters are used as additional inputs for ALF. A final ALF sample is derived by weighting the regular ALF and the filter applied to the samples before the deblocking filter. Specifically, a filtered sample is derived as

where is the clipped difference between a neighboring sample and current sample , is the clipped difference between an intermediate sample and current sample and is the clipped difference between a neighboring sample before DBF and current sample . The filter coefficients are signalled. In this test, 3x3 diamond shape is applied to the samples before the deblocking filter. In an APS, a flag is signalled to indicate whether samples before DBF are used for ALF. In this proposal, this flag is always set as true at encoder.

**Test 5.3: Combination for ALF related tests**

Test 5.3a is a combination of Test 5.1a and Test 5.2

Test 5.3b is a combination of Test 5.1b and Test 5.2



Test 5.1: The cross-shape filter has same number of coefficients, but might require additional line buffers. Current ECM does not use virtual boundary processing, such that it is unlikely that the usage of samples which are farther away from current sample could have effect on visual quality at CTU row boundaries. The effect of switching between the cross shape and the diamond shape appears low in terms of benefit (comparing Tests 5.1a/b and 5.3a/b, aso the encoder runtime is faster in the b tests that always use the cross shape).

Test 5.2: The usage of samples before deblocking requires more computation (additional 3x3 filter shape). The fact that it requires to store samples both before and after deblocking might be a larger concern in a hardware implementation, it also might have some impact on the pipeline of filter stages.

Combination tests 5.3x indicate that the gains are adding up.

Decision: Adopt Test 5.3b from JVET-AA0095.

### EE2 contributions: Enhanced compression beyond VVC capability (22)

Contributions in this area were discussed in session X at XXXX–XXXX UTC on XXday XX July 2022 (chaired by JRO).

[JVET-AA0042](https://jvet-experts.org/doc_end_user/current_document.php?id=11717) EE2-2.8, 2.9, 2.10: Longer luma and chroma filters for RPR [K. Andersson, R. Yu (Ericsson)]

[JVET-AA0192](https://jvet-experts.org/doc_end_user/current_document.php?id=11881) Crosscheck of JVET-AA0042-v1 (EE2-2.8, 2.9, 2.10: Longer luma and chroma filters for RPR) [C. S. Coban, H. Golestani (Qualcomm) [late]

[JVET-AA0057](https://jvet-experts.org/doc_end_user/current_document.php?id=11733) EE2-1.1a: Convolutional cross-component intra prediction model [P. Astola, J. Lainema, R. G. Youvalari, A. Aminlou, K. Panusopone (Nokia)]

[JVET-AA0156](https://jvet-experts.org/doc_end_user/current_document.php?id=11844) Crosscheck of JVET-AA0057 (EE2-1.1a: Convolutional cross-component intra prediction model) [Y.-J. Chang (Qualcomm)] [late]

[JVET-AA0058](https://jvet-experts.org/doc_end_user/current_document.php?id=11734) EE2-2.7: GPM adaptive blending (JVET-Z0059, JVET-Z0137) [Y. Kidani, H. Katou, K. Unno, K. Kawamura (KDDI), N. Yan, X. Xiu, W. Chen, H.-J. Jhu, C.-W. Kuo, X. Wang (Kwai)]

[JVET-AA0159](https://jvet-experts.org/doc_end_user/current_document.php?id=11847) Crosscheck of JVET-AA0058 (EE2-2.7: GPM adaptive blending (JVET-Z0059, JVET-Z0137) [Z. Deng (Bytedance)] [late]

[JVET-AA0061](https://jvet-experts.org/doc_end_user/current_document.php?id=11737) EE2-3.1: IBC Merge Mode with Block Vector Differences [N. Zhang, J. Xu, K. Zhang, M. Salehifar, L. Zhang (Bytedance)]

[JVET-AA0201](https://jvet-experts.org/doc_end_user/current_document.php?id=11890) Crosscheck of JVET-AA0061 (EE2-3.1: IBC Merge Mode with Block Vector Differences) [D. Ruiz Coll, V. Warudkar (Ofinno)] [late]

[JVET-AA0062](https://jvet-experts.org/doc_end_user/current_document.php?id=11738) EE2-3.3: Combination of EE2-3.1 and EE2-3.2 [N. Zhang, Z. Deng, K. Zhang, J. Xu, M. Salehifar, L. Zhang (Bytedance)]

[JVET-AA0203](https://jvet-experts.org/doc_end_user/current_document.php?id=11892) Crosscheck of JVET-AA0062 (EE2-3.3: Combination of EE2-3.1 and EE2-3.2) [D. Ruiz Coll, V. Warudkar (Ofinno)] [late] [miss]

[JVET-AA0225](https://jvet-experts.org/doc_end_user/current_document.php?id=11915) Crosscheck of JVET-AA0062 (EE2-3.3: Combination of EE2-3.1 and EE2-3.2) [Y. Kidani, K. Kawamura (KDDI)] [late]

[JVET-AA0070](https://jvet-experts.org/doc_end_user/current_document.php?id=11746) EE2-3.2: Reconstruction-Reordered IBC for screen content coding [Z. Deng, K. Zhang, L. Zhang (Bytedance)]

[JVET-AA0202](https://jvet-experts.org/doc_end_user/current_document.php?id=11891) Crosscheck of JVET-AA0070 (EE2-3.2: Reconstruction-Reordered IBC for screen content coding) [D. Ruiz Coll, V. Warudkar (Ofinno)] [late] [miss]

[JVET-AA0224](https://jvet-experts.org/doc_end_user/current_document.php?id=11914) Crosscheck of JVET-AA0070 (EE2-3.2: Reconstruction-Reordered IBC for screen content coding) [Y. Kidani, K. Kawamura] (KDDI)] [late]

[JVET-AA0072](https://jvet-experts.org/doc_end_user/current_document.php?id=11748) EE2-2.4: ARMC with refined motion [Y. Wang, K. Zhang, N. Zhang, Z. Deng, L. Zhang (Bytedance)]

[JVET-AA0078](https://jvet-experts.org/doc_end_user/current_document.php?id=11754) EE2-1.6: Weighted chroma prediction [J.-Y. Huo, Z.-Y. Zhang, H.-Q. Du, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.), J. Ren, M. Li (OPPO)]

[JVET-AA0164](https://jvet-experts.org/doc_end_user/current_document.php?id=11852) Crosscheck of JVET-AA0078 (EE2-1.6: Weighted chroma prediction) [X. Li (Alibaba)] [late]

[JVET-AA0092](https://jvet-experts.org/doc_end_user/current_document.php?id=11768) EE2-2.5: ARMC improvements [G. Laroche, P. Onno (Canon)]

[JVET-AA0158](https://jvet-experts.org/doc_end_user/current_document.php?id=11846) Crosscheck of JVET-AA0092 (EE2-2.5: ARMC improvements) [Y. Wang (Bytedance)] [late]

[JVET-AA0093](https://jvet-experts.org/doc_end_user/current_document.php?id=11769) EE2-2.6: Combination tests of Test2.3, Test 2.4 and Test 2.5 [G. Laroche, P. Onno (Canon), Y. Wang, M. Salehifar, K. Zhang, Y. He, N. Zhang, Z. Deng, L. Zhang (Bytedance)]

[JVET-AA0213](https://jvet-experts.org/doc_end_user/current_document.php?id=11903) Crosscheck of JVET-AA0093 (EE2-2.6: Combination tests of Test 2.3, Test 2.4 and Test 2.5) [G. Li (Tencent)] [late]

[JVET-AA0095](https://jvet-experts.org/doc_end_user/current_document.php?id=11771) EE2-5: Adaptive filter shape switch and using samples before deblocking filter for adaptive loop filter [N. Hu, V. Seregin, M. Karczewicz (Qualcomm), W. Yin, K. Zhang, L. Zhang (Bytedance)]

[JVET-AA0168](https://jvet-experts.org/doc_end_user/current_document.php?id=11856) Crosscheck of JVET-AA0095 (EE2-5.2: Using sample before deblocking filter for adaptive loop filter) [W. Yin (Bytedance)] [late]

[JVET-AA0215](https://jvet-experts.org/doc_end_user/current_document.php?id=11905) Crosscheck of JVET-AA0095 (EE2-5.3: Combination of Adaptive filter shape switch for ALF/Longer filter length for ALF and Using sample before deblocking filter for adaptive loop filter) [K. Andersson (Ericsson)] [late]

[JVET-AA0096](https://jvet-experts.org/doc_end_user/current_document.php?id=11772) EE2-2.2: Motion compensated picture boundary padding [F. Le Léannec, P. Andrivon, M. Radosavljević (Xiaomi), Z. Zhang, H. Huang, C-C. Chen, Y-J. Chang, Y. Zhang, V. Seregin, M. Coban, M. Karczewicz (Qualcomm)]

[JVET-AA0206](https://jvet-experts.org/doc_end_user/current_document.php?id=11895) Crosscheck of JVET-AA0096 (EE2-2.2: Motion compensated picture boundary padding) [V. Rufitskiy, A. Filippov (Ofinno)] [late]

[JVET-AA0106](https://jvet-experts.org/doc_end_user/current_document.php?id=11782) EE2-1.7: IBC Adaptation for Camera-Captured Content [J. Xu, Y. Wang (Bytedance)]

[JVET-AA0205](https://jvet-experts.org/doc_end_user/current_document.php?id=11894) Crosscheck of JVET-AA0106 (EE2-1.7: IBC Adaptation for Camera-Captured Content) [A. Filippov, V. Rufitskiy (Ofinno)] [late]

[JVET-AA0107](https://jvet-experts.org/doc_end_user/current_document.php?id=11783) EE2-2.1: Regression based affine candidate derivation [Y. Zhang, H. Huang, V. Seregin, M. Coban, M. Karczewicz (Qualcomm)]

[JVET-AA0210](https://jvet-experts.org/doc_end_user/current_document.php?id=11900) Crosscheck of JVET-AA0107 (EE2-2.1: Regression based affine candidate derivation) [G. Li (Tencent)] [late]

[JVET-AA0116](https://jvet-experts.org/doc_end_user/current_document.php?id=11792) EE2-2.3: MMVD and Affine MMVD Extension [M. Salehifar, Y. He, K. Zhang, N. Zhang, L. Zhang (Bytedance)]

[JVET-AA0212](https://jvet-experts.org/doc_end_user/current_document.php?id=11902) Crosscheck of JVET-AA0116 (EE2-2.3: MMVD and Affine MMVD Extension) [G. Li (Tencent)] [late]

[JVET-AA0118](https://jvet-experts.org/doc_end_user/current_document.php?id=11794) EE2-1.4: Spatial GPM [F. Wang, Y. Yu, H. Yu, D. Wang (OPPO)]

[JVET-AA0151](https://jvet-experts.org/doc_end_user/current_document.php?id=11839) Crosscheck of JVET-AA0118 (EE2-1.4: Spatial GPM) [K. Naser (InterDigital)] [late]

[JVET-AA0125](https://jvet-experts.org/doc_end_user/current_document.php?id=11801) EE2-1.1b and 1.2: Filter-based linear model and gradient linear model [C.-W. Kuo, H.-J. Jhu, X. Xiu, N. Yan, W. Chen, X. Wang (Kwai)]

[JVET-AA0162](https://jvet-experts.org/doc_end_user/current_document.php?id=11850) Crosscheck of JVET-AA0125 (EE2-1.1b and 1.2: Filter-based linear model and gradient linear model) [X. Li (Alibaba)] [late]

[JVET-AA0126](https://jvet-experts.org/doc_end_user/current_document.php?id=11802) EE2-1.1c, 1.3a and 1.3b: Combined tests of EE2-1.1a, 1.1b and 1.2 [P. Astola, J. Lainema, R. G. Youvalari, A. Aminlou, K. Panusopone (Nokia), C.-W. Kuo, H.-J. Jhu, X. Xiu, N. Yan, W. Chen, X. Wang (Kwai)]

[JVET-AA0160](https://jvet-experts.org/doc_end_user/current_document.php?id=11848) Crosscheck of JVET-AA0126 (EE2-1.3a: Combined tests of EE2-1.1a and 1.2) [Z. Deng (Bytedance)] [late]

[JVET-AA0193](https://jvet-experts.org/doc_end_user/current_document.php?id=11882) Crosscheck of JVET-AA0126 (EE2-1.1c, 1.3a and 1.3b) B. Vishwanath (??) [late]

[JVET-AA0209](https://jvet-experts.org/doc_end_user/current_document.php?id=11899) Crosscheck of JVET-AA0126 (EE2 Test 1.1c and 1.3b) [G. Li (Tencent)] [late]

[JVET-AA0133](https://jvet-experts.org/doc_end_user/current_document.php?id=11809) EE2-4.1: Inter MTS optimization [B. Ray, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-AA0182](https://jvet-experts.org/doc_end_user/current_document.php?id=11870) Crosscheck of JVET-AA0133 (EE2-4.1: Inter MTS optimization) [Z. Fan, Y. Yasugi, T. Ikai (Sharp)] late]

[JVET-AA0135](https://jvet-experts.org/doc_end_user/current_document.php?id=11811) EE2-1.5: Chroma intra modes derived from collocated luma blocks and neighbouring chroma blocks [Y.-J. Chang, K. Cao, B. Ray, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-AA0152](https://jvet-experts.org/doc_end_user/current_document.php?id=11840) Crosscheck of JVET-AA0135 (EE2-1.5: Chroma intra modes derived from collocated luma blocks and neighbouring chroma blocks) [K. Naser (InterDigital)] [late]

[JVET-AA0153](https://jvet-experts.org/doc_end_user/current_document.php?id=11841) EE2-1.6b: Convolutional cross-component intra prediction model + weighted chroma prediction [P. Astola, J. Lainema (Nokia)] [late]

[JVET-AA0157](https://jvet-experts.org/doc_end_user/current_document.php?id=11845) Crosscheck of JVET-AA0153 (EE2-1.6b: Convolutional cross-component intra prediction model + weighted chroma prediction) [Y.-J. Chang (Qualcomm)] [late]

[JVET-AA0176](https://jvet-experts.org/doc_end_user/current_document.php?id=11864) EE2 2.1d: Combination test of EE2 2.1b and 2.5 [Y. Zhang, H. Huang, V. Seregin, M. Coban, M. Karczewicz (Qualcomm), G. Laroche, P. Onno (Canon)] [late]

[JVET-AA0211](https://jvet-experts.org/doc_end_user/current_document.php?id=11901) Crosscheck of JVET-AA0176 (EE2 2.1d: Combination test of EE2 2.1b and 2.5) [G. Li (Tencent)] [late]

### EE2 related contributions (10)

Contributions in this area were discussed in session 7 at 1330–1500 UTC on Thursday 14 July 2022 and in session 11 at 1300–1405 UTC on Friday 15 July 2022 (chaired by JRO).

[JVET-AA0045](https://jvet-experts.org/doc_end_user/current_document.php?id=11720) [EE2-1.4 related] Reduced Complexity Spatial GPM [K. Naser, Y. Chen, A. Robert, K. Reuzé (InterDigital)]

This contribution is related to spatial geometrical partition mode (SGPM) that is currently being studied in EE2. It is proposed to reduce the number of tested combinations of split and intra modes by selecting only two intra modes from TIMD process. Furthermore, reduced signaled SGPM modes and reduced block sizes are tested to achieve a reasonable compromise between the gain and coding complexity.

The following results are obtained:

Test1: #intraModes = 2 + reduced SGPM modes + reduced block sizes

* + AI: -0.07%, -0.06%, -0.05% with 105%% EncT and 100%DecT

Test2: #intraModes = 2 + reduced block sizes

* + AI: -0.18%, -0.14%, -0.14% with 120%% EncT and 106%DecT

The EE2 proposal had -0.2% in AI, with encoding time increase of 23%. Test2 is not much different from that.

Related: JVET-AA0119 and JVET-AA0149.

See further notes under JVET-AA0149

[JVET-AA0245](https://jvet-experts.org/doc_end_user/current_document.php?id=11935) Cross-check of JVET-AA0045 (EE2-1.4 related: Reduced Complexity Spatial GPM) [C. Bonnineau, M. Abdoli (IRT b-com)] [late]

[JVET-AA0103](https://jvet-experts.org/doc_end_user/current_document.php?id=11779) EE2-related: CCLM with non-linear term [X. Li, Y. Ye, R.-L. Liao, J. Chen (Alibaba)]

In this contribution, a non-linear term of the down-sampled reconstructed luma value is introduced into the CCLM method. A chroma sample is predicted based on a linear term and a non-linear term of the reconstructed luma sample value, where the model parameters are derived from two lines left and above adjacent chroma samples and the corresponding down-sampled luma samples based on the LDL decomposition method used in EE2 Test 1.1a. It is reported that on top of ECM-5.0, when the CCLM\_LT mode is replaced by the proposed method, the overall coding performance impact for {Y, U, V, EncT, DecT} is {-0.12%, -0.17%, -0.37%, 101%, 101%} for AI and {-0.17%, -0.15%, -0.75%, 101%, 101%} for RA.

Gain comes mainly from class A1, and there from Campfire and Tango.

Would there still be gain after adoption of CCCM? It was answered that CCCM has its largest benefit in other sequences. On the other hand, this might lead to a situation where the gain on sequences other than Campfire and Tango is even lower, which might lead to an assessment that such a proposal would not be adopted.

It was pointed out that conceptually, the proposed is using a similar nonlinear mapping as CCCM, could be interpreted as a subset. However, the proposal is to use it instead of the existing CCLM\_LT when slope adjustment is equal to zero, so no additional signalling is required.

Has it been tried to signal as an additional mode? No – this is recommended to be tested. It is also recommended to test with identical parameter derivation as used in CCCM.

Investigate in EE.

[JVET-AA0231](https://jvet-experts.org/doc_end_user/current_document.php?id=11921) Crosscheck of JVET-AA0103 (EE2-related: CCLM with non-linear term [Z. Xie (OPPO)] [late]

[JVET-AA0114](https://jvet-experts.org/doc_end_user/current_document.php?id=11790) [EE2-related] Division-free operation and mean-compensation for convolutional cross-component model (CCCM) [A. Aminlou, J. Lainema, P. Astola (Nokia)]

This contribution proposes two modifications to convolutional cross-component model (CCCM) based chroma prediction with an intention to simplify the implementation. The first modification removes the mean values of luma and chroma samples. Mean values are calculated using the neighboring samples which are used to derive the model. This is asserted to reduce the dynamic range of data and the bitdepth required for model derivation. The second modification replaces the division operation in CCCM by a piece-wise polynomial (power of 2) function.

The impact on coding efficiency and runtimes over ECM-5.0 is reportedly {for Y, U, V, EncT, DecT}:

AI { -1.38%, -2.83%, -2.81%, 102%, 99%} RA { -0.79%, -2.09%, -2.47%, 102%, 94%}

The impact on coding efficiency and runtimes over the CCCM implementation of EE2 Test 1.1a is reportedly {for Y, U, V, EncT, DecT}:

AI { 0.01%, -0.02%, 0.02%, 100%, 98%} RA { 0.00%, 0.01%, -0.05%, 101%, 95%}

It is commented that the first modification (mean removal) is undesirable as it introduces an additional pipeline stage which cannot be considered as a simplification. Further study might be desirable on the division simplification. However in general, implementation optimization is not of primary importance at this stage of exploration activity.

[JVET-AA0175](https://jvet-experts.org/doc_end_user/current_document.php?id=11863) Crosscheck of JVET-AA0114 ([EE2-related] Division-free operation and mean-compensation for convolutional cross-component model (CCCM)) [Y.-J. Chang (Qualcomm)] [late]

[JVET-AA0119](https://jvet-experts.org/doc_end_user/current_document.php?id=11795) EE2-1.4a-related: Modifications of Spatial GPM [F. Wang, Y. Yu, H. Yu, D. Wang (OPPO)]

This contribution proposes two modifications to the SGPM method presented in the EE2-1.4a test. One modification is to extend SGPM to smaller blocks, and the other is to apply an adaptive SGPM blending in the transition region between the two partitions. It is reported that with these modifications, on top of ECM-5.0, the simulation results are as follows:

AI {-0.22% Y, -0.17% U, -0.19% V, 113% EncT, 103% DecT}

RA {-0.12% Y, -0.06% U, -0.18% V, x% EncT, x% DecT}

Adaptive blending is adapted based on block size, no signalling.

The encoding time decrease is achieved by reducing number of RDO decisions (number also dependent on block size), while the compression gain of EE is retained by introducing new elements. There is also an increase in the decoder run time.

Additional block sizes (without reducing RDO) would give 0.1% on top of the EE2 method. Not known how much gain comes from the adaptive blending. It is noted that in the inter case, making the adaptive blending dependent on block size did not have an effect. This may different in intra.

What would be the gain by just reducing RDO without additional elements? Not known.

See further notes under JVET-AA0149

[JVET-AA0226](https://jvet-experts.org/doc_end_user/current_document.php?id=11916) Crosscheck of JVET-AA0119 (EE2-1.4a-related: Modifications of Spatial GPM) [H.-J. Jhu, X. Xiu (Kwai)] [late]

[JVET-AA0127](https://jvet-experts.org/doc_end_user/current_document.php?id=11803) EE2-1 related: Encoder optimization of EE2-1.2 and 1.3 [C.-W. Kuo, H.-J. Jhu, X. Xiu, N. Yan, W. Chen, X. Wang (Kwai)]

(include abstract)

Was discussed in context of EE2 – see notes there.

[JVET-AA0161](https://jvet-experts.org/doc_end_user/current_document.php?id=11849) Crosscheck of JVET-AA0127 (EE2-1 related: Encoder optimization of EE2-1.2 and 1.3) [Z. Deng (Bytedance)] [late]

[JVET-AA0163](https://jvet-experts.org/doc_end_user/current_document.php?id=11851) Crosscheck of JVET-AA0127 (EE2-1 related: Encoder optimization of EE2-1.2 and 1.3) [X. Li (Alibaba)] [late]

[JVET-AA0128](https://jvet-experts.org/doc_end_user/current_document.php?id=11804) EE2-related: On regression based affine candidate derivation [W. Chen, X. Xiu, C.-W. Kuo, H.-J. Jhu, N. Yan, X. Wang (Kwai)]

This contribution proposes two modifications to the EE2 Test-2.1 on regression based affine candidate derivation:

* Remove the redundant linear regression process for candidate 1 derivation
* Increase the number of CABAC contexts of affine merge index from 1 to 3.

Several subtests are performed on top of the EE2 test 2.1b:

* Test a: the removal of the redundant linear regression process for candidate 1

RA: {-0.xx%, -0.xx %, -0.xx %, xxx%, xxx%}; LDB: {-0.xx%, -0.xx %, -0.xx %, xxx%, xxx%};

* Test b: the increased contexts for affine merge index

RA: {-0.xx%, -0.xx %, -0.xx %, xxx%, xxx%}; LDB: {-0.xx%, -0.xx %, -0.xx %, xxx%, xxx%}.

* Test c: the combination of Test a and b

RA: {-0.xx%, -0.xx %, -0.xx %, xxx%, xxx%}; LDB: {-0.xx%, -0.xx %, -0.xx %, xxx%, xxx%}.

No full results yet.

Test a: According to opinion of the cross-checker (which was not completed yet), the statement about the regression being redundant is not entirely justified. This also seems at a level of small-scale optimization which is not that relevant at this moment; neither runtime nor compression performance are significantly changed.

Test b: The increased number of CABAC contexts seems to give some gain (0.06% in RA, 0.08% in LB) from preliminary results. Cross-check not complete yet.

Was again reviewed in session 24 after full results and crosscheck were available. Test b had initially been considered as potential candidate for adoption on top of the adopted EE proposal. It was also reported that gain is achieved in a variety of sequences.

RA: {-0.06%, -0.08 %, -0.03 %, 100%, 102%}; LDB: {-0.08%, 0.08 %, -0.19 %, 100%, 102%}

Decision: Adopt JVET-AA0128 test b (increased number of CABAC contexts only)

[JVET-AA0240](https://jvet-experts.org/doc_end_user/current_document.php?id=11930) Crosscheck of JVET-AA0128 (EE2-related: on regression based affine candidate derivation) [Y. Zhang (Qualcomm)] [late]

[JVET-AA0138](https://jvet-experts.org/doc_end_user/current_document.php?id=11814) EE2-related: On Gradient Linear Model (GLM) [X. Li, Y. Ye, R.-L. Liao, J. Chen (Alibaba)]

In this contribution, a new GLM mode is proposed where the chroma sample is predicted based on both the gradient of luma samples and the reconstructed value of the down-sampled luma sample with different parameters. The model parameters are derived by the same adjacent samples as the CCLM modes in ECM-5.0 based on the LDL decomposition method used in EE2 Test 1.1. Multiple variants of the proposed method with different coding gain and complexity trade-offs are tested on top of ECM-5.0. The overall coding performance impact for {Y, U, V, EncT, DecT} are:

Test 1: two GLM modes with 16 gradient patterns

Class A to E:{-0.90%, -0.68%, -0.93%, 117%, 99%} AI, {x.xx%, x.xx%, x.xx%, xxx%, xxx%} RA;

Class F: {-0.75%, -4.28%, -4.27%, 108%, 100%} AI, {-1.60%, -4.14%, -4.14%, 106%, 101%} RA;

Class TGM: {-4.24%, -13.07%, -12.31%, 109%, 97%} AI, {-2.81%, -6.88%, -6.44%, 107%, 101%} RA;

Test 2: one GLM mode with 16 gradient patterns

Class A to E:{-0.88%, -0.58%, -0.87%, 111%, 99%} AI, {x.xx%, x.xx%, x.xx%, xxx%, xxx%} RA;

Class F: {-0.81%, -3.84%, -3.75%, 105%, 100%} AI, {-1.59%, -3.84%, -4.08%, 104%, 101%} RA;

Class TGM: {-4.21%, -12.06%, -11.32%, 106%, 98%} AI, {-2.77%, -6.37%, -6.07%, 104%, 100%} RA;

Test 3: two GLM modes with 4 gradient patterns

Class A to E:{-0.79%, -0.75%, -0.96%, 107%, 100%} AI, {x.xx%, x.xx%, x.xx%, xxx%, xxx%} RA;

Class F: {-0.81%, -4.58%, -4.37%, 103%, 100%} AI, {-1.62%, -4.08%, -4.13%, 102%, 101%} RA;

Class TGM: {-4.36%, -13.22%, -12.28%, 102%, 98%} AI, {-2.90%, -6.99%, -6.54%, 102%, 100%} RA;

Test 4: one GLM mode with 4 gradient patterns

Class A to E:{-0.78%, -0.68%, -0.90%, 105%, 100%} AI, {x.xx%, x.xx%, x.xx%, xxx%, xxx%} RA;

Class F: {-0.81%, -4.09%, -3.86%, 102%, 100%} AI, {-1.67%, -4.02%, -4.16%, 101%, 101%} RA;

Class TGM: {-4.35%, -12.11%, -11.32%, 102%, 98%} AI, {x.xx%, x.xx%, x.xx%, xxx%, xxx%} RA;

The proposal EE1.2 was mainly adopted due to its additional benefit for screen content. The most simplified version (test 4) seems to further improve gain in classes F and TGM, while reducing encoding time. It is however noted that the encoder speedup of the adopted EE method was not used in the contribution. According to the opinion of the cross-checker, the new mode 2 is not a simplification due to increased number of parameters, but the additional gain is asserted as interesting.

Investigate in EE.

[JVET-AA0173](https://jvet-experts.org/doc_end_user/current_document.php?id=11861) Crosscheck of JVET-AA0138 (EE2-related: On Gradient Linear Model (GLM)) [C.-W. Kuo (Kwai)] [late]

[JVET-AA0140](https://jvet-experts.org/doc_end_user/current_document.php?id=11816) EE2-related: Self-Aware Filter Estimation for CCLM [K. Zhang, Z. Deng, L. Zhang (Bytedance)]

This contribution presents a self-aware filter estimation CCLM (SAFE-CCLM) method to ECM. With SAFE-CCLM, N candidate luma down-sampling filters are predefined. When SAFE-CCLM is applied, a linear model between luma and chroma component is derived in the same way as that in ECM for each candidate filter first. Second, prediction values are calculated with each linear model in a testing region including one-column left neighbouring samples and one-row above neighbouring samples. Third, a SAD cost between the reconstructed samples and their corresponding prediction values in the testing region is computed for each filter candidate. Finally, the filter candidate with the least SAD cost is selected as the down-sampling filter to perform the CCLM prediction for the current block. When CCLM is indicated to be used for a block, a SAFE-CCLM flag is signaled to further indicate whether SAFE-CCLM is applied. On top of ECM-5.0, simulation results of the proposed method are reported as below:

AI: {-0.80%, -0.03%, -0.23%, 103%, 103%}; RA:{-0.38%, 0.03%, -0.13%, 100%, 100%}.

Simulation results on TGM category are reported as below:

AI: {-3.31%, -6.73%, -6.45%, 102%, 104%}; RA:{-2.11%, -3.40%, -3.23%, 102%, 100%}.

Not known how much gain is retained after the adoptions from EE2-1.3a\*. It is likely that there is a large overlap. Investigate in EE.

[JVET-AA0221](https://jvet-experts.org/doc_end_user/current_document.php?id=11911) Crosscheck of JVET-AA0140 (EE2-related: Self-Aware Filter Estimation for CCLM) [X. Li (Alibaba)] [late] [miss]

[JVET-AA0147](https://jvet-experts.org/doc_end_user/current_document.php?id=11823) EE2-Related: Extended Offline-Filtering Taps for ALF [W. Yin, K. Zhang, Z. Deng, L. Zhang (Bytedance)]

In the current adaptive loop filter (ALF) design, online-trained filters contain 2 kinds of taps: spatial taps and offline-filtered taps. In ECM-5.0, an online-trained filter contains 2 offline-filtered taps. In this contribution, the number of offline-filtered taps is extended to 8.

On top of ECM-5.0, simulation results are reported as below:

AI: -0.10%, 0.01%, 0.01%, 102%, 103%.

RA: -0.14%, 0.07%, -0.01%, 101%, 102%.

LB: -0.20%, 0.00%, -0.02%, 102%, 103%.

When combining with EE2-5.1b, simulation results compared with ECM-5.0 are reported as below:

AI: -0.13%, 0.02%, 0.01%, 103%, 102%.

RA: -0.22%, -0.02%, -0.07%, 101%, 102%.

LB: -0.41%, -0.16%, -0.05%, 103%, 103%.

The proposal is filtering the output of the fixed filter (“offline filter”) again. This might likely require additional line buffering, in addition to the additional computations.

Not known if the gain is retained in combination with EE2-5.2 which was adopted.

Investigate in EE

[JVET-AA0149](https://jvet-experts.org/doc_end_user/current_document.php?id=11825) EE2-1.4 related: Improvements on Spatial GPM [A. Natesan, J. N. Shingala, J. R. Arumugam, V. Valvaiker (Ittiam), T. Lu, P. Yin, F. Pu, T. Shao, A. Arora, S. McCarthy (Dolby)]

This contribution proposes modifications on EE2 Test 1.4 for spatial GPM, aiming at simplifying the algorithm and improving the coding efficiency. There are three aspects: 1) Introduce a fast GPM partitioning list derivation algorithm to adaptively determine the GPM partitioning mode and to reduce the number of GPM partitions to check for template-based partition type and intra mode selection; 2) Use IPM list as an alternative to the MPM list used in Test 1.4 and exclude TIMD in SGPM intra mode; 3) Reduce the template size used in Test 1.4 from 4 to 1.

The impact on coding efficiency and runtimes over ECM-5.0 is reportedly {for Y, U, V, EncT, DecT}:

AI: { -0.21%, -0.11%, -0.06%, 113%, 106%}

RA: {-0.12%, -0.07%, -0.05%, 104%, 100%}

Setting 2 with additional simplification (optimized candidate partitioning types for testing, excluded usage of TIMD mode etc.) is tested and the results over ECM-5.0 are

AI: { xx%, xx%, xx%}

RA: { xx%, xx%, xx%}

It is commented by the cross-checker that reduction of template size may not be useful.

Encoding runtime is decreased by reducing the number of GPM partition types and intra prediction modes (normatively)

Investigate JVET-AA0045, JVET-AA0119, JVET-AA0149 in EE, in particular identify the benefit of

* Strategies for reducing RDO decisions (also for previous test 1.4)
* Reducing the number of GPM partition types and intra mode derivation
* Smaller and bigger block sizes
* Adaptive blending

[JVET-AA0230](https://jvet-experts.org/doc_end_user/current_document.php?id=11920) crosscheck of JVET-AA0149 (EE2-1.4 related: Improvements on Spatial GPM) [K. Naser (InterDigital)] [late] [miss]

[JVET-AA0248](https://jvet-experts.org/doc_end_user/current_document.php?id=11939) Crosscheck of JVET-AA0149 (EE2-1.4 related: Improvements on Spatial GPM) [F. Wang (OPPO)] [late]

### ECM modifications beyond EE2 (27)

Contributions in this area were discussed in session 11 at 1410–1510 UTC on Friday 15 July 2022, in session 14 at 1530–1700 UTC on Monday 18 July 2022 (chaired by JRO), in session 17 1415-1530 on Tuesday 19 July 2022, in session 18 1550-1725 on Tuesday 19 July 2022 (chaired by Y. Ye), and in session 22 at 1520–1730 UTC on Wednesday 20 July 2022 (chaired by JRO).

[JVET-AA0043](https://jvet-experts.org/doc_end_user/current_document.php?id=11718) IntraTMP Adaptation for Camera Captured Contents [K. Naser, T. Poirier, F. Galpin, A. Robert (InterDigital)]

This contribution proposes enabling IntraTMP for all CTC classes. To well balance the bitrate gain and the coding time, certain adaptations of IntraTMP are proposed. Specifically, it is proposed to speed-up the template matching process by subsampling the search range and perform iterative refinement. Two variants are studied:

Variant 1:

AI: -0.23% -0.30% -0.26% with 103% EncT and 103% DecT

RA: xx% xx% xx% with xx% EncT and xx% DecT

Variant 2:

-0.21% -0.22% -0.26% with 103% EncT and 102% DecT

RA: xx% xx% xx% with xx% EncT and xx% DecT

Subsampling 0 (every pixel position) would be identical with the current IntraTMP. In that case, no refinement step would be applied.

Investigate in EE. It should also be investigated how the subsampling approach would work for screen content.

It was commented that it would not be a simplification if the decoder would need to implement both downsampled and not downsampled versions.

It was also commented that the gain is almost zero for class A1. There seems to be sequence dependency.

[JVET-AA0204](https://jvet-experts.org/doc_end_user/current_document.php?id=11893) Crosscheck of JVET-AA0043 (IntraTMP Adaptation for Camera Captured Contents) [D. Ruiz Coll, V. Warudkar (Ofinno)] [late]

[JVET-AA0044](https://jvet-experts.org/doc_end_user/current_document.php?id=11719) IntraTMP for chroma Components [K. Naser, T. Dumas, T. Poirier, F. Galpin (InterDigital)]

This contribution proposes enabling IntraTMP for chroma components when chroma separated tree is activated in intra slice. Specifically, a CU level flag is signaled to indicate if IntraTMP is applied for the two chroma components. On top of ECM-5.0, the following results are obtained (classF and TGM):

AI: -0.17%%, -0.15%%, -0.20%% for Y, U and V respectively with 101% encoding time and 101% decoding time.

RA: -0.11%%, -0.20%%, -0.16%% for Y, U and V respectively with 98% encoding time and 101% decoding time.

Practically no gain for class F, only for TGM. For the case of screen content, the gain appears relatively small

Search is performed independently for luma and chroma. This is necessary due to the dual-tree coding. It was pointed out that in single-tree case it would be better to re-use the luma vector for chroma.

Investigate in EE (SC category) along with JVET-AA0053.

[JVET-AA0171](https://jvet-experts.org/doc_end_user/current_document.php?id=11859) Crosscheck of JVET-AA0044 (IntraTMP for chroma Components) [W. Lim, S.-C Lim (ETRI)] [late]

[JVET-AA0053](https://jvet-experts.org/doc_end_user/current_document.php?id=11729) AHG12: Using block vector derived from IntraTMP for IBC [W. Lim, D. Kim, S.-C. Lim, J. S. Choi (ETRI)]

This contribution proposes to store and use block vector (BV) derived from intra template matching prediction (IntraTMP) for intra block copy (IBC). IntraTMP finds the best matched template position in the reconstructed area of the current picture and uses the corresponding block as the prediction block of the current block. However, there is no process for storing block vector of the IntraTMP in ECM-5.0. In this contribution, IntraTMP block vectors are stored and added as spatial candidates to block vector candidate list of IBC in order to improve the coding efficiency. Compared to ECM-5.0, it is reported that the proposed method shows average BD-rates of -0.13% and -0.17% for class F and TGM under AI case and average BD-rates of -0.06% and -0.04% for class F and TGM under RA case with negligible complexity impact. And, it is suggested to include the proposed method into the next version of ECM.

It is proposed to use a BV from a neighboring block that was coded in IntraTMP mode in the candidate list of IBC. Even though the gain is relatively small (considering gains for screen content are typically larger), this does not seem to introduce additional dependency compared to current ECM which already uses refined IBC BV in the list derivation.

Gain is homogeneous for classes F and TGM.

Investigate in EE in SC category.

It is commented that it would be interesting to investigate what is the current gain of IntraTMP in the screen content classes, considering the fact that many more SC tools have been added since its adoption.

[JVET-AA0155](https://jvet-experts.org/doc_end_user/current_document.php?id=11843) Cross-check of JVET-A0053 "AHG12: Using block vector derived from IntraTMP for IBC" [F. Le Léannec (Xiaomi)] [late]

[JVET-AA0064](https://jvet-experts.org/doc_end_user/current_document.php?id=11740) AHG12: A study on non-separable primary transform [J. Choi, M. Koo, J. Lim, J. Zhao, S. Kim (LGE)]

In the current ECM, multiple types of transforms are available for primary transform selection. All the available primary transforms are separable transforms. In general, the structural restriction of separable transforms could be limited to efficiently represent the highly dynamic image statistics from textures, such as directed edges in natural image and video sequences. Therefore, this contribution is to provide an initial progress of our investigation on non-separable primary transform (NSPT). Specifically, current NSPT design includes 35 or 67 transform sets depends on TB size and intra prediction modes, and only one NSPT kernel is applied per transform set. At this stage, the same test sequence is applied for training and testing. From the experimental results based on ECM4.0, 3.43% coding gain has been observed in AI configuration.

Only for luma, not is combination with LFNST.

No results on RA, but probably would also have benefit.

Decoder time (102%) is unreliable. Encoder (122%) is more realistic.

Why higher gain in class E? It was answered that the class E is better predictable.

How much less performance could be expected when training with data outside the test set?

Why using more transform sets (67 instead 35) for non-square blocks? Better performance

Currently applied to block sizes up to 16x16. It was pointed out that restricting to 8x8 might be interesting, as this would significantly reduce worst case complexity in terms of computation and storage.

Coefficients are 8 bit. Total memory required for storage?

Proponents think it would be premature to investigate in EE – first necessary to perform training outside of test set.

[JVET-AA0069](https://jvet-experts.org/doc_end_user/current_document.php?id=11745) Non-EE2: AmvpMerge for low delay [H. Jang, J. Nam, N. Park, J. Lim, S. Kim (LGE)]

In the ECM-5.0, the amvpMerge is enabled only when a reference picture pair includes both a forward picture and a backward picture because amvpMerge is based on bilateral matching. Considering that amvpMerge could be effective even for the low-delay case without bilateral matching-based reordering and refinement, this contribution suggests enabling the amvpMerge for low delay case. In order to enable amvpMerge for low delay case this contribution proposes three aspects below:

1. The amvpMerge is enabled even the reference picture pair is not true bi-directional pictures for low-delay case. Additionally the restriction condition on the resampled, long-term and WP reference picture are relaxed for low-delay picture cases.
2. BM cost of an amvpMerge candidate which is not true bi-direction is set to be the maximum value during merge candidate reordering.
3. The decoder side motion vector refinement process of the conventional amvpMerge is bypassed for the low delay case.

With combination of above mentioned aspects, following two alternative methods are tested to support the amvpMerge in low delay case.

* Method 1 : aspect 1 + aspect 2
* Method 2 : aspect 1 + aspect 2 + aspect 3

The proposed test also includes !MR184 to prevent an empty merge candidate list for amvpMerge.

The simulation result of Method1 shows {%d, %d, %d / Enc%d, Dec%d} and Method2 shows {%d, %d, %d / Enc%d, Dec%d} for LDB configuration. The negligible performance impact on RA is observed for both proposed methods.

Results not complete yet. Partial results (classes C, D, E) show 0.3%-0.5% bit rate reduction. Both methods almost identical. Method 2 would be preferable for LD cases, according to opinion of proponents and crosscheckers, removing decoder side MV refinement.

Proposal is straightforward and clean implementation is available, according to crosschecker’s opinion could be included in ECM.

Investigate in EE.

[JVET-AA0232](https://jvet-experts.org/doc_end_user/current_document.php?id=11922) Crosscheck of JVET-AA0069 (Non-EE2: AmvpMerge for low delay) [Z. Zhang (Qualcomm)] [late]

[JVET-AA0073](https://jvet-experts.org/doc_end_user/current_document.php?id=11749) Non-EE2: Modification of LFNST for MIP coded block [J.-Y. Huo, W.-H. Qiao, X. Hao, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.), J. Ren, M. Li (OPPO)]

In ECM5.0, for a MIP coded block, LFNST transform set 0 is used and LFNST transpose flag is equal to 0. In this contribution, modifications of LFNST for MIP coded blocks are proposed. It is proposed to use DIMD to derive a traditional intra predition mode based on the prediction samples before interpolation of MIP modes, then the LFNST transform set and LFNST transpose flag are both determined by the derived intra mode. Moreover, LFNST is enable for MIP coded blocks of width and height greater than or equal to 4. It is reported that the coding performance to ECM 5.0 are as below:

For AI configuration: -0.11%, -0.08%, and -0.12%, with 105% EncT, 100% DecT.

For RA configuration: -0.07%, -0.06%, and -0.12%, with xxx% EncT, xxx% DecT.

Follow-up of JVET-Z0048. Complexity is decreased relative to previous proposal by performing the mode derivation before upsampling.

Crosschecker reports matching of results. He points out that the modification allows parallel processing DIMD and upsampling. Also the changes give additional gain, without modifying the encoder decision relative to the previous proposal.

Appropriate modification of previous proposal.

Investigate in EE.

[JVET-AA0243](https://jvet-experts.org/doc_end_user/current_document.php?id=11933) Crosscheck of JVET-AA0073 (Non-EE2: Modification of LFNST for MIP coded block) [[X. Li (Alibaba)](mailto:sid.lxw@alibaba-inc.com)] [late]

[JVET-AA0075](https://jvet-experts.org/doc_end_user/current_document.php?id=11751) Non-EE2: Template matching based BCW index derivation for merge mode [R.-L. Liao, J. Chen, Y. Ye, X. Li (Alibaba)]

This contribution proposes to derive BCW index for merge coded CUs based on template matching cost instead of inferring from neighboring blocks. Given a selected merge candidate, the TM cost values are calculated with different BCW indices setting, and then, the BCW index with minimum TM cost value is used to predict the merge CU. In addition, it is proposed to remove negative weight and use three weights {3/8, 4/8, 5/8} for both low delay and non-low delay pictures. It is reported that on top of ECM-5.0, the overall coding performance impact for {Y, U, V, EncT, DecT} is {-0.10%, -0.15%, -0.18%, 100%, 100%} for RA and {-0.13%, -0.12%, -0.18%, 101%, 101%} for LDB.

Template matching is added both at decoder and encoder, otherwise the reduction of weights would not give gain.

JVET-AA0134 also relates to a BCW modification (which so far was not changed relative to VVC). It was also reported that BCW gives much less gain in ECM than it was the case in VTM. See further notes under JVET-AA0134

Investigate in EE

[JVET-AA0187](https://jvet-experts.org/doc_end_user/current_document.php?id=11875) Crosscheck of JVET-AA0075 (Non-EE2: Template matching based BCW index derivation for merge mode) [D. Kim, W. Lim, S.-C. Lim (ETRI)] [late]

[JVET-AA0196](https://jvet-experts.org/doc_end_user/current_document.php?id=11885) Crosscheck of JVET-AA0075 (Non-EE2: Template matching based BCW index derivation for merge mode) [B. Vishwanath (Bytedance)] [late]

[JVET-AA0233](https://jvet-experts.org/doc_end_user/current_document.php?id=11923) Crosscheck of JVET-AA0075 (Non-EE2: Template matching based BCW index derivation for merge mode) [Z. Zhang (Qualcomm)] [late]

[JVET-AA0097](https://jvet-experts.org/doc_end_user/current_document.php?id=11773) ECM fix for block-level out-of-bound checking [F. Le Léannec, P. Andrivon, M. Radosavljević (Xiaomi)]

This contribution proposes to fix the ECM software implementation, with regards to the process that checks if a reference block of a bi-predicted inter block is out-of-bounds or not.

Indeed, it is stated the source code implementation of these checks does not exactly correspond to the description of the adopted methods described in JVET-Z0136 and JVET-Y0125. The ECM-5.0 implementation leads to a higher number of blocks detected as OOB than it should be during the bi-prediction process, leading to per-sample OOB checks for a larger number of blocks than is needed.

A fix is proposed in this contribution. It provides reduced runtime with no change in ECM CTC results.

Confirmed by original contributors who performed the cross-check.

Results are unchanged because in the previous SW implementation a wrong OOB classification at block level was corrected by subsequent sample-level change (which however would not have been exectuted if the block-level check would have correct).

Is the sample-by-sample OOB check still needed with this fix? Yes.

Decision (BF/SW): Adopt JVET-AA0097

[JVET-AA0167](https://jvet-experts.org/doc_end_user/current_document.php?id=11855) Crosscheck of JVET-AA0097: ECM fix for block-level out-of-bound checking [X. Xiu (Kwai)] [late]

[JVET-AA0098](https://jvet-experts.org/doc_end_user/current_document.php?id=11774) AHG12: Encoder configuration proposal to reduce worst case encoding time [F. Le Léannec, P. Andrivon, M. Radosavljević (Xiaomi)]

See section 4.3

[JVET-AA0104](https://jvet-experts.org/doc_end_user/current_document.php?id=11780) Non-EE2: On planar horizontal mode and planar vertical mode [X. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba)]

In this contribution, two new additional planar modes are proposed where only the horizontal interpolation or only the vertical interpolation are used to obtain the predicted samples. It is reported that on top of ECM-5.0, the overall coding performance impact for {Y, U, V, EncT, DecT} is {-0.08%, -0.01%, 0.01%, 117%, 100%} for AI and {-0.03%, 0.03%, -0.03%, 103%, 100%} for RA.

Only for non-ISP blocks

Signalling about the three planar modes is done separately via an additional flag, followed by an index.

Could there be too much signalling overhead? How often are the new modes used? Similar frequently as the existing one.

Encoder runtime is critical. With the current tradeoff, the proposal would not be attractive for adoption.

Investigate in EE.

[JVET-AA0184](https://jvet-experts.org/doc_end_user/current_document.php?id=11872) Crosscheck of JVET-AA0104 Non-EE2: On planar horizontal mode and planar vertical mode [W. Jia (Bytedance)] late]

Results match, simple decoder modification.

[JVET-AA0108](https://jvet-experts.org/doc_end_user/current_document.php?id=11784) AHG12: IBC AMVP candidates clustering [D. Ruiz Coll, V. Warudkar (Ofinno)] [late][withdrawn]

This contribution was withdrawn

[JVET-AA0120](https://jvet-experts.org/doc_end_user/current_document.php?id=11796) Non-EE2: Template-based multiple reference line intra prediction [L. Xu, Y. Yu, H. Yu, D. Wang (OPPO)]

This contribution proposes a template-based multiple reference line intra prediction (TMRL) mode. The proposed mode combines reference line and prediction mode together and uses a template matching method to construct a list of candidate combinations. An index to the candidate combination list is coded to indicate which reference line and prediction mode is used in coding the current block. In this proposal, the regular multiple reference line (MRL) for the non-TIMD part is replaced by the proposed mode.

Compared with ECM5.0 anchor, the proposed method shows coding performance gains as follows.

Test-a:

AI -0.08%/-0.01%/-0.04%, Enc: 102%, Dec: 105%,

RA -0.03%/-0.08%/-0.13%, Enc: 101%, Dec: 101%,

Test-b:

AI -0.13%/-0.07%/-0.05%, Enc: 105%, Dec: 104%,

RA -0.05%/0.02%/-0.05%, Enc: 101%, Dec: 100%.

The proposed modification replaces the MRL in ECM-5.0 when TIMD is not selected, and when MRL is selected. This contribution is related to JVET-AA0137 and JVET-AA0246.

See notes under JVET-AA0246.

Investigate in EE.

[JVET-AA0166](https://jvet-experts.org/doc_end_user/current_document.php?id=11854) Crosscheck of JVET-AA0120 (Non-EE2: Template-based multiple reference line intra prediction) [X. Li (Alibaba)] [late]

[JVET-AA0228](https://jvet-experts.org/doc_end_user/current_document.php?id=11918) Crosscheck of JVET-AA0120 (Non-EE2: Template-based multiple reference line intra prediction) [K. Cao (Qualcomm)] [late]

[JVET-AA0121](https://jvet-experts.org/doc_end_user/current_document.php?id=11797) Non-EE2: Template-based MIP [Z. Xie, Y. Yu, H. Yu, D. Wang, Y. Liu, M. Li (OPPO), J. Huo, W. Qiao, X. Hao, Y. Ma, F. Yang (Xidian University)]

This contribution proposes a template-based MIP method to improve the coding efficiency of ECM. It is proposed to use the top and left templates of the current coding block to derive the MIP mode and the transpose flag. Moreover, DIMD is used to derive a traditional intra-prediction mode based on the MIP prediction samples, then both the LFNST transform set and LFNST transpose flag are determined by the derived traditional intra-prediction mode. Compared with the ECM-5.0 anchor, it is reported that the simulations on the proposed method showed the following coding performance results for {Y, U, V, EncT, DecT}:

AI: -0.17%, -0.06%, -0.08%; 116%, 101%

RA: -0.11%, -0.01%, -0.20%; 103%, 100%

Built on top of JVET-AA0073, extending by an additional template matching. Compared to the other proposal, not a good tradeoff compression vs. encoding time. Further study to decrease encoding time is recommended.

[JVET-AA0165](https://jvet-experts.org/doc_end_user/current_document.php?id=11853) Crosscheck of JVET-AA0121 (Non-EE2: Template-based MIP) [X. Li (Alibaba)] [late]

[JVET-AA0124](https://jvet-experts.org/doc_end_user/current_document.php?id=11800) Non-EE2: Enable amvpMerge mode on scaled reference pictures when DMVD is disabled [Z. Zhang, H. Huang, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes to enable AmvpMerge mode when DMVD is disabled and RPR is enabled. The proposed method was implemented on top of ECM-5.0 and tested with EE2 RPR configuration settings 1.5x and 2.0x with DMVD additionally disabled in both anchor and test. The test results are summarised as follows:

RA 1.5x: -0.14% (Y), -0.19% (U), -0.22% (V), 101% (EncT), 100% (DecT)

RA 2.0x: -0.10% (Y), -0.15% (U), -0.09% (V), 101% (EncT), 100% (DecT)

In ECM-5.0, when RPR is enabled and DMVD is disabled, AmvpMerge mode is disabled.

It was commented that RPR pictures as well as weighted prediction pictures and long term reference pictures are currently not in the reference picture list for AmvpMerge in ECM-5.0.

This contribution proposes to allow RPR pictures to be used for AmvpMerge mode, without bilateral matching. This brings some coding performance gain in RPR without impact on runtime, and does not affect CTC.

In the April meeting, it was decided to enable AmvpMerge when DMVD is disabled. This contribution extends the decision to the case when RPR is enabled, which seems straightforward.

Decision (cleanup): adopt JVET-A0124 (merge request 162)

[JVET-AA0236](https://jvet-experts.org/doc_end_user/current_document.php?id=11926) Crosscheck of JVET-AA0124 (Non-EE2: Enable amvpMerge mode on scaled reference pictures when DMVD is disabled) [H. Jang (LGE)] [late]

[JVET-AA0129](https://jvet-experts.org/doc_end_user/current_document.php?id=11805) Non-EE2/AHG10: Improved inter hash RDO considering OBMC off in ECM5.0 [X. Xiu, C.-W. Kuo, H.-J. Jhu, W. Chen, N. Yan, X. Wang (Kwai)]

This document presents one encoder optimization for inter-hash-based motion estimation by enabling the R-D cost calculation of disabling OBMC at CU-level. For TGM sequences, the proposed encoder improvement reportedly shows 0.99% and 1.45% BD-rate reduction over ECM5.0 anchor for RA and LDB configurations, respectively, without noticeable impacts on both encoding and decoding run-time.

This is an encoder only optimization for ECM, specifically modifying the rate distortion optimization process by using non-OBMC samples in motion estimation for screen content. The proposed change only applies to ECM and not to VTM, as the VTM does not have the OBMC tool.

For Class F, the gains are 0.13% for RA and 0.35% for LB.

Cross checker confirms the results and suggests that the benefits are obvious.

Decision (SW): adopt JVET-AA0129

[JVET-AA0169](https://jvet-experts.org/doc_end_user/current_document.php?id=11857) Cross-check of JVET-AA0129 [F. Le Léannec (Xiaomi)] [late]

[JVET-AA0132](https://jvet-experts.org/doc_end_user/current_document.php?id=11808) AHG6: ECM software configuration parameters for template matching tools [C.-C. Chen, H. Huang, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution provides a software package on top of ECM5 that adds configuration parameters to turn on and off template matching tools to facilitate evaluating coding tool performance without the need of touching default macro settings of template matching tools and re-compiling the software. It is reported that some of the template matching tools (e.g., TM-OBMC, TM-MMVD) in ECM are not configurable using configuration parameters solely instead of macros since such configuration option is not implemented in the software. The provided software package fixes the aforementioned software issues and is capable of handling template matching tools through configuration parameters without the need of software re-compilation and macro changing. Experimental result shows comparison between macro-off and config-off settings for template matching tools. It is reported that both settings provide identical PSNR results.

This contribution proposes to add 7 template matching related SPS flags on top of the 4 template matching related SPS flags in ECM-5.0.

In ECM-5.0, some template matching tools can only be turned on/off by switching macros and re-compiling.

It was asked if it is possible to only add configuration parameters and not add SPS flags to achieve the same purpose. However this would require to re-compile the software.

It was commented that the proposed software modification could allow more flexible combination of different template matching tools.

It was commented that it may be also beneficial to use one global SPS flag to turn on/off all inter template matching tools. The proponent volunteered to add this global flag during software integration.

It was asked if there are additional template matching tools adopted at this meeting, and it seems that one ARMC-related modification was adopted but may or may not require a separate control.

Multiple experts expressed support for adding the proposed modification to enable more testing of template matching tools.

This contribution focuses only on inter template matching tools. There is no proposed modification to control intra template matching tools.

Decision: adopt JVET-AA0132.

[JVET-AA0237](https://jvet-experts.org/doc_end_user/current_document.php?id=11927) Crosscheck of JVET-AA0132 (AHG6: ECM software configuration parameters for template matching tools) [H. Jang (LGE)] [late]

[JVET-AA0134](https://jvet-experts.org/doc_end_user/current_document.php?id=11810) Non-EE2: POC based BCW weights derivation [Z. Zhang, H. Huang, C.-C. Chen, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes to add BCW weights derived considering POC difference between reference and current pictures. The proposed change is used for AMVP and merge modes. It was implemented on top of ECM-5.0 and tested with and without BcwFast encoder speedup configuration. The test results are summarised as follows:

BcwFast is set to 1:

RA: -0.10% (Y), -0.08% (U), -0.12% (V), 103% (EncT), 100% (DecT).

LDB: -0.05% (Y), -0.07% (U), 0.13% (V), 106% (EncT), 100% (DecT).

BcwFast is set to 0:

RA: -0.xx% (Y), -0.xx% (U), -0.xx% (V), 1xx% (EncT), 1xx% (DecT).

LDB: -0.xx% (Y), -0.xx% (U), -0.xx% (V), 1xx% (EncT), 1xx% (DecT).

Compared to JVET-A0075, which removes elements from BCW, the change suggested in this contribution (adding more weights, and use different weights based on POC difference) comes with a worse tradeoff regarding encoding time vs. performance. On the other hand, the decoder change is simpler (just changing syntax parsing, no template matching).

During the discussion, it is pointed out that there could be potential that the two BCW related proposals could be combined and add up gains, as both are targeting different aspects of BCW (e.g. using TM for reordering of weights from A0075, or omit negative weights in the POC difference criterion).

Investigate in EE

[JVET-AA0190](https://jvet-experts.org/doc_end_user/current_document.php?id=11878) Crosscheck of JVET-AA0134 (Non-EE2: POC based BCW weights derivation) [R.-L. Liao (Alibaba)] [late]

[JVET-AA0136](https://jvet-experts.org/doc_end_user/current_document.php?id=11812) Non-EE2: On CCCM improvement [Y.-J. Chang, V. Seregin, M. Karczewicz (Qualcomm)]

In ECM, there are three types of templates for CCLM to derive the linear modes: full template, top-only template, and left-only template. This template selection scheme is extended to CCCM in this contribution. In ECM, the predictors of non-CCLM modes can be fused with the predictors of MMLM. In this contribution, MMLM fusion is replaced with multi-model CCCM fusion. On top of ECM-5.0, the experimental results are summarized as follows:

AI: -0.05 % (Y), -1.14 % (U), -1.09 % (V), 103 % (EncT), 101 % (DecT)

RA: -0.03 % (Y), -0.79 % (U), -0.9 % (V), 100 % (EncT), 98 % (DecT)

There are two aspects in this contribution, first one is to extend CCCM with top-only and left-only templates, and the second aspect is to replace MMLM predictor with CCCM predictor in MMLM fusion.

Currently results are only available for the combination of both aspects, but not for each aspect individually.

The proposal mainly benefits chroma, and has chroma performance gain for all sequences.

Investigate in EE. Also provide results for each aspect.

[JVET-AA0183](https://jvet-experts.org/doc_end_user/current_document.php?id=11871) Crosscheck of JVET-AA0136 (Non-EE2: On CCCM improvement) [J. Lainema (Nokia)] late]

[JVET-AA0137](https://jvet-experts.org/doc_end_user/current_document.php?id=11813) Non-EE2: Intra Prediction Fusion [K. Cao, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes an intra prediction fusion method that blends intra predictors derived from two different reference lines. The proposed method was implemented on top of ECM-5.0 software, it reportedly provides Y, U, V-BD rate reduction with encoder and decoder runtime as follows:

AI: -0.13%, -0.09%, -0.05%, 106% (EncT), 102% (DecT)

RA: -0.04%, -0.03%, -0.10%, 101% (EncT), 100% (DecT)

This contribution is related to JVET-AA0120 and JVET-AA0246.

The proposed intra prediction fusion replaces the current intra prediction process with the following three restrictions:

* CU size must be greater than 16 in total number of luma samples
* CU is not predicted using ISP mode
* Only applied to angular prediction modes with non-integer slope

Two reference lines (default reference line as indicated by the MRL mode, which could be the reference line immediately adjacent to the current CU when MRL is not used, and the one reference line top/left to the default reference line) are used in fusion with fixed weights (3/4, 1/4). This extends by one line the reference area required in ECM-5.0 within the current CTU, so no buffering beyond CTU boundary is needed.

The proposed fusion method also applies to DIMD and TIMD modes.

The current proposed method performs two predictions and then combines them. It was commented that an alternative is to fuse the two reference lines together first and then perform only one prediction.

See further notes under JVET-AA0246.

Investigate in EE.

[JVET-AA0207](https://jvet-experts.org/doc_end_user/current_document.php?id=11896) Crosscheck of JVET-AA0137 (Non-EE2: Intra Prediction Fusion) [K. Kim (WILUS)] [late]

[JVET-AA0234](https://jvet-experts.org/doc_end_user/current_document.php?id=11924) Crosscheck of JVET-AA0137 (Non-EE2: Intra Prediction Fusion) [L. Xu (OPPO)] [late]

[JVET-AA0139](https://jvet-experts.org/doc_end_user/current_document.php?id=11815) Non-EE2: Longer deblocking filter for luma [K. Andersson, J. Enhorn (Ericsson)]

ECM can operate at a larger CTU and transform size than VVC. To enable stronger deblocking of smooth large blocks it is proposed to increase the maximum luma deblocking length from 7 to 15 for a boundary segment that has a size orthogonal to the block boundary of 64 samples or more. The filter design is similar to the VVC deblocking design for length greater than 3. The deblocking decision for a longer filter (length 15 on at least one side) is checked if length 7 deblocking decisions are passed. The additional decision for length 15 is similar to VVC but a bit more restrictive.

It is asserted that subjectively improvements can be seen on both SDR and HDR material especially at lower bitrates.

BD rate impact compared to ECM-5.0:

RA: 0.00%, LDB: -0.03%, AI: 0.00%

Encoding and decoding time is asserted to be similar as ECM.

The proposal includes adding a longer deblocking for luma, and changing filter selection decision for some larger blocks with size greater than or equal to 64 in one dimension.

Cross checker reported that some subjective improvement can be observed for Campfire, QP 42. Subjective improvements for other cases are not as obvious.

The proponent verbally reported subjective improvements can be observed in more cases.

More subjective viewing is needed to better understand the subjective benefits.

It was commented that deblocking is more “down-stream” compared to other coding tools in ECM, and it might be too early to consider such changes.

Further study is encouraged.

[JVET-AA0177](https://jvet-experts.org/doc_end_user/current_document.php?id=11865) Crosscheck of JVET-AA0139 (Non-EE2: Longer deblocking filter for luma) [N. Hu (Qualcomm)] [late]

[JVET-AA0141](https://jvet-experts.org/doc_end_user/current_document.php?id=11817) Non-EE2: Enhanced temporal motion information derivation [L. Zhao, K. Zhang, L. Zhang (Bytedance)]

This proposal proposes two aspects to improve the temporal motion information derivation process for sub-block-based TMVP (sbTMVP) and AMVP modes. Firstly, two collocated frames are utilized to provide temporal motion information. Secondly, the motion shift to locate TMVP is adaptively determined from multiple locations according to template costs. On top of ECM-5.0, simulation results of the proposed method are reported as below:

RA: {-0.10%, -0.10%, -0.16%; 105%, 103%};

LDB: {-0.20%, -0.20%, -0.23%; 112%, 108%}.

It was asked why LDB runtime increase is so much larger than RA, and the proponent suggested TMVP related tools tend to have larger LDB runtime increase.

The proponent said that currently implementation is not optimal which could result in more-than-necessary runtime increase.

There is no cross check for this contribution. Qualcomm and Tencent volunteered to perform cross check.

It was asked if results are available that do not use the template matching aspect. No such results are currently available.

It was commented that sbTMVP has 0.43% RA gain in VTM-10.0, but has only 0.18% RA gain in ECM-5.0.

The proposal adds one more sbTMVP candidate on top of the sbTMVP candidate.

Multiple experts expressed interest to study the following aspects from this proposal:

* sbTVMP aspect without template matching
* more optimized implementation and better encoder to reduce runtime increase

Investigate in EE.

[JVET-AA0142](https://jvet-experts.org/doc_end_user/current_document.php?id=11818) AHG12/Non-EE2: Picture-Level Geometry Transform [W. Jia, K. Zhang, Y. Wang, T. Fu, Y. Li, L. Zhang (Bytedance)]

In this contribution, a method of picture-level geometry transform is proposed. Three kinds of geometry transform: horizontal flip, vertical flip and 180°-rotation can be applied at picture level. It is signaled in the slice header to indicate whether a geometry transform is applied, and which transform is applied. On top of ECM-5.0, simulation results of the proposed method are reported as below:

AI: -0.19%, -0.08%, -0.10%. Encoder run time is increased by 1%

In a first step, a fast algorithm is run only on 16x16 block size which decides the geometry transform to use. It is commented that it would be interesting to hear how much better a brute-force approach doing multiple full encodings might become.

Similar methods had been proposed before (even for AVC), but usually came with much higher encoder run time.

The method is only applied on I slices. It is pointed out that an extension to other slice types might cause many complications.

An additional buffer may be necessary, as the picture is stored in original orientation in the reference picture buffer. An additional pass is necessary at least at the encoder.

Tradeoff of gain vs. encoder run time is attractive.

How often is the geometry modification applied?

It was asked why this is signalled in the slice header and not picture header. A comment is made that the overhead is small, and there may be situations where multiple slices exist and are treated differently, e.g. in some 360 packing schemes.

It was suggested also to test in screen content classes.

Investigate in EE.

[JVET-AA0143](https://jvet-experts.org/doc_end_user/current_document.php?id=11819) Non-EE2: Simplification methods for OBMC [K. Kim, D. Kim, J.-H. Son, J.-S. Kwak (WILUS)] [late]

Initial version rejected as “placeholder”

This contribution proposes several simplification methods for OBMC. The first method is to remove the ‘obmc\_flag’ syntax which is signalled in every CU. Encoder would save memory associated with OBMC on/off decision. The second method is to disable a sub-block OBMC according to the result of a template matching based OBMC. If a template matching based OBMC has not been performed on any CU boundary block, a sub-block OBMC is also not applied for the CU. The third method is to disable OBMC in coding blocks with large block size. When the block size cbW \* cbH is greater than or equal to 32(ECM-5.0) and cbW and cbH is less than 128 (additional restriction), OBMC is applied. On top of ECM-5.0, simulation results of the proposed method are reported as below:

RA: {0.03%, 0.15%, 0.05%, 99%, 88%}

LDB: {0.04%, 0.24%, 0.10%, 101%, 96%}

Some small gain in screen content classes.

A clarification was asked about the second method.

It was commented that in screen content better results are usually achieved by turning OBMC off.

The main target is reducing decoder run time. However, in the current exploration experiments, decoder runtime is not a major problem.

The cross-checker comments that the first element (removing the flag) is not useful, and also not a simplification.

The proposal might be relevant if at a later stage a more practical design of OBMC would be needed.

[**JVET-AA0244**](https://jvet-experts.org/doc_end_user/current_document.php?id=11934) **Crosscheck of JVET-AA0143 (Non-EE2: Simplification methods for OBMC) [X. Xiu (Kwai)] [late]**

[JVET-AA0144](https://jvet-experts.org/doc_end_user/current_document.php?id=11820) Non-EE2: DMVR for affine merge coded blocks [J. Chen, R.-L. Liao, X. Li, Y. Ye (Alibaba)]

This contribution proposes to apply DMVR on the blocks coded with affine merge mode. For simplicity, only the CU level base motion vector refinement is proposed in this contribution. It is reported that applying proposed DMVR on affine blocks achieves {-0.13%(Y), -0.03%(U), -0.07%(V)} coding gain in RA on average, and luma coding gain of -0.17% on class A1 and -0.27% on classA2.

In v2, a new test (test 2) with less running time is provided.

Test 2 reduces the search range to 3, and some encoder optimization is performed.

Encoding time increase is by 8% in test 2. Decoding time increased by 6%.

SAD is used as criterion (same as in first pass of multi-pass DMVR).

Both cross-checkers report matching of results, and support investigation in the EE. They see good potential to further reduce the run time both of encoder and decoder (for example, usage of 12-tap filter may not be necessary).

Investigate in EE

[JVET-AA0195](https://jvet-experts.org/doc_end_user/current_document.php?id=11884) Crosscheck of JVET-AA0144 (Non-EE2: DMVR for affine merge coded blocks) [B. Vishwanath (Bytedance)] [late]

[JVET-AA0208](https://jvet-experts.org/doc_end_user/current_document.php?id=11898) Cross-check of JVET-AA0144 (Non-EE2: DMVR for affine merge coded blocks) [H. Huang (Qualcomm)] [late]

[JVET-AA0146](https://jvet-experts.org/doc_end_user/current_document.php?id=11822) AHG12/Non-EE2: Fixes on ECM for 360-degree video coding [Y. Wang, K. Zhang, Z. Deng, L. Zhang (Bytedance)]

This contribution presents two fixes to solve the crash issue of ECM for 360-degree video coding. The horizontal wrap around motion compensation, a 360-specific coding tool, is not fully considered by new coding tools in ECM. ECM-5.0 crashes when the horizontal wrap around motion compensation is enabled. Two fixes are proposed to solve the issue. First, clipMv() is replaced by wrapClipMv() in LIC, ARMC, MVD sign prediction, and multi-pass DMVR when the horizontal wrap around motion compensation is enabled. Second, out-of-boundary checking is skipped when the horizontal wrap around motion compensation is enabled. Compared to padded equi-rectangular projection (PERP) coding using VTM-17.0, simulation results of PERP coding using ECM-5.0 with the proposed fixes are reported as below:

End-to-end WS-PSNR: -13.21%, -24.59%, -25.88%;

End-to-end S-PSNR-NN: -13.20%, -24.62%, -25.92%.

The second fix (disabling OOB) would be more than a SW bug fix, would be normative (but extremely simple HLS modification, e.g. to be imposed by bitstream restriction), but is obviously solving a problem. Alternatively, OOB might be redesigned to only use it on top and bottom of the picture in case of wraparound, but that might not have any merit in the case of PERP, as the top and bottom areas are very homogeneous due to the projection.

It is interesting being able to use ECM also for this type of content. It is however noted that the gain for the 360 classes seems to be lower than for other classes in RA.

Decision (SW/BF+description doc): Adopt JVET-AA0146

[JVET-AA0189](https://jvet-experts.org/doc_end_user/current_document.php?id=11877) Crosscheck of JVET-AA0146 (AHG12/Non-EE2: Fixes on ECM for 360-degree video coding) [R.-L. Liao (Alibaba)] [late]

[JVET-AA0148](https://jvet-experts.org/doc_end_user/current_document.php?id=11824) Non-EE2: On MHP (Multi-Hypothesis Prediction) [K. Sato, Y. Yu, H. Yu, D. Wang (OPPO)]

This contribution proposes to extend add\_hyp\_weight\_idx for multi-hypothesis prediction currently in ECM. In particular, two more index values are added for add\_hyp\_weight\_idx to signal additional α values equal to 1/8 and -1/4, respectively.

It is reported that the simulations showed the following results -- Y: -0.02 %, U: 0.04 %, V: -0.03% with EncT: xxx % and DecT: xxx % for RA CTC, and Y: -0.09 %, U: -0.28 %, V: -0.19% with EncT: 103 % and DecT: 101 % for LDB CTC.

The gain is higher with LD, but with RA no significant loss is observed. With “Slide Editing” and “SlideShow” in LD CTC, gain by -0.73% and -1.04% are observed, respectively. Both in RA and LD no significant increase in EncT/DecT can be observed.

It is requested that the proposed method be studied under the next round of EE.

Very simple change of syntax and decoder, but more checks are necessary at the encoder and encoding time increases by 3% (in LDB), versus a gain of approx. 0.1%. Gain in RA is lower, partial encoding time results indicate that it may be around 1% increase.

Investigate in EE

[JVET-AA0188](https://jvet-experts.org/doc_end_user/current_document.php?id=11876) Crosscheck of JVET-AA0148 (Non-EE2: On MHP (Multi-Hypothesis Prediction)) [D. Kim, W. Lim, S.-C. Lim (ETRI)] [late]

[JVET-AA0150](https://jvet-experts.org/doc_end_user/current_document.php?id=11826) AHG12: On CIPF (CABAC Initialization from the Previous Frame) [K. Sato, Y. Yu, H. Yu, D. Wang (OPPO)]

In the JVET-Z meeting, CIPF (CABAC Initialization from the Previous Frame) was adopted to ECM5.0. In the adopted method, CABAC context tables of the same slice-type, same sliceQp, and same Tid from a previously coded frame can be inherited and used as the initial tables for coding the current frame. The amount of buffer for CIPF is restricted to 5.

However, encoders equipped with the capabilities of rate control and perceptual Qp optimization for real-world applications like VVenC usually do not use a constant SliceQp even among the frames in the same temporal layer.

This contribution first analyses the behaviours of the CIPF buffer under two configurations: (1) with RA CTC where SliceQp is constant within a temporal layer, and (2) with a RA nonCTC QP setting where SliceQp is not constant within a temporal layer. Then this contribution proposes a method to operate the CIPF buffer in a proper manner even when the SliceQps in one temporal layer are not constant.

The proposed method does not change the behaviour of CIPF under CTC.

It is recommended that the proposed method be studied under AHG12 or EE2.

Gain under non-CTC: -0.06% luma in RA. It was asked how in terms of performance this non-CTC condition would compare to CTC. It causes 0.5% loss on average.

It was suggested that for supporting more general cases (including non-constant QP coding) another solution might be to signal explicitly from which reference picture the CIPF inheritance is used.

An example of slice QP changes is taken from VVEnc, but it is commented that in many encoders, much larger changes of QP are practically used.

It was suggested that another option could be just increasing the buffer size and storing information for several different QP values.

It was further suggested that also investigating LDB case could be relevant.

It is generally agreed that the current CIPF design has been optimized for CTC (which is appropriate for the EE exploration), and that probably changes would be necessary in case it would be used in a standard by a later time to make it more generally beneficial also for other encoder configurations. A change is not of high importance at this moment, but further study is welcome.

[JVET-AA0242](https://jvet-experts.org/doc_end_user/current_document.php?id=11932) Crosscheck of JVET-AA0150 (AHG12: On CIPF (CABAC Initialization from the Previous Frame)) [H. Golestani (Qualcomm)] [late]

Results match

[JVET-AA0191](https://jvet-experts.org/doc_end_user/current_document.php?id=11880) [AHG12] On CTC for Low Complexity ECM [K. Naser, A. Robert, T. Poirier, P Le Guyadec, S. Puri, F. Galpin (InterDigital)] [late]

See section 4.3

[JVET-AA0246](https://jvet-experts.org/doc_end_user/current_document.php?id=11936) Non-EE2: Combination of JVET-AA0120 and JVET-AA0137 [L. Xu, Y. Yu, H. Yu, D. Wang (OPPO), K. Cao, V. Seregin, M. Karczewicz (Qualcomm)] [late]

This contribution proposes to combine the template-based multiple reference line intra prediction in JVET-AA0120[1] and the intra prediction fusion method proposed in JVET-0137 [2]. The proposed method was implemented on top of ECM-5.0 software, and it reportedly provides Y, U, V-BD rate reduction with encoder and decoder runtime as follows:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | |
|  | **Over ECM-5.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -0.11% | -0.06% | -0.11% | 110% | 103% |
| Class A2 | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
| Class B | -0.20% | -0.12% | -0.10% | 109% | 104% |
| Class C | -0.24% | -0.29% | -0.08% | 107% | 110% |
| Class E | -0.22% | -0.07% | -0.22% | 109% | 103% |
| **Overall** | #VALUE! | #VALUE! | #VALUE! | #DIV/0! | #DIV/0! |
| Class D | -0.21% | -0.11% | -0.25% | 106% | 109% |
| Class F | -0.09% | -0.07% | -0.20% | 106% | 105% |

This contribution is related to JVET-AA0120 and JVET-AA0137.

The proponents assert that the coding gains from JVET-AA0120 and JVET-AA0137 are almost additive, and AA0246 is a straightforward combination of the two proposals. In this contribution, when TMRL in JVET-AA0120 is selected, intra prediction fusion is also applied to the selected TMRL mode. During template matching, only one reference line is used. During prediction, two reference lines are used as in JVET-AA0137.

Test a in JVET-AA0120 (approximately 0.08% gain in AI) is used in this combination.

The following aspects could be further investigated:

* performance vs. runtime from JVET-AA0120 and JVET-AA0137 separately, and also in combination
* Investigate reference line fusion instead of prediction fusion for the proposed intra prediction fusion method in JVET-AA0137 (see notes under JVET-AA0137)

Investigate in EE.

It was agreed that it should be targeted to run subjective tests on ECM6 vs. VTM at the next meeting. Rate matching should be done with VTM. Expecting a bit rate saving of close to 20%, this might translate into a QP+1 change for VTM. Otherwise, there is the established procedure of changing the QP at some point towards the end the sequence. ECM should not have significantly higher rate than VTM. Details to be figured out in telco(s) between AHG4 (and AG 5), AHG6, and AHG12. One early telco (before the finalization of ECM6) is recommended, plus at least one thereafter.

# High-level syntax (HLS) and related proposals (17)

## AHG9: SEI message studies and proposals (5)

Contributions in this area were discussed in session 5 at 0500–0730 UTC and session 6 at 0750-0830 UTC on Thursday 14 July 2022 (chaired by JRO).

[JVET-AA0079](https://jvet-experts.org/doc_end_user/current_document.php?id=11755) AHG9: Decoded picture hash SEI message extension [P. Bordes, F. Galpin, P. DeLagrange, E. François (InterDigital)]

This contribution proposes syntax and semantics for extension of the Decoded picture hash SEI message to support several checksum hashes at different steps of the decoded pictures reconstruction process.

Cases of post processing that are proposed to be supported are colour transform, film grain, post filter, and clean area after GDR.

It was commented that an expression is missing about which samples would be subject to the check, e.g. cropped window, or specific positions.

It was commented that commonly new SEI messages are defined rather than extending existing ones for new purposes.

It was further commented that bit-exact reproduction of post processing is currently not required in other SEI messages, and might also be difficult to achieve. For example, in film grain, it might depend on the random number generator, in case of NN based post processing, it is known from EE1 that deviations occur depending on the floating point format of the processing device.

It was commented that an SEI like this could be useful to check the output in cases where bit-exact reproduction is possible (e.g., neural network with integer implementation and without overflow), without necessarily requiring that it is mandatory to produce such an output in the corresponding other SEI message.

It was also commented that currently the sequence of multiple post processing steps is not defined, such that the check “after last PP step” could give different results (JVET-AA0102 defines specification of such sequence).

Further study is recommended.

[JVET-AA0091](https://jvet-experts.org/doc_end_user/current_document.php?id=11767) AHG9: Resolution Change Information SEI message [V. Drugeon, K. Abe, T. Toma (Panasonic)]

Dynamic resolution encoding can be used for dynamically adapting the resolution of a video on a scene-by-scene basis to reach the desired bitrate while preserving the quality. However, many receivers have been developed with the assumption that the picture resolution within a bitstream stays constant, for example within an MPEG-DASH Representation.

It is suggested to consider signalling of dynamic resolution encoding in bitstreams, together with the resolutions used and optionally which pictures use each resolution. The authors of the present document believe that such a signalling would greatly simplify implementation of receivers in environments where bitstreams may use dynamic resolution encoding.

The SEI message is proposed for VVC, HEVC and AVC. The intention is to let a decoding device know which resolutions will appear in a bitstream. It was commented that hypothetically this information could be drawn from PPS, but that might be only possible for VVC (with RPR). For AVC and HEVC, it would be necessary to start a new CVS for a resolution change. In case of CMAF, it is done that way, such that the necessary information is available at the systems layer.

It was commented that there may be problems when splicing of bitstreams is performed – when the splicer changes the properties that are specified in the SEI. Existing devices would not know about this SEI messages, such that “promises might be broken”. This could be resolved if the language in the semantics would be softened, similar as in the manifest SEI message.

It was commented that it might be another option to convey such information at systems layer rather than SEI. In the case of broadcast such as DVB, it is however currently not the case that the information is available at the beginning (as would be the case in CMAF).

Further study recommended.

[JVET-AA0102](https://jvet-experts.org/doc_end_user/current_document.php?id=11778) AHG9: SEI processing order SEI message [P. Yin, S. McCarthy, W. Husak, K. Konstantinos, T. Lu, F. Pu, A. Arora, T. Shao (Dolby)]

An SEI message for carrying information on the preferred order of processing SEI messages is proposed. An encoder (i.e., content author) could use the proposed SEI message to indicate that SEI messages of a particular payloadType should be processed by the receiving system either before or after SEI messages of another payloadType to enable an adequate user experience. As an example, the SEI message processing order SEI message could be used to indicate that information in a colour transform information SEI message should be processed before neural-network post-filtering using information in NNPF SEI messages. In this case, for example, the processing order could matter because the NNPF was trained using original (unmapped) luma and chroma sample values rather than mapped luma and chroma sample values. The proposed SEI processing order SEI message is modelled on the SEI manifest SEI message.

Version 2 of this contribution contains refinements in response to review of version 1 of this contribution, including:

• A section was added to describe the function of the proposed SEI message, including the behaviour when an indicated SEI message is not present or when a receiver is not capable of processing the indicated SEI message

• Improvements to the syntax element that specifies the number of types of SEI messages for which information is provided in the SEI processing order SEI message

• A brief discussion of other comments noted during the review of the first version of this contribution

JVET-AA0101 is also related, specifically for the case of NN based post processing.

The SEI message would be suitable for VVC rather than VSEI.

Why 16 bits for number of payload types? (was changed to 8 bit in version 2 of the contribution)

It was commented that hypothetically the loop could be running without knowing the number at the decoder until the SEI ends. (a syntax proposal for this is included in 4.1 of version 2 of the contribution).

The manifest message indicates that certain SEI messages are likely to be present but they don’t need to appear in the end. The same concept is followed here, but when one of the SEIs in the sequence is missing, the whole sequence would become obsolete. Nevertheless, the decoder could still execute the remaining ones in arbitrary sequence. Same case if a decoding device is unable to support one of the steps.

As the necessity of such an SEI message is very specific for some SEI messages mostly from the domain of post processing, would it be more appropriate to specify less generally? It was later argued that it is better keeping it generic, either for future SEI messages or existing ones (360 degree messages were mentioned as an example where a concatenation sequence might be important as well).

Version 2 with some changes was reviewed in session 24 0800 UTC

If one of the concatenated SEI messages is not available (or not implemented), the sequence of theother ones should still be maintained.

Also cases are supported where the order of some of the concatenated SEI messages may not be in preferred order, e.g. either A-B-C or B-A-C might be expected.

It was emphasized that this SEI proposal is only about preferred concatenation, not about an expected “normative output”. A decoder can still use any order, or omit certain SEI messages.

It was asked again why not just interpret the order of transmission as the preferred order. However, the current spec also does not contain any statement about that. Further, that might be problematic when the SEI messages have different persistence, e.g. some are valid for the entire sequence and others sent later for a specific picture.

Decision: Adopt JVET-AA0102, syntax as from 4.1 of version 2, semantics from 3.1.4.

Generate an output document “preliminary WD of …” (JVET-AA2027, and WG 5 doc), which hypothetically could be included already in te other ongoing amendment/next edition of VVC at a later meeting.

[JVET-AA0105](https://jvet-experts.org/doc_end_user/current_document.php?id=11781) AHG9: Metadata for display on transparent screens based on ACI SEI messages [E. Thomas, P. Andrivon, F. Le Léannec, M. Radosavljević, M.-L. Champel (Xiaomi)]

Following up on JVET-Y0104 and JVET-Z0129, this contributions asserts that the signaling of metadata for assisting with the adjustment of the transparency effect of video sequences on transparent screens has two main benefits. First, it is asserted that this metadata would prevent a degraded quality of experience when displaying conventional content on transparent screens due to the possible loss of important visual content made unintentionally transparent (e.g. missing or barely visible logo, text, objects, etc.). Second, it is also asserted that a content creator may leverage the metadata to use the transparency effect as a way to enhance the narrative effect (e.g. visual content fading in and out in the viewer’s surroundings, highlighting an important part of the image such as a character, a face, an object, etc..).

Following the recommendations from JVET 26th, this contribution provides a technical solution to carry this metadata and enabling the use case by reusing the alpha plane auxiliary pictures and the associated alpha channel information (ACI) SEI messages. The changes introduce a new mode for the parameter alpha\_channel\_use\_idc as well as semantic constraints for this new mode for the values of the parameters alpha\_transparent\_value and alpha\_opaque\_value. Lastly, an informative note is provided to describe the possible post decoding operations for this new mode as it is done for the existing modes. As comparison point, an approach is also presented for enabling the use case with the same benefits using a new type of auxiliary pictures.

It was commented that the equation suggested for NOTE4 is basically identical to the existing blending equation. The note may not be needed at all.

It was discussed if a new mode is needed at all. The exact adjustment of the transparency is depending on display type, viewing environment, etc. Otherwise, the alpha map just gives information where a relevant foreground object is.

Unclear why a new mode is needed. More information would be necessary what is different from other usage. The only difference is that it is prescribed in this mode that transparent and opaque levels shall be identical. For the latter, it is unclear why this flexibility is removed.

[JVET-AA0110](https://jvet-experts.org/doc_end_user/current_document.php?id=11786) AHG9: SEI message with sample phase indication for consistent rendering [J. Samuelsson-Allendes, S. Deshpande (Sharp)]

An SEI message containing a sample phase indication is proposed. It is asserted to enable more consistent rendering, in particular when rendering at a higher resolution and when switching between multiple resolutions as is common in streaming scenarios where bandwidth may fluctuate.

Compared to JVET-Y0156-v1 presented and discussed at the 25th JVET meeting, this version includes an improved text description including the definition of persistence, as well as numerical examples.

Presented in session 6, 0750 UTC

Various ratios of downsampling are supported, and examples are given in the presentation.

The proponents suggest that the original resolution and separate chroma phase do not need to specified. The latter can be deduced by VUI.

Proposed for VSEI. This might require some re-wording with regard to cropping.

Is a common denominator for horizontal/vertical useful? Should be split into two different syntax element.

It is also suggested to use a “-1” syntax for the denominators, and give up the current semantics that zero value means unknown phase.

Is a precision of 1/512 sample precision needed? Probably not, but it is of no harm to use more precision than for the currently foreseen use cases. Visibility of the effect is highly content dependent, mostly visible for still video.

Candidate for inclusion in VSEI extension. The proponents are asked to provide text considering the modifications suggested above.

Decision: Adopt JVET-AA0110 for VSEI extensions draft 2

## Neural-network post filter (8)

Contributions in this area were discussed in session 6 at 0830–0930 UTC on Thursday 14 July 2022, and in sessions 9 and 10 at 0500-0700 and 0720-0845 on Friday 15 July 2022 (chaired by JRO).

[JVET-AA0054](https://jvet-experts.org/doc_end_user/current_document.php?id=11730) AHG9: On Neural-network Post-filter Characteristics SEI Message [S. Deshpande (Sharp)]

Some modifications are proposed to the Neural-network post-filter characteristics SEI. Message. Following is proposed:

* Proposal 1: It is proposed to specify signalling of external URI information in Neural network post-filter characteristics SEI message. It is asserted that this allows defining neural network information externally.
* Proposal 2: It is proposed to signal one flag instead of two flags (nnpfc\_out\_sub\_width\_c\_flag and nnpfc\_out\_sub\_height\_c\_flag) to specify output chroma format information. This saves 1 bit.
* Proposal 3: A sanity check constraint is proposed to disallow signalling invalid values. In particular, it is proposed to be disallowed to signal nnpfc\_purpose equal to 2 or 4 indicating upsampling of chroma format, when the input is monochrome or 4:4:4.

Proposal 1: It is proposed to resolve the signalling of external URI by a tag URI (such a mechanism is defined by IETF in RFC4151, without need for registration authority) as identifier. This is also used in ATSC.

If a decoding device cannot interpret the definition found at the URI, it would not be able to process it. It was pointed out that it cannot be controlled if the method found at that URI is publicly known and could be interpreted by everybody.

Several experts expressed opinion that this would be a suitable mechanism.

Proposal 2: One bit is saved but signalling only one flag for chroma upsampling instead of two separately for width and height. This is justified due to the fact that currently only a limited amount of upsampling ratios are defined in the given purpose definitions, and the input chroma format is also available. For future purposes, other syntax elements would need to be added anyway.

Proposal 3 disallows an unreasonable combination of syntax elements and purposes.

Further editorial remarks:

* for null-terminated UTF-8, see the definition of st(v) in the HEVC standard.
* inpSubWidthC and inpSubHeightC do not need to be defined as input parameters. See subclause 7.3 of VSEI. Equivalent parameters are already defined in VSEI. Note that the decoded picture hash SEI message uses those variables without defining them as input parameters since they are already derived in VSEI from the VUI.

Decision: Adopt JVET-AA0054 (proposals 1-3)

[JVET-AA0055](https://jvet-experts.org/doc_end_user/current_document.php?id=11731) AHG9: Comments on Neural-network Post-filter Characteristics SEI Message [S. Deshpande, A. Sidiya (Sharp)]

This document contains comments on Neural-network post-filter characteristics SEI message. Following is proposed:

* Proposal A: It is proposed to allow signalling additional padding types for the neural network.
* Proposal B: A new value is proposed to indicate the type of neural network parameters. Also, the syntax element nnpfc\_parameter\_type is changed from a flag to a two-bit idc for this.
* Proposal C: It is proposed to signal frame rate upsampling as an additional purpose for post processing filter. Additionally, it is proposed to conditionally signal two syntax elements for frame rate upsampling purpose.
* Proposal D: Some sanity check constraints are proposed for patch size related syntax elements (nnpfc\_constant\_patch\_size\_flag, nnpfc\_patch\_width\_minus1, nnpfc\_patch\_height\_minus1) to disallow signalling invalid values.

Proposal A adds two new input padding types which are used internally in some NN processing packages. “Wraparound padding” is padding left to right, top to bottom. It was asked what would be the use case of top/bottom? For example, ERP only uses left to right. As this could be an input to the network, wraparound should be only left/right. Currently, no use case of top/bottom padding is known.

For fixed padding value, it was suggested in the discussion to define separate values for luma and the two chroma components.

Proposal B was agreed, removing “may” in case of type 2.

Proposal C was asserted as a relevant purpose type, but the suggested signalling is not sufficient. It was commented that a corresponding definition of the output tensor would be needed. Further study on this.

During the discussion, it was noted that some change of expressing the number format of the input and output tensors might need consideration. For example, depending on the purpose, a device could convert the output to whatever is needed, and only needs to know which format comes from the network output, which should be known from the NNR description.

Proposal D was agreed. The same is also proposed as one element in JVET-AA0067 (problem 2). Some editorial rewording may be useful.

Decision: Adopt JVET-AA0055 Proposal A with wraparound padding only L/R, fixed padding separate values for luma/chroma; Proposals B/D as written above.

[JVET-AA0056](https://jvet-experts.org/doc_end_user/current_document.php?id=11732) AHG9: On syntax gating in the neural-network post-filter characteristics SEI message [M. M. Hannuksela, F. Cricri, M. Santamaria (Nokia), T. Chujoh, Y. Yasugi, T. Ikai (Sharp), S. McCarthy, A. Arora, T. Shao, P. Yin, T. Lu, F. Pu, W. Husak (Dolby)]

This contribution proposes to add nnpfc\_purpose\_and\_formatting\_flag in the neural-network post-filter characteristics (NNPFC) SEI message and change some constraints of the NNPFC SEI message. It is asserted that the proposed changes achieve the following impacts:

1. Enabling the signalling of syntax elements for the filtering purpose, input tensor formatting, output tensor formatting, and complexity for an externally specified filter.
2. Avoiding redundant signalling of syntax elements for the filtering purpose, input tensor formatting, output tensor formatting, and complexity in an NNPFC SEI message containing an MPEG Neural Network Representation (NNR) update.
3. Enabling an NNPFC SEI message with an NNR update of an externally specified filter.

It was agreed that the introduction of a gating flag is useful, both for external means and NNR update.

Decision: Adopt JVET-AA0056.

During the discussion, the previous decision on JVET-AA0054 proposal 1 was questioned, as well as more generally the option of supporting external means was put into question.

[JVET-AA0067](https://jvet-experts.org/doc_end_user/current_document.php?id=11743) AHG9: Some specification improvements for neural-network post-filter characteristics SEI message [T. Chujoh, Y. Yasugi, T. Ikai (Sharp), M. Hannuksela, F. Cricri (Nokia), S. McCarthy, A. Arora, T. Shao, P. Yin, T. Lu, F. Pu, W. Husak (Dolby)]

In this contribution, some specification improvements for neural-network post-filter characteristics SEI message have been proposed. Since five problems regarding the range of the syntax elements and the variables in the current working draft were found, how to fix them is presented.

Problem 1: The chroma format of outSubWidthC equal to 1 and outSubHeightC equal to 2 does not exist, but those values are not prohibited. -> was resolved, as those syntax elements were removed, see under JVET-AA0054

Problem 2: It is possible to define the values of patch size as greater than picture size. Those values should be related to picture size. -> was resolved, see under JVET-AA0055

Problem 3: The conditions between BitDepthY/BitDepthC and outTensorBitDepth of equation (83) are incorrect. -> bug in description, agreed

Problem 4: It is possible to define the value of maxNumParameters as more extensive than the range of unsigned 64-bit integers. -> not a bug, but agreed that the suggested change is appropriate

Problem 5: The maximum value of the syntax element nnpfc\_num\_kmac\_operations\_idc is not defined. -> agreed

Decision: Adopt JVET-AA0067 Problems 2-5 (problem 2 also resolved JVET-AA0055, but the differences are purely editorial)

[JVET-AA0083](https://jvet-experts.org/doc_end_user/current_document.php?id=11759) AHG9: NNR post-filter SEI message extension for flexible decoding capabilities [F. Galpin, T. Dumas, P. Bordes, E. François (InterDigital)]

This contribution proposes to amend the NNR post-filter SEI message proposed in JVET\_Z0244. The NNR post-filter SEI message may carry an ISO/IEC 15938-17 bitstream that specifies a neural-network-based post-filter. It also carries information necessary for a decoder to infer the neural network. It is reported that the current design limits flexibility in the inference implementation, that can be for instance required because of the targeted platform capabilities. The NNR post-filter SEI message extension proposed in this contribution allows signaling information used at the decoder side to choose between several Neural Network inference strategies.

The problem stated here is that some processing platforms may not be capable to process a block size specified in the SEI. It was suggested to add signalling on receptive field, furthermore a superblock/subblock concept is introduced.

This would only be needed if an exact match of the output is required (may not be necessary in post processing). Otherwise, processing in smaller subblocks could be executed by an implementation, where even precise match could be achieved if the network uses integer and the overlap is as large as the receptive field (which might however be inefficient in terms of computation with small subblocks).

A script is delivered with the contribution, which is however only testing with relatively small models.

Further study is suggested. Would be interesting to demonstrate how large deviations could become. Furthermore, as the signalling might not be needed when no precise match between encoder and decoder is required, the information about the receptive field might still be useful to let the decoder know if it is capable to process the given model when it can only support smaller subblocks or smaller subblocks.

Investigate in EE with a post filter architecture, exercising that an implementation uses smaller subblocks than anticipated.

[JVET-AA0100](https://jvet-experts.org/doc_end_user/current_document.php?id=11776) AHG9: On auxiliary input and separate colour description in the neural-network post-filter characteristics SEI message [T. Shao, A. Arora, P. Yin, S. McCarthy, T. Lu, F. Pu, W. Husak (Dolby), Miska M. Hannuksela, Francesco Cricri, Maria Santamaria Gomez (Nokia), T. Chujoh, Y. Yasugi, T. Ikai (Sharp)]

The following syntax elements for the neural-network post-filter characteristics SEI message are proposed to be added to the next draft of Additional SEI messages for VSEI (JVET-Z2006). The neural-network post-filter characteristics SEI message was adopted during the 26th meeting of JVET in April 2022.

1. nnpfc\_auxiliary\_inp\_idc to indicate that auxiliary input data may be present in the neural network input tensor for any allowed luma-only, chroma-only, and luma-chroma configuration. Currently, auxiliary input data may be present in only one configuration of a luma-chroma input tensor (nnpfc\_inp\_order\_idc equal to 3) and not in any other luma-only, chroma-only, or luma-chroma configuration (nnpfc\_inp\_order\_idc equal to 0, 1, and 2, respectively)
2. nnpfc\_separate\_colour\_description\_present\_flag to indicate that the colour primaries, transfer characteristics, and matrix coefficients of the picture that results from the neural-network post filtering may be different than for the input to the filter. When nnpfc\_separate\_colour\_description\_present\_flag is equal to 1, the syntax elements nnpfc\_colour\_primaries, nnpfc\_transfer\_characteristic, and nnpfc\_matrix\_coeffs specify the colour primaries, transfer characteristics, and matrix coefficients of the picture that results from the neural-network post filtering.

About 1. It is agreed that enabling the auxiliary input (only slice QP currently) also for nnpfc\_inp\_order\_idc equal to 0, 1, and 2 is useful. This is a commonly used additional input in post filtering.

It is further discussed that many more types of auxiliary input (including local information such as boundary strength, partitioning and residue) are already experimented in EE1. Though it is too early to include such elements in the SEI message, it might not be simple and very specific for a given decoding standard how these might be linked and described, as those are not specified to be commonly output by decoders.

About 2. The intent is to have a different colour format at the output of the network, and this would be described in the same way as currently in VUI (specified in VSEI). The motivation is that the training is done including the purpose for an optimization of a certain colour space which is different from the colour space of the decoder output.

Decision: Adopt JVET-A0100

[JVET-AA0101](https://jvet-experts.org/doc_end_user/current_document.php?id=11777) AHG9: On processing order in the neural-network post-filter activation SEI message [T. Shao, A. Arora, P. Yin, S. McCarthy, T. Lu, F. Pu, W. Husak (Dolby)]

The following syntax elements for the neural-network post-filter activation SEI message are proposed to be added to the next draft of Additional SEI messages for VSEI (JVET-Z2006). The neural-network post-filter activation SEI message was adopted during the 26th meeting of JVET in April 2022.

1. nnpfa\_independent\_flag to indicate preference that the neural-network post-filter signalled in the SEI be either independent of other neural-network post-filters that may also be used for the current picture, or dependent on the output of one or more such neural-network post-filters
2. nnpfa\_num\_preceding\_nnpfa\_ids\_minus1 to indicate the number of neural-network post-filters on which the current neural-network post-filter may depend
3. nnpfa\_preceding\_nnpfa\_id[ i ] to specify the identifying number, nnpfa\_id, of the i-th neural-network post-processing filter on which the current neural-network filter may depend

It was discussed that it might be premature at this moment to think about concatenation of different neural networks. For example, in case of a network performing upsampling, it is trained end-to-end, such that it implicitly considers the quality of the low resolution input and provides optimized output. Additional denoising before or after that process may not even be necessary.

[JVET-AA0145](https://jvet-experts.org/doc_end_user/current_document.php?id=11821) AHG9: On decoupling neural-network post-filter activation SEI message [H.-B. Teo, J. Gao, C.-S. Lim, K. Abe, V. Drugeon (Panasonic)]

This contribution proposes to decouple the neural-network post-filter activation SEI message in the working draft of VSEI amendment from the current picture by introducing a picture identifier to the SEI message. Two variants of the solution are proposed. Three implementations of the solution are suggested.

The motivation behind this proposal is the possible large size of NNR streams, such that it would be better to have them available early enough and be able to load the model parameters before they are activated for some given picture. Therefore, several variants are proposed which give “advance notice” that a certain model is activated for one of the next pictures to be decoded.

It is mentioned that this might be primarily relevant for low delay applications, as otherwise there is typically some delay between finalization of decoding and output.

More study would be needed to understand the problem. Which size of models is realistic for real-time decoding and display? How frequently can new models be loaded in that case? If switching between models is made (e.g. for intra or inter, or in latter case, for different temporal levels), how many are needed?

## Film grain synthesis (3)

Contributions in this area were discussed in session 10 at 0845–0935 UTC on Friday 15 July 2022, and in session 17 at 1300–1350 UTC on Tuesday 19 July 2022 (chaired by JRO).

[JVET-AA0051](https://jvet-experts.org/doc_end_user/current_document.php?id=11727) AHG13: Proposed text: Film grain synthesis technology for video applications (Draft 2) [D. Grois (Comcast), Y. He (Qualcomm), W. Husak (Dolby), P. de Lagrange (InterDigital), 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 end-to-end processing steps for film grain synthesis, that may include content analysis, noise/film grain removal and film grain model parameter estimation, parameter encoding, encapsulation, and decoding, and film grain synthesis and blending for consumer distribution applications.

This version reorganized the technical report in the following order: overview, synthesis, analysis, parameter descriptions and examples while also adding text in areas that were thinly populated. Several figures and pictures were added to provide visual examples in order to better instruct the reader. Examples have been moved to an Annex.

Annexes A.2/A.3 to be filled with the analysis/synthesis algorithms that are implemented in software. Currently the software is already available in the latest version of ITU-T version of VVC reference software. An update of HEVC reference software should be done accordingly (merge request to HM, but not merged yet) – could be at next meeting.

Otherwise, there are sections still to be filled regarding AR models.

[JVET-AA0052](https://jvet-experts.org/doc_end_user/current_document.php?id=11728) AHG13: On VSEI film grain profiles [Y. He, M. Coban, M. Karczewicz (Qualcomm), M. Radosavljević (Xiaomi), M. Raulet (Aterme)]

This contribution proposes two VSEI profiles, frequency filtering film grain synthesis profile and auto-regressive film grain synthesis profile, for the film grain characteristics (FGC) SEI message. The profiles specify constraints on FGC SEI and the requirements for the capabilities needed to enable the use of FGC SEI.

Specification text (annex of VSEI) is attached to the contribution, and presented.

It is commented that the sentence “The definition of "perceptually significant differences" is beyond the scope of this Specification and may be specified by the specifications that reference this Specification for the constraints of the profiles.” is questionable, as it leaves the true definition of conformance potentially to other organisation. Would it then not be better that those organizations would by themselves define the profile. One participant commented that this is not what he understood under the term “soft conformance” in the AHG discussion.

Definition of “soft conformance” as described in the AHG report might be reworded as: An implementer of the XXX profile must implement a film grain synthesis process which may or may not be the process specified as an example synthesis process in the SEI message. Output frames shall be in the same order and should not have perceptually significant differences with the frames that would be produced by the example film grain synthesis process.

It is discussed whether defining such a profile in VSEI would have benefit at all. It would be decoupled from VVC profiles, and decoder implementers would have the choice to build a device which is either compliant only to a VVC profile, or to a profile combination VVC+VSEI-FGS. Basically, the implementation of the SEI message is already an option for any decoder implementer.

It is also pointed out that nobody might be able to test “soft conformance”.

One argument is that by defining such a profile in VSEI would be that application standards such as DVB etc. would just take it up rather than defining it by themselves without having knowledge about the matter.

On the other hand, it was reported that for the case of HDR consistency between different application standards was achieved.

Follow-up discussion in session 24 0845:

A new version of the draft text was presented.

The specification text is asserted to be improved, but the main aspect of the debate had been if it makes sense to define profile(s) at all.

It was commented that defining two different profiles for film grain may even be more confusing than defining nothing at all.

No consensus could be reached that the definition of such profiles would be beneficial.

[JVET-AA0235](https://jvet-experts.org/doc_end_user/current_document.php?id=11925) AHG13: AOMedia technical report on film grain synthesis technology and AFGS1 specification [A. Norkin (Netflix)] [late]

JVET is currently working on a technical report describing practices for film grain technology. This document communicates that AOMedia is working on a standalone film grain synthesis specification (AFGS1). This draft specification uses the ITU-T T.35 user registered metadata, which make it possible to use this technology with video codecs that support the ITU-T T.35 SEI messages. In addition to that, AOMedia is working on a technical report that describes how the AFGS1 technology can be used in a video compression system. Both draft documents are publicly available online and can be used as sources of information and references in the JVET report on the film grain synthesis technologies.

References listed here could be referred to in the DTR, and be replaced by more final documents later, which might be expected to be available when the TR is due (planned for January 2023).

The TR should include some high-level description of the referenced method, also highlighting the specific differences of AR methods, and also the analysis. Appoint A. Norkin as an additional editor of the TR (recommendation).

## Non-SEI HLS aspects (1)

Contributions in this area were discussed in session 17 at 1350–1405 UTC on Tuesday 19 July 2022 and in session 19 at 0615–XXXX UTC on Wednesday 20 July 2022 (chaired by JRO).

[JVET-AA0099](https://jvet-experts.org/doc_end_user/current_document.php?id=11775) AHG9: On subpictures order [Hendry, S. Kim, S. Lee (LGE)]

This contribution asserted that the order of subpictures can be in an order different from raster scan without breaking the availability rules that are specified in the current specification. When subpictures’ order is not in raster scan order, it may cause additional complexity for implementation. It is further asserted that while such flexibility may be desired (i.e., allowing subpictures to be not in raster scan order because of its structure/shape), most common use-case scenarios / applications can be supported with subpictures that are in raster scan order.

It is proposed that subpicture order is simplified such that if the structure of subpicture allows both raster scan order and non-raster scan order, the order shall be in raster scan order. This can be done by defining a constraint or use a new flag in general constraints information (GCI).

Question raised: Is it possible to modify VVC such that raster-scan becomes restricted. Potentially, there might be already encoders and bitstreams supporting other than raster-scan order which would not be compliant any more? (backward compatibility)

Presentation was continued in session 19.

Two solutions are proposed: Bitstream restriction (which would not be backward compatible), or new GCI flag.

Is the problem only existing in case of dependent subpictures, or when loop filtering is performed across subpicture boundaries? Yes.

It was commented that there is definitely no bug. It was a design decision made by purpose, and the flexibility was intended and well understood.

Also solution 2 would not help, as compliant decoders would need to support non-raster-scan anyway.

It was commented that the subpicture order is known by being signalled in SPS, so the decoder knows it ahead and can react accordingly in terms of buffer usage.

The request for the change appears to be related to a specific way of implementing subpictures, particularly for simplifying buffering.

No support by other experts.

# Plenary meetings, joint meetings, BoG reports, and liaison communications

## JVET plenaries (update)

No intermediate plenaries were held, as document review and decisions were made in single-track mode at this meeting.

Some of the discussions and actions at closing plenary sessions are noted in this section.

* Thu. 28 Apr.
  + 2100–2300 UTC (session 18): Review of remaining docs and revisits
* Fri. 29 Apr.
  + 0500–0710 UTC General plenary wrap-up: See notes under sections 8, 9, 10, and 11.
  + 0005+1–0020(+1) WG 5 Closing plenary: Approval of meeting recommendations

## Information sharing meetings

Information sharing sessions with other WGs and AGs of the MPEG community were held on Monday 18 July 0500–0730, Wednesday 20 July 0500–0600, and Friday 22 July 2100–2300. The status of the work in the MPEG WGs and AGs was reviewed at these information sharing sessions.

## Joint meetings with AG 5 and VCEG 2100–2300 on Wednesday 27 April and 0800–0920 on Thursday 27 April

Two joint meeting sessions with SC 29/AG 5 (Visual Quality Assessment) and VCEG were held during the current meeting.

The following topics were discussed in this joint session.

* Test material (see the notes in section 4.5)
* Quality assessment (see the notes in section 4.6)

## BoGs (0)

Three break-out groups were established at this meeting to conduct discussion and develop recommendations on specific topics. See sections 5.2.1 and 5.3.1 and the notes for JVET-Z0115 in section 6.2 for details.

## Liaison communications

The JVET received a liaison letter from SC 29/WG 1 JPEG ([m59796](https://dms.mpeg.expert/doc_end_user/current_document.php?id=82797&id_meeting=190)) related to its JPEG AI project. The letter was sent to provide information to various MPEG WGs informing them about the scope and CfP for the project, and no specific action was requested in the letter. It was considered not necessary to send a response from JVET.

Incoming liaison statement from JPEG (m60532) – initial text drafted in session 24.

# Project planning

## Software timeline

ECM6 software (including all adoptions) was planned to be available 3 weeks after the meeting.

The first release candidate of the NN codebase software was planned to be available 3 weeks after the meeting, and the final release was planned 2 weeks thereafter.

VTM17.1 software was planned to be available on 2022-08-05. (Note that further updates may be released later)

HM16.26 software was planned to be available on 2022-08-05. A new version of HTM was targeted for release before the next meeting.

## Core experiment and exploration experiment planning

An EE on neural network-based video coding was established, as recorded in output document JVET-AA2023.

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-AA2024.

Initial versions of these documents were presented and approved in the first plenary session on Friday 22 July.

## Drafting of specification text, encoder algorithm descriptions, and software

The following agreement has been established: the editorial team has the discretion to not integrate recorded adoptions for which the available text is grossly inadequate (and cannot be fixed with a reasonable degree of effort), if such a situation hypothetically arises. In such an event, the text would record the intent expressed by the committee without including a full integration of the available inadequate text.

## Plans for improved efficiency and contribution consideration

The group considered it important to have the full design of proposals documented to enable proper study.

Adoptions need to be based on properly drafted working draft text (on normative elements) and HM/VTM encoder algorithm descriptions – relative to the existing drafts. Proposal contributions should also provide a software implementation (or at least such software should be made available for study and testing by other participants at the meeting, and software must be made available to cross-checkers in EEs).

Suggestions for future meetings included the following generally-supported principles:

* No review of normative contributions without draft specification text
* VTM algorithm description text is strongly encouraged for non-normative contributions
* Early upload deadline to enable substantial study prior to the meeting
* Using a clock timer to ensure efficient proposal presentations (5 min) and discussions

As general guidance, it was suggested to avoid usage of company names in document titles, software modules etc., and not to describe a technology by using a company name.

## General issues for experiments

It was emphasized that those rules which had been set up or refined during the 12th JVET meeting should be observed. In particular, for some CEs of some previous meetings, results were available late, and some changes in the experimental setup had not been sufficiently discussed on the JVET reflector.

Group coordinated experiments have been planned as follows:

* “Core experiments” (CEs) are the coordinated experiments on coding tools which are deemed to be interesting but require more investigation and could potentially become part of a draft standard by the next meeting or in the near future.
* “Exploration experiments” (EEs) are also coordinated experiments. These are conducted on technology which is not foreseen to become part of a draft standard in the near future. The investigating methodology for assessment of such technology can also be an important part of an EE. (Further general rules for EEs, as far as deviating from the CE rules below, should be discussed in a future meeting. For the current meeting, procedures as described in the EE description document are deemed to be sufficient.)
* A CE is a test of a specific fully described technology in a specific agreed way. It is not a forum for thinking of new ideas (like an AHG). The CE coordinators are responsible for making sure that the CE description is complete and correct and has adequate detail. Reflector discussions about CE description clarity and other aspects of CE plans are encouraged.
* A description of each experiment is to be approved at the meeting at which the experiment plan is established. This should include the issues that were raised by other experts when the tool was presented, e.g., interference with other tools, contribution of different elements that are part of a package, etc. The experiment description document should provide the names of individual people, not just company names.
* Software for tools investigated in a CE will be provided in one or more separate branches of the software repository. Each CE will have a “fork” of the software, and within the CE there may be multiple branches established by the CE coordinator. The software coordinator will help coordinate the creation of these forks and branches and their naming. All JVET members will have read access to the CE software branches (using shared read-only credentials as described below).
* During the experiment, revisions of the experiment plans can be made, but not substantial changes to the proposed technology.
* The CE description must match the CE testing that is done. The CE description needs to be revised if there has been some change of plans.
* The CE summary report must describe any changes that were made in the process of finalizing the CE.
* By the next meeting it is expected that at least one independent cross-checker will report a detailed analysis of each proposed feature that has been tested and confirm that the implementation is correct. Commentary on the potential benefits and disadvantages of the proposed technology in cross-checking reports is highly encouraged. Having multiple cross-checking reports is also highly encouraged (especially if the cross-checking involves more than confirmation of correct test results). The reports of cross-checking activities may (and generally should) be integrated into the CE report rather than submitted as separate documents.
* It is mandatory to report encoder optimizations made for the benefit of a tool, and if an equivalent optimization could be applied on the anchor, a comparison against the improved anchor shall be provided.

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 system at:

<https://www.itu.int/ifa/t/2017/sg16/exchange/wp3/q06/vceg_account.txt>

Some agreements relating to CE activities were established as follows:

* Only qualified JVET members can participate in a CE.
* Participation in a CE is possible without a commitment of submitting an input document to the next meeting. Participation is requested by contacting the CE coordinator.
* All software, results, and documents produced in the CE should be announced and made available to JVET in a timely manner.
* A JVET CE reflector will be established and announced on the main JVET reflector. Discussion of logistics arrangements, exchange of data, minor refinement of the test plans, and preparation of documents shall be conducted on the JVET CE reflector, with subject lines prefixed by “[CEx: ]”, where “x” is the number of the CE. All substantial communications about a CE other than such details shall take place on main JVET reflector. In the case that large amounts of data are to be distributed, it is recommended to send a link to the data rather than the data itself, or upload the data as an input contribution to the next meeting.

General timeline for CEs

T1= 3 weeks after the JVET meeting: To revise the CE description and refine questions to be answered. Questions should be discussed and agreed on JVET reflector. Any changes of planned tests after this time need to be announced and discussed on the JVET reflector. Initially assigned description numbers shall not be changed later. If a test is skipped, it is to be marked as “withdrawn”.

T2 = Test model software release + 2 weeks: Integration of all tools into a separate CE branch of the VTM is completed and announced to JVET reflector.

* Initial study by cross-checkers can begin.
* Proponents may continue to modify the software in this branch until T3.
* 3rd parties are encouraged to study and make contributions to the next meeting with proposed changes

T3: 3 weeks before the next JVET meeting or T2 + 1 week, whichever is later: Any changes to the CE test branches of the software must be frozen, so the cross-checkers can know exactly what they are cross-checking. A software version tag should be created at this time. The name of the cross-checkers and list of specific tests for each tool under study in the CE plan description shall be documented in an updated CE description by this time.

T4: Regular document deadline minus 1 week: CE contribution documents including specification text and complete test results shall be uploaded to the JVET document repository (particularly for proposals targeting to be promoted to the draft standard at the next meeting).

The CE summary reports shall be available by the regular contribution deadline. This shall include documentation about crosscheck of software, matching of CE description and confirmation of the appropriateness of the text change, as well as sufficient crosscheck results to create evidence about correctness (crosscheckers must send this information to the CE coordinator at least 3 days ahead of the document deadline). Furthermore, any deviations from the timelines above shall be documented. The numbers used in the summary report shall not be changed relative to the description document.

CE reports may contain additional information about tests of straightforward combinations of the identified technologies. Such supplemental testing needs to be clearly identified in the report if it was not part of the CE plan.

New branches may be created which combine two or more tools included in the CE document or the VTM (as applicable).

It is not necessary to formally name cross-checkers in the initial version of the CE description document. To adopt a proposed feature at the next meeting, JVET would like to see comprehensive cross-checking done, with analysis of whether the description matches the software, and a recommendation of the value of the tool and given tradeoffs.

The establishment of a CE does not indicate that a proposed technology is mature for adoption or that the testing conducted in the CE is fully adequate for assessing the merits of the technology, and a favourable outcome of CE does not indicate a need for adoption of the technology into a standard or test model.

Availability of specification text is important to have a detailed understanding of the technology and also to judge what its impact on the complexity of the specification will be. There must also be sufficient time to study this in detail. CE contributions without sufficiently mature draft specification text in the CE input document should not be considered for adoption.

Lists of participants in CE documents should be pruned to include only the active participants. Read access to software will be available to all members.

# Establishment of ad hoc groups

The ad hoc groups established to progress work on particular subject areas until the next meeting are described in the table below. The discussion list for all of these ad hoc groups was agreed to be the main JVET reflector ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de)). The previously approved rules for MPEG ad hoc groups established in document [SC29/AG2 N 46](https://www.mpegstandards.org/wp-content/uploads/2022/01/ISO-IECJTC1-SC29-AG2_N0046_AhG.pdf) were agreed to apply to these ad hoc groups.

Review of AHG plans was conducted during the closing plenary on Friday 29 April 2022 at 0500 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-Z1003, JVET-Z1005, JVET-Z1008, JVET-Z2005, and JVET-Z2006). * 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-Z1004 errata output collection. * Produce and finalize JVET-Z2002 VVC Test Model 17 (VTM 17) Algorithm and Encoder Descriptions. * Coordinate with the test model software development AhG to address issues relating to mismatches between software and text. * Collect and consider errata reports on the texts. | B. Bross, C. Rosewarne (co-chairs), F. Bossen, J. Boyce, A. Browne, S. Kim, S. Liu, J.‑R. Ohm, G. J. Sullivan, A. Tourapis, Y.-K. Wang, Y. Ye (vice-chairs) | N |
| **Test model software development (AHG3)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Coordinate development of test models (VTM, HM, SCM, SHM, HTM, MFC, MFCD, JM, JSVM, JMVM, 3DV-ATM, 360Lib, and HDRTools) software and associated configuration files. * Produce documentation of software usage for distribution with the software. * Enable software support for recently standardized additional SEI messages. * Discuss and make recommendations on the software development process. * Propose improvements to the guideline document for developments of the test model software. * Perform comparative tests of test model behaviour using common test conditions. * Suggest configuration files for additional testing of tools. * Investigate how to minimize the number of separate codebases maintained for group reference software. * Coordinate with AHG on Draft text and test model algorithm description editing (AHG2) to identify any mismatches between software and text, and make further updates and cleanups to the software as appropriate. * Prepare drafts of merged CTC documents for HM and VTM, as applicable. | F. Bossen, X. Li, K. Sühring (co-chairs), Y. He, K. Sharman, V. Seregin, A. Tourapis (vice‑chairs) +HDR | N |
| **Test material and visual assessment (AHG4)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Consider plans for additional verification testing of VVC capability, particularly target establishing a test plan for VVC scalability features by the next meeting. * Maintain the video sequence test material database for testing the VVC and HEVC standards and potential future extensions, as well as exploration activities. * Study coding performance and characteristics of available and proposed video test material. * Identify and recommend appropriate test material for testing the VVC standard and potential future extensions, as well as exploration activities. * Identify and characterize missing types of video material, solicit contributions, collect, and make available a variety of video sequence test material. * Maintain and update the directory structure for the test sequence repository as necessary. * Collect information about test sequences that have been made available by other organizations. * Prepare and conduct remote expert viewing for purposes of subjective quality evaluation. * Coordinate with AG 5 in studying and developing further methods of subjective quality evaluation, e.g. based on crowd sourcing. * Prepare availability of viewing equipment and facilities arrangements for future meetings. * Prepare subj. test … | V. Baroncini, T. Suzuki, M. Wien (co-chairs), S. Liu, G. Martin-Cocher, A. Segall, P. Topiwala, S. Wenger, J. Xu, Y. Ye (vice-chairs) | Y (2 weeks notice) |
| **Conformance testing (AHG5)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the JVET-Y2026 draft conformance testing for operation rage extensions and investigate the need for improvements. * Study the requirements of VVC, HEVC, and AVC conformance testing to ensure interoperability. * Maintain and update the conformance bitstream database. * Study additional testing methodologies to fulfil the needs for VVC conformance testing. | D. Rusanovskyy, I. Moccagatta (co-chairs), F. Bossen, K. Kawamura, T. Hashimoto, 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-5.0 software version and the reference configuration encodings according to JVET-Y2017 common test conditions. * Investigate encoder speedup and other encoder software optimization. * Coordinate with ECM algorithm description editors to identify any mismatches between software and text, make further updates and cleanups to the software as appropriate. * Prepare subj. test … | V. Seregin (chair), J. Chen, F. Le Léannec, K. Zhang (vice-chairs) | Y (2 weeks notice) |
| **Low latency and constrained complexity (AHG7)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Identify additional application scenarios and their requirements for low latency and constrained complexity, taking into account aspects of real-time encoding and decoding. * Discuss requirements of already identified scenario such as cloud gaming, game casting, video conferencing, video surveillance and remote control of systems. * Evaluate and refine new CTC for low latency and constrained complexity application scenarios, and investigate a set of tools that provide a reasonable tradeoff regarding complexity vs. compression, as well as latency constraints. * Conduct tests with ECM and VTM to determine the impact of discussed configurations on coding efficiency and run time, and conduct visual tests in coordination with AHG4, including the support for timely availability of test materials and test subjects. * Review current test sequences and if necessary collect new test materials that are suitable for the intended application domains, and establish an applicable dataset in coordination with AHG4. * Coordinate with AHG3 and AHG12 to discuss and recommend configuration(s) applicable to ECM and VTM, taking into account complementarity with existing CTCs. | A. Duenas, T. Poirier, S. Liu (co-chairs), L. Wang, J. Xu (vice-chairs) | N |
| **High bit depth, high bit rate, and high frame rate coding (AHG8)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the benefits and characteristics of VVC coding tools for high bit depth, high bit rate, and high frame rate coding. * Study the JVET-U2018 testing conditions for high bit depth, high bit rate, and high frame rate coding, and suggest improvements as applicable, and coordinate with AHG3 towards a combination with JCTVC-AF1100. * Contribute to the development of software and conformance testing for operation range extensions in coordination with AHG3 and AHG5. | A. Browne, T. Ikai (co-chairs), D. Rusanovskyy, X. Xiu, Y. Yu (vice-chairs) | N |
| **SEI message studies (AHG9)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the SEI messages in VSEI, VVC, HEVC and AVC. * Collect software and showcase information for SEI messages, including encoder and decoder implementations and bitstreams for demonstration and testing. * Identify potential needs for additional SEI messages. * Investigate the possible need of mandatory post processing in the context of SEI messages * Study SEI messages defined in HEVC and AVC for potential use in the VVC context. * Coordinate with AHG3 for software support of SEI messages. | S. McCarthy, Y.-K. Wang (co-chairs), T. Chujoh, S. Deshpande, C. Fogg, P. de Lagrange, G. J. Sullivan, A. Tourapis, S. Wenger (vice-chairs) | N |
| **Encoding algorithm optimization (AHG10)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the impact of using techniques such as tool adaptation and configuration, and perceptually optimized adaptive quantization for encoder optimization. * Study the impact of non-normative techniques of pre processing for the benefit of encoder optimization. * Study encoding techniques of optimization for objective quality metrics and their relationship to subjective quality. * Study optimized encoding for reference picture resampling and scalability modes in VTM. * Consider neural network-based encoding optimization technologies for video coding standards. * Investigate other methods of improving objective and/or subjective quality, including adaptive coding structures and multi-pass encoding. * Study methods of rate control and rate-distortion optimization and their impact on performance, subjective and objective quality. * Study the potential of defining default or alternate software configuration settings and test conditions optimized for either subjective quality, or higher objective quality, and coordinate such efforts with AHG3 and AHG6. * Study the effect of varying configuration parameters depending on temporal layer, such as those related to deblocking, partitioning, chroma QP. | P. de Lagrange, A. Duenas, R. Sjöberg, A. Tourapis (co-chairs) | N |
| **Neural network-based video coding (AHG11)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Evaluate and quantify the performance improvement potential of NN-based video coding technologies compared to existing video coding standards such as VVC, including both individual coding tools and novel architectures. * Refine the test conditions for NN-based video coding. Generate and distribute anchor encoding, and develop supporting software as needed. * Study the impact of training (including the impact of loss function) on the performance of candidate technologies, and identify suitable materials for testing and training. * Analyse complexity characteristics, perform complexity analysis, and develop complexity reductions of candidate technology. * Study and maintain the SADL (Small Adhoc Deep-Learning Library). Identify gaps in functionality and develop improvements as needed. * Finalize and discuss the EE on neural network-based video coding. * Coordinate with other relevant groups, including SC29/AG5 on the evaluation and assessment of visual quality and AHG12 on the interaction with ECM coding tools. In case of conducting subjective viewing, support shall be given to AHG4 for timely availability of test materials and test subjects. * Investigate common software for development and verification NN-based video coding technologies. | E. Alshina, S. Liu, A. Segall (co‑chairs), F. Galpin, J. Pfaff, S. S. Wang, Z. Wang, M. Wien, P. Wu, J. Xu (vice‑chairs) | N |
| **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 ECM5 algorithm description JVET-Z2025. * 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-Z2024 in coordination with the EE coordinators. * Coordinate with AHG11 to study the interaction with neural network-based coding tools. * Prepare subj. test … | M. Karczewicz, Y. Ye, L. Zhang (co-chairs), B. Bross, X. Li, K. Naser, H. Yang (vice-chairs) | Y (2 weeks notice) |
| **Film grain technologies (AHG13)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the benefits and characteristics of film grain technologies, including autoregressive and frequency-filtering technologies. * Discuss the JVET-Y2020 draft of the Technical Report on Film grain synthesis technology for video applications, and suggest improvements as necessary. * Study alternative film grain models and their associated documentation. * Study preprocessing and encoder technologies for determining values for FGC (Film Grain Characteristics) SEI message syntax elements. * Identify potential need for additional film grain technology and signalling, if needed. * Study categorization and/or classification of FGC implementations. * Study the implication of identifying FGC reference implementations. * Coordinate development of film grain technology software and configuration files. * Coordinate with AHG3 for software support of the FGC SEI message. | W. Husak, M. Radosavljević, W. Wan (co-chairs), D. Grois, Y. He, P. de Lagrange, A. Segall, A. Tourapis (vice-chairs) | Y (2 weeks notice) |

It was confirmed that the rules which can be found in document ISO/IEC JTC 1/‌SC 29/‌AG 2 [N 046](https://www.mpegstandards.org/wp-content/uploads/2022/01/ISO-IECJTC1-SC29-AG2_N0046_AhG.pdf) “Ad hoc group rules for MPEG AGs and WGs” (available at <https://www.mpegstandards.org/adhoc/>), are consistent with the operation mode of JVET AHGs. It is pointed out that JVET does not maintain separate AHG reflectors, such that any JVET member is implicitly a member of any AHG. This shall be mentioned in the related WG Recommendations. The list above was also issued as a separate WG 5 document (ISO/IEC JTC 1/‌SC 29/‌WG 5 [N 136](https://dms.mpeg.expert/doc_end_user/current_document.php?id=82006&id_meeting=189)) in order to make it easy to reference.

# Output documents

The following documents were agreed to be produced or endorsed as outputs of the meeting. Names recorded below indicate the editors responsible for the document production. Where applicable, dates of planned finalization and corresponding parent-body document numbers are also noted.

It was reminded that in cases where the JVET document is also made available as a WG 5 output document, a separate version under the WG 5 document header should be generated. This version should be sent to GJS and JRO for upload.

The list of JVET ad hoc groups was also issued as a WG 5 output document WG 5 [N 136](https://dms.mpeg.expert/doc_end_user/current_document.php?id=82006&id_meeting=189), as noted in section 9.

[JVET-AA1000](https://jvet-experts.org/doc_end_user/current_document.php?id=11704) Meeting Report of the 27th JVET Meeting [J.-R. Ohm] [WG 5 N 140] (2022-08-19)

Initial versions of the meeting notes (d0 … d8) were made available on a daily basis during the meeting.

Remains valid – not updated: [JCTVC-H1001](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=5095) HEVC software guidelines [K. Sühring, D. Flynn, F. Bossen (software coordinators)]

Remains valid – not updated: [JVET-Y1002](https://jvet-experts.org/doc_end_user/current_document.php?id=11463) High Efficiency Video Coding (HEVC) Test Model 16 (HM 16) Encoder Description Update 16 [C. Rosewarne (primary editor), K. Sharman, R. Sjöberg, G. J. Sullivan (co-editors)] [WG 5 [N 103](https://dms.mpeg.expert/doc_end_user/current_document.php?id=82085&id_meeting=189)]

[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 CD N 147] [G. J. Sullivan, A. Tourapis] (2022-08-XX)

New JVET version, or just convert Draft 1 into CD?

This includes identifiers for YCoCg-R colour representation with equal luma and chroma bit depths.

A request for a new edition (WG 5 N 131) was reviewed Friday 29 April at 0700 UTC

Target dates: CD July 2022, DIS October 2022, FDIS July 2023.

[JVET-AA1004](https://jvet-experts.org/doc_end_user/current_document.php?id=11706) Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR [B. Bross, I. Moccagatta, C. Rosewarne, G. J. Sullivan, Y. Syed, Y.-K. Wang] (2022-09-30, near next meeting)

New: AVC on MV range, VVC conformance V1, …

Remains valid – not updated: [JVET-Z1005](https://jvet-experts.org/doc_end_user/current_document.php?id=11707) New levels for HEVC (Draft 3) [T. Suzuki, A. Tourapis, Y.-K. Wang] (2022-05-06)

A DoCR for ballot responses on CDAM 2 (WG 5 N 127) was reviewed Friday 29 April at 0635 UTC.

This was integrated into the DIS of HEVC 5th edition (WG 5 N 128), targeting FDIS and ITU-T consent in October, and a recommendation was formulated to integrate the previous Amd.2 into a new edition. A request for the new edition was made in recommendation 3.1.2 of WG 5 (see annex C).

Post-meeting note: In post-meeting consultation with the SC 29 Committee Manager, it was clarified that the approval of the plan to produce a new edition rather than an amendment will cause a delay in the processing of this text for ballot on the ISO/IEC side, such that the timeline envisioned during the meeting is not feasible, thus probably resulting in a delay of one meeting cycle for the ISO/IEC approval process. This does not necessarily affect the ITU-T approval timeline.

Upon consideration of NB comment US-006 on the CDAM 2 ballot, it was decided to increase the maximum frame rate to 960 fps (same value as in the corresponding VVC levels) in the high tier for levels 7.0 and higher.

Remains valid – not updated [JVET-T1006](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10538) Annotated regions and shutter interval information SEI messages for AVC (Draft 2) [J. Boyce, S. McCarthy, Y.-K. Wang]

Remains valid – not updated: [JCTVC-V1007](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=10312) SHVC Test Model 11 (SHM 11) Introduction and Encoder Description [G. Barroux, J. Boyce, J. Chen, M. M. Hannuksela, Y. Ye] [WG 11 N 15778]

Remains valid – not updated: [JVET](https://jvet-experts.org/doc_end_user/current_document.php?id=11708)-Z1008 Additional colour type identifiers for AVC and HEVC (Draft 1) [G. J. Sullivan, A. Tourapis] (2022-06-03)

This includes identifiers for YCoCg-R colour representation with equal luma and chroma bit depths.

Remains valid – not updated: [JCTVC-X1009](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=10572) Common Test Conditions for SHVC [V. Seregin, Y. He]

Remains valid – not updated [JCTVC-O1010](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=8511) Guidelines for Conformance Testing Bitstream Preparation [T. Suzuki, W. Wan]

No output: JVET-T1011 through JVET-T1013

JVET-AA1011 HEVC multiview profiles supporting extended bit depth (draft 1) [A. Tourapis, W. Husak] [WG 5 preliminary WD N 143] (2022-08-XX)

From JVET-AA0239

Remains valid – not updated [JCTVC-V1014](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=10316) Screen Content Coding Test Model 7 Encoder Description (SCM 7) [R. Joshi, J. Xu, R. Cohen, S. Liu, Y. Ye] [WG 11 N 16049]

Remains valid for HM – not updated: [JCTVC-Z1015](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=10689) Common Test Conditions for Screen Content Coding [H. Yu, R. Cohen, K. Rapaka, J. Xu]

No output: JVET-X1016 through JVET-X1019

[JVET-AA1100](https://jvet-experts.org/doc_end_user/current_document.php?id=11466) Common Test Conditions for HM Video Coding Experiments [K. Sühring, K. Sharman] (2022-08-19)

CTC for high bit depth and high bit rate to be removed, unified in JVET-AA2018

This specifies only the CTC for non-4:2:0 colour formats. The corresponding document for VVC is JVET-T2013, no unification yet

Remains valid – not updated: [JVET-T2001](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10540) Versatile Video Coding Draft 10 [B. Bross, J. Chen, S. Liu, Y.-K. Wang]

Remains valid – not updated: [JVET-Z2002](https://jvet-experts.org/doc_end_user/current_document.php?id=11709) Algorithm description for Versatile Video Coding and Test Model 17 (VTM 17) [A. Browne, Y. Ye, S. Kim] [WG 5 N 130] (2022-06-30, near next meeting)

No new elements from decisions of the current meeting. Description of SEI (in particular for the SEI processing order) should be added after availability of reference SW (next meeting)

It is noted that the list above may not be complete; if some adoption is missing that is recorded somewhere else in the meeting notes it shall also be considered included.

Remains valid – not updated: [JVET-N1003](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=6638) Guidelines for VVC reference software development [K. Sühring]

Remains valid – not updated: [JVET-T2004](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10542) Algorithm descriptions of projection format conversion and video quality metrics in 360Lib (Version 12) [Y. Ye, J. Boyce]

[JVET-AA2005](https://jvet-experts.org/doc_end_user/current_document.php?id=11710) New level and systems-related supplemental enhancement information for VVC (Draft 3) [B. Bross, E. François, A. Tourapis, Y.-K. Wang] [WG 5 DAM 1 N 145] (2022-XX-XX)

A DoCR for ballot responses on CDAM 1 (WG 5 N 144) was reviewed Friday 22 July at XXXX UTC.

See notes under JVET-AA0048.

The corresponding SEI message indicators relating to JVET-AA2006 are to be included.

[JVET-AA2006](https://jvet-experts.org/doc_end_user/current_document.php?id=11711) Additional SEI messages for VSEI (Draft 2) [S. McCarthy, T. Chujoh, M. M. Hannuksela, G. J. Sullivan, Y.-K. Wang] [WG 5 CDAM N 141] (2022-XX-XX)

Elements from JVET-AA0054, JVET-AA0055, JVET-AA0056, JVET-AA0067, JVET-AA0100, JVET-AA0110 are added.

The corresponding variables in VVC are included in JVET-AA2005.

It was requested that VTM software implementations of the new SEI messages shall be made available by the next meeting.

Remains valid – not updated: [JVET-S2007](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9679) Versatile supplemental enhancement information messages for coded video bitstreams Draft 5 [J. Boyce, V. Drugeon, G. J. Sullivan, Y.-K. Wang]

Remains valid – not updated: [JVET-X2008](https://jvet-experts.org/doc_end_user/current_document.php?id=11228) Conformance testing for versatile video coding (Draft 7) [J. Boyce, F. Bossen, K. Kawamura, I. Moccagatta, W. Wan]

Remains valid – not updated: [JVET-Y2009](https://jvet-experts.org/doc_end_user/current_document.php?id=11470) Reference software for versatile video coding (Draft 3) [F. Bossen, K. Sühring, X. Li]

This had been issued as ISO/IEC FDIS 23090-16 as WG 5 [N 112](https://dms.mpeg.expert/doc_end_user/current_document.php?id=82000&id_meeting=189), and submitted for ITU-T consent.

Remains valid – not updated: [JVET-Y2010](https://jvet-experts.org/doc_end_user/current_document.php?id=11471) VTM and HM common test conditions and software reference configurations for SDR 4:2:0 10 bit video [F. Bossen, X. Li, V. Seregin, K. Sharman, K. Sühring]

Remains valid – not updated: [JVET-Z2011](https://jvet-experts.org/doc_end_user/current_document.php?id=11712) VTM and HM common test conditions and evaluation procedures for HDR/WCG video [A. Segall, E. François, W. Husak, S. Iwamura, D. Rusanovskyy] (2022-05-13)

This includes a merge of HM/VTM CTC of HDR as per JVET-Z0175. No need to review this during the closing plenary was identified.

Remains valid – not updated: [JVET-U2012](https://jvet-experts.org/doc_end_user/current_document.php?id=10681) JVET common test conditions and evaluation procedures for 360° video [Y. He, J. Boyce, K. Choi, J.-L. Lin] (2021-03-31)

Remains valid – not updated: [JVET-T2013](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10546) VTM common test conditions and software reference configurations for non-4:2:0 colour formats [Y.-H. Chao, Y.-C. Sun, J. Xu, X. Xu]

Remains valid – not updated: [JVET-Q2014](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9683) JVET common test conditions and software reference configurations for lossless, near lossless, and mixed lossy/lossless coding [T.-C. Ma, A. Nalci, T. Nguyen]

Remains valid – not updated: [JVET-Q2015](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9684) JVET functionality confirmation test conditions for reference picture resampling [J. Luo, V. Seregin]

[JVET-AA2016](https://jvet-experts.org/doc_end_user/current_document.php?id=11713) Common Test Conditions and evaluation procedures for neural network-based video coding technology [E. Alshina, R.-L. Liao, S. Liu, A. Segall] (2022-08-05)

Update in the cross-check of training procedures as suggested in the BoG report JVET-AA0247. Initial version reviewed during the closing plenary.

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]

May need update in case of a new optional class G, as well as JVET-Y2010.

[JVET-AA2018](https://jvet-experts.org/doc_end_user/current_document.php?id=10683) Common test conditions for high bit depth and high bit rate video coding [A. Browne, T. Ikai, D. Rusanovskyy, M. Sarwer, X. Xiu]

Unified for VTM and HM

No output: JVET-X2019

[JVET-AA2020](https://jvet-experts.org/doc_end_user/current_document.php?id=11475) Film grain synthesis technology for video applications (Draft 2) [D. Grois, Y. He, W. Husak, A. Norkin, M. Radosavljević, A. Tourapis, W. Wan] [WG 5 DTR [N 142](https://dms.mpeg.expert/doc_end_user/current_document.php?id=82207&id_meeting=189)]

WG 5 recommendation of adding A. Norkin as editor.

See discussion about the proposed draft JVET-AA0051.

No outputs: JVET-X2021, JVET-X2022

These numbers are retained for future purposes of planning possible additional VVC verification testing reports and other purposes.

[JVET-AA2023](https://jvet-experts.org/doc_end_user/current_document.php?id=11701) Exploration Experiment on Neural Network-based Video Coding (EE1) [E. Alshina, W. Chen, F. Galpin, Y. Li, Z. Ma, H. Wang, L. Wang] [WG 5 N 148] (2022-08-05)

An initial draft of this document was reviewed and approved at XXXX UTC on Friday 22 July.

[JVET-AA2024](https://jvet-experts.org/doc_end_user/current_document.php?id=11703) Exploration Experiment on Enhanced Compression beyond VVC capability (EE2) [V. Seregin, J. Chen, G. Li, K. Naser, J. Ström, M. Winken, X. Xiu, K. Zhang] [WG 5 N 149] (2022-08-19)

An initial draft of this document was reviewed and approved on Friday 22 July at XXXX UTC.

Categories are intra prediction, inter prediction, screen content coding, transforms, and in-loop filters.

It was discussed if it would be possible to provide software for proposals included in EE earlier than T3 (based on ECM4 then). It is however argued that this would create additional effort in creating duplicate branches in the repository.

It is further mentioned that there might be problems in imposing such a rule in “ad hoc” fashion at a meeting, as some companies might not have expected it and need time to clear software disclosure before providing it.

It is generally understood that for transparency reasons, contributors of technology are always encouraged to provide software openly. At this stage of exploration, it however appears difficult to make that mandatory, as it might even happen that proposals are withdrawn from EE if no clearance of disclosure is reached within a company.

[JVET-AA2025](https://jvet-experts.org/doc_end_user/current_document.php?id=11714) Algorithm description of Enhanced Compression Model 6 (ECM 6) [M. Coban, F. Le Léannec, K. Naser, J. Ström, L. Zhang] [WG 5 N 150] (2022-09-30)

New elements from notes elsewhere in this report:

* From BoG JVET-Z0210: Include the following proposal in the next release of ECM: JVET-Z0131, JVET-Z0127 (option 2)
* Decision: Adopt JVET-Z0050 (Test 1.3b)
* Decision: Adopt JVET-Z0054 (Test 2.1b)
* Decision: Adopt JVET-Z0136 (test EE2 – 2.2a)
* Decision: Adopt JVET-Z0061 (test EE2 – 2.3)
* Decision: Adopt JVET-Z0117, version where no switching between 4-tap and 6-tap filters depending on sequence is applied (no SPS flag, but retain configurability via parameter file or macro
* Decision: Adopt JVET-Z0118 (software, and potentially ECM description, as far as there are deviations from the VVC GDR)
* Decision: Adopt JVET-Z0056 (test EE2 – 2.4)
* Decision: Adopt JVET-Z0139 (version EE2-2.7c)
* Decision: Adopt JVET-Z0139 (version EE2-2.7c)
* Decision: Adopt JVET-Z0153 (3.2), JVET-Z0075 (3.3), JVET-Z0084 (3.4), JVET-Z0160 (3.5b)
* Decision: Adopt JVET-Z0135, Test 4.3b
* Decision: Adopt JVET-Z0085
* Decision: Adopt JVET-Z0102
* Decision: Adopt JVET-Z0105 (not CTC)
* Potential impact of JVET-Z0072 and JVET-Z0099, as applicable (see notes above for VTM JVET-Z2002)

It is noted that the list above may not be complete; if some adoption is missing that is recorded somewhere else in the meeting notes it shall also be considered included.

Remains valid – not updated: [JVET-Y2026](https://jvet-experts.org/doc_end_user/current_document.php?id=11477) Conformance testing for VVC operation range extensions (Draft 3) [WG 5 DAM [N 110](https://dms.mpeg.expert/doc_end_user/current_document.php?id=81998&id_meeting=189)] [D. Rusanovskyy, T. Hashimoto, H.-J. Jhu, I. Moccagatta, Y. Yu]

Editorial improvements are recorded in JVET-AA0109, no need for a new output at the currentmeeting.

JVET-AA2027 SEI processing order SEI message in VVC (draft 1) [S. McCarthy, M. Hannuksela, Y.-K. Wang] [WG 5 preliminary WD N 146] (2022-08-26)

From JVET-AA0102.

It was requested that a VTM software implementations of the SEI message shall be made available by the next meeting.

# Future meeting plans, expressions of thanks, and closing of the meeting

Future meeting plans were established according to the following guidelines (assuming face-to-face meetings):

* Meeting under ITU-T SG 16 auspices when it meets (ordinarily starting meetings on the Wednesday of the first week and closing it on the Wednesday of the second week of the SG 16 meeting – a total of 8 meeting days), and
* Otherwise meeting under ISO/IEC JTC 1/‌SC 29 auspices when its MPEG WGs meet (ordinarily starting meetings on the Friday prior to the main week of such meetings and closing it on the same day as other MPEG WGs – a total of 8 meeting days).

In cases where an exceptionally high workload is expected for a meeting, an earlier starting date may be defined. In cases of online meetings, no sessions should be held on weekend days. This may imply an earlier starting date as well.

Some specific future meeting plans (to be confirmed) were established as follows:

* Fri. 21 – Fri. 28 October 2022, 28th meeting under ITU-T SG16 auspices in XXXX, XX.
* Wed. 11 – Fri. 13 and Mon. 16 – Fri. 20 January 2023, 29th meeting under ISO/IEC JTC 1/‌SC 29 auspices, to be held as teleconference meeting.
* During Wed. 19 – Fri. 28 April 2023, 30th meeting under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.
* During July 2023, 31st meeting under ITU-T SG16 auspices, date and location t.b.d.
* During October 2023, 32nd meeting under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.
* During January 2024, 33rd meeting under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.
* During April 2024, 34th meeting under ITU-T SG16 auspices, date and location t.b.d.
* During July 2024, 34th meeting under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.

The agreed document deadline for the 27th JVET meeting was planned to be XXday XX October 2022.

Mathias Wien and Johannes Sauer were thanked for planning, organizing and conducting the remote expert viewings related to the exploration experiment on neural network-based video compression, and to assessment of new test sequences in the category of screen/gaming content.

Alibaba, InterDigital, and Youku were thanked for offering new test materials that can be used for developing and testing video technology standards.

Thanks were expressed to Christian Tulvan for his engagement in maintaining the site jvet-experts.org. Institut Mines-Télécom was thanked for hosting the site.

It was suggested that in a future meeting, perspectives should be discussed for the ongoing exploration, in terms of potentially developing standardization projects and realistic timelines. From the current status of the JVET-internal explorations, there does not seem to be sufficient evidence to embark on standardization soon. Potential requirements need also to be discussed with the parent bodies.

The 27th JVET meeting was closed at approximately 00XX hours UTC on Saturday 23 July 2022.

# Annex A to JVET report: List of documents

# Annex B to JVET report: List of meeting participants

The participants of the twenty-seventh meeting of the JVET, according to the participation records from the Zoom teleconferencing tool used for the meeting sessions (approximately XXX people in total, not including those who attended only the joint sessions with other groups), were as follows:



# Annex C to JVET report: Recommendations of the 8th 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 139**

**Recommendations of the 8th WG 5 meeting**

1. [] L. Wang, X. Xu, S. Liu, F. Galpin. EE1-1.1: neural network based in-loop filter with 2 models. JVET- Z0094, 26th Meeting, by teleconference, 20–29 April 2022. [↑](#endnote-ref-2)