|  |  |
| --- | --- |
| **Joint Video Experts Team (JVET)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29**  26th Meeting, by teleconference, 20–29 April 2022 | Document: JVET-Z\_Notes\_d9 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Meeting Report of the 26th Meeting of the Joint Video Experts Team (JVET), by teleconference, 20–29 April 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-sixth meeting during 20–29 April 2022 as an online-only meeting. It had previously been planned to be in Alpbach, AT, but this plan was changed due to the difficulties 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 seventh 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 20 April 2022. Meeting sessions were held on all days except the weekend days of Saturday and Sunday 23 and 24 April 2022, until the meeting was closed at approximately 0020 hours UTC on Saturday 30 April 2022. Approximately XXX people attended the JVET meeting, and approximately XXX input documents (not counting crosschecks), 13 AHG reports, 2 EE summary reports, 2 BoG reports, and 1 report on expert viewing conducted prior to the meeting 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-fifth JVET meeting in producing the following documents:

* JVET-Y1002 High Efficiency Video Coding (HEVC) Test Model 16 (HM 16) Encoder Description Update 16
* JVET-Y1004 Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR
* JVET-Y1005 New levels for HEVC (Draft 2), also issued as WG 5 CDAM
* JVET-Y1100 Common Test Conditions for HM Video Coding Experiments
* JVET-Y2002 Algorithm description for Versatile Video Coding and Test Model 16 (VTM 16)
* JVET-Y2005 VVC operation range extensions (Draft 6), also integrated into VVC version 2 which was submitted as WG 5 FDIS and for ITU consent
* JVET-Y2006 Additional SEI messages for VSEI (Draft 6), also integrated into VSEI version 2 which was submitted as WG 5 FDIS and for ITU consent
* JVET-Y2009 Reference software for versatile video coding (Draft 3), also submitted as WG 5 FDIS and for ITU consent
* JVET-Y2010 VTM common test conditions and software reference configurations for SDR video
* JVET-Y2011 VTM common test conditions and evaluation procedures for HDR/WCG video
* JVET-Y2017 Common test conditions and evaluation procedures for enhanced compression tool testing
* JVET-Y2019 New level and systems-related supplemental enhancement information for VVC (Draft 1)
* JVET-Y2020 Film grain synthesis technology for video applications (Draft 1)
* JVET-Y2023 Exploration Experiment on neural network-based video coding (EE1)
* JVET-Y2024 Exploration Experiment on enhanced compression beyond VVC capability (EE2)
* JVET-Y2025 Algorithm description of Enhanced Compression Model 4 (ECM 4)
* JVET-Y2026 Conformance testing for VVC operation range extensions (Draft 3), also issued as WG 5 DAM.

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

* 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 color 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-
* JVET-Z2011 VTM 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 Wed. 13 – Fri. 15 and Mon. 18 – Fri. 22 July 2022 under ISO/IEC JTC 1/‌SC 29 auspices, to be conducted as a teleconference meeting; during 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.; and during April 2024 under ITU-T SG16 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 site <http://phenix.int-evry.fr/jvet/> is still accessible, but was 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 the 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-sixth meeting during 20–29 April 2022 as an online-only meeting, using Zoom teleconferencing tools. For ISO/IEC purposes, JVET is alternatively designated ISO/IEC JTC 1/‌SC 29/‌WG 5, and this was the seventh 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_04_Z_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-Y1002 High Efficiency Video Coding (HEVC) Test Model 16 (HM 16) Encoder Description Update 16
* JVET-Y1004 Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR
* JVET-Y1005 New levels for HEVC (Draft 2), also issued as WG 5 CDAM
* JVET-Y1100 Common Test Conditions for HM Video Coding Experiments
* JVET-Y2002 Algorithm description for Versatile Video Coding and Test Model 16 (VTM 16)
* JVET-Y2005 VVC operation range extensions (Draft 6), also integrated into VVC version 2 which was submitted as WG 5 FDIS and for ITU consent
* JVET-Y2006 Additional SEI messages for VSEI (Draft 6), also integrated into VSEI version 2 which was submitted as WG 5 FDIS and for ITU consent
* JVET-Y2009 Reference software for versatile video coding (Draft 3), also submitted as WG 5 FDIS and for ITU consent
* JVET-Y2010 VTM common test conditions and software reference configurations for SDR video
* JVET-Y2011 VTM common test conditions and evaluation procedures for HDR/WCG video
* JVET-Y2017 Common test conditions and evaluation procedures for enhanced compression tool testing
* JVET-Y2019 New level and systems-related supplemental enhancement information for VVC (Draft 1)
* JVET-Y2020 Film grain synthesis technology for video applications (Draft 1)
* JVET-Y2023 Exploration Experiment on neural network-based video coding (EE1)
* JVET-Y2024 Exploration Experiment on enhanced compression beyond VVC capability (EE2)
* JVET-Y2025 Algorithm description of Enhanced Compression Model 4 (ECM 4)
* JVET-Y2026 Conformance testing for VVC operation range extensions (Draft 3), also issued as WG 5 DAM.

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 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 and editorial 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, 13 April 2022. Any documents uploaded after 1159 hours Paris/Geneva time on Thursday 14 April 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-Z0149 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 was 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 was 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-Z0077 (a proposal on an end-to-end method in NN-based video coding), uploaded 04-18,
* JVET-Z0124 (a proposal on spatial GPM), uploaded 04-15,
* JVET-Z0164 (a proposal on YCoCg-R identifier for CICP), uploaded 04-18,
* JVET-Z0184 (a proposal on IBC merge/AMVP list construction), uploaded 04-20,
* JVET-Z0219 (a proposal on SPS flag for TM-based tools), uploaded 04-21,
* JVET-Z0244 (a proposal on NN-based post-filter SEI), uploaded 04-27,
* JVET-Z0249 (a proposal on a fix for a ticket), uploaded 04-28,
* JVET-Z0250 (a proposal on a fix for a ticket), uploaded 04-28.

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

* JVET-Z0114 (a document on CTC for low-latency in cloud gaming applications), uploaded 04-14,
* JVET-Z0116 (a document on CTC for low-latency in cloud gaming applications), uploaded 04-14,
* JVET-Z0141 (a document on GDR implementation for ECM), uploaded 04-20,
* JVET-Z0175 (a document on merging CTCs for HDR video), uploaded 04-20,
* JVET-Z0209 (a document on encoder optimization for 3D-HEVC), uploaded 04-20.

All cross-verification reports at this meeting except for JVET-Z0071 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, neither were they provided within 2 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-Z0076, JVET-Z0080, JVET-Z0081, JVET-Z0090, JVET-Z0148, JVET-Z0168, JVET-Z0172, … .

“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-Z0124 and JVET-Z0141, which were both categorized as late in the list above, depending on 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-Y1000, the High Efficiency Video Coding (HEVC) Test Model 16 (HM 16) encoder description update 16 JVET-Y1002, the Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR JVET-Y1004, the New levels for HEVC (Draft 1) JVET-Y1005, the Common Test Conditions for HM video coding experiments JVET-Y1100, the Algorithm description for Versatile Video Coding and Test Model 16 (VTM 16) JVET-Y2002, the Operation range extensions for VVC (Draft 6) JVET-Y2005, the Additional SEI messages for VSEI (Draft 6) JVET-Y2006, the Reference software for Versatile Video Coding (draft 3) JVET-Y2008, the VTM common test conditions and software reference configurations for SDR video JVET-Y2010, the VTM common test conditions and evaluation procedures for HDR/WCG video JVET-Y2011, the Common Test Conditions and evaluation procedures for enhanced compression tool testing JVET-Y2017, the New level and systems-related supplemental enhancement information for VVC (Draft 1) JVET-Y2019, the Film grain synthesis technology for video applications (Draft 1) JVET-Y2020, 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, the Algorithm description of Enhanced Compression Model 4 (ECM 4) JVET-Y2026, and the Conformance testing for VVC operation range extensions (Draft 3) JVET-Y2026, had been completed and were approved. It was noted that JVET-Y2020 was delivered only shortly before the meeting, and any concerns should be raised after studying it. No such concerns were raised later during the meeting. It was noted that the WG 5 output document N 98 (AVC 10th edition text) is still in the phase of editorial finalization. The software implementations of VTM (version 16.0), ECM (version 4.0) were also approved.

Only minor editorial issues were found in the meeting report JVET-Y1000 – 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, as 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.
* Avoid overloading people’s internet connections by not using video for the teleconferencing calls – only voice and screen sharing. Extensive use of screen sharing is encouraged.

## 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
* Code of conduct policy reminder
* IPR policy reminder and declarations
* Contribution document allocation
* Review of results of the previous meeting
* 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
* Coordination of visual quality testing
* Coordination activities with other organizations
* 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, meeting planning, other planning issues
* Other business as appropriate for consideration.

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

* 0500–0700 1st “morning” session [break after 2 hours]
* 0720–0920 2nd “morning” session
* [“overday” break – nearly 12 hours]
* 2100–2300 1st “night” session [break after 2 hours]
* 2320–0120+1 2nd “night” session

It was also pointed out that the session times had been changed from meeting to meeting, such that different time zones of the world might be treated approximately equally fairly either in one meeting or another. For the current meeting, the same session times were used as in the 23rd JVET meeting (which had been the sixth 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 1365 (as of 18 April 2022). All future discussions (including those on HEVC, VVC, CICP, etc.) shall be conducted on the JVET reflector rather than the JCT-VC reflector, while the old reflectors (including JVT, JCT-VC, and JCT-3V) 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
* **AML**: … merge list
* **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
* **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
* **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
* **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
* **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
* **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**: …
* **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 2020-10-13
    - 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 editing period, pending submission to ITTF
    - 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 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 (during the current meeting)
    - ISO/IEC 23008-2:2020 FDAM 1 ballot closed 2021-06-03 (shutter interval information SEI message) published 2021-07-12
    - ISO/IEC 23008-2:2020 CDAM 2 High-range levels output of 25th meeting of January 2022, CDAM ballot closed 2022-04-15, ballot results in [m59308](https://dms.mpeg.expert/doc_end_user/documents/138_OnLine/wg11/m59308-v1-m59308.zip)
    - 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, to close 2022-04-28
    - ISO/IEC 23090-3:202x (Ed. 2) with operation range extensions, FDIS approval at WG level 2022-01-21
    - Conformance testing
      * H.266.1 V1 Consented 2022-01-28, Last Call began 2022-04-01, to close 2022-04-28
      * ISO/IEC 23090-15 V1 FDIS approval at WG level 2022-10-15
      * ISO/IEC 23090-15 DAM 1 operation range extensions – DAM from previous meeting, DAM ballot pending, no action at his meeting
    - Reference software
      * H.266.2 V1 Consented 2022-01-28, Last Call began 2022-04-01, to close 2022-04-28
      * ISO/IEC 23090-16 V1 FDIS approval at WG level 2022-01-21, FDIS ballot pending
  + 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, to close 2022-04-28
    - ISO/IEC 23002-7:202x (Ed. 2) FDIS approval at WG level 2022-01-21, FDIS ballot pending
  + CICP V2 (includes errata items)
    - ISO/IEC 23091-2 V2 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
  + 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 WD issued 2022-01, pending submission of output document
  + 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 2022-04-17. (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 included in 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-Y1005) – ISO/IEC 23008-2 CDAM 2 ballot closed 2022-04-15. Issue DIS of new edition of HEVC, incorporating Amd.1 and corrigenda items (FDIS in January 2023, ITU-T consent in October); note that Amd.1 = shutter interval is already included in latest ITU-T edition of H.265
  + New level and systems-related supplemental enhancement information (from JVET-Y2019) – New VVC amendment was requested from 25th meeting, CDAM (JVET draft 2) to be issued
  + Film grain synthesis technology for video applications (from JVET-Y2020) – JVET draft 2 to be issued, plan for DTR from the 27th meeting
  + The 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 Recommendations. A request form also needs to be filled out (but 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 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-Y1000 were reviewed. The following small copy-paste issues in the meeting report were noted and were not considered sufficient to warrant issuing a revision:
  + In the list of output documents in section 1, JVET-X2017 should have been JVET-Y2017. In sections 2.3 and 2.4, some numbering errors occurred: the VTM output of the prior meeting should have been 16 instead 15, and the ECM output should have been 4 instead 3. The latter should also be numbered as JVET-X2026 rather than JVET-W2026 in section 2.4.3. In the output document list in section 10, the status of the additional SEI messages for VSEIv2 should have been shown as FDIS rather than DIS. On the document site, everything is correct.
* 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 quite a few documents registered where authors’ given names were not abbreviated, and company affiliation was missing in the authors’ list. Participants were reminded to stick to JVET’s conventions.
* The primary goals of the meeting were
  + Errata
  + Conformance for version 2 of VVC (under DAM ballot)
  + New levels for HEVC (DAM -> DIS)
  + New level and systems-related SEI for VVC (CDAM)
  + Preparation of TR for film grain (draft 2, which was later cancelled)
  + Exploration Experiments
    - Neural network-based video coding
    - Enhanced compression beyond VVC
* Liaison communication: JPEG has sent a liaison [m59796](https://dms.mpeg.expert/doc_end_user/current_document.php?id=82797&id_meeting=190) to various MPEG WGs (including WG 5) informing them about the scope and CfP on JPEG-AI. No specific request is made – not necessary to send a response.
* Number of documents slightly lower than last meeting (140->135)
* 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, in particular regarding the plan of a hybrid/physical meeting in October 2022

## Scheduling of discussions

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

* 0500–0700 1st “morning” session [break after 2 hours]
* 0720–0920 2nd “afternoon” session
* [“overday” break – nearly 12 hours]
* 2100–2300 1st “night” session [break after 2 hours]
* 2320–0120+1 2nd “night” session

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

* Wed. 20 Apr., 1st day
  + Session 1:
    - 0500–0550 Opening remarks, review of practices, agenda, IPR reminder
    - 0550–0700 Reports of AHGs 1–9
  + Session 2:
    - 0720–0750 Reports of AHGs 10–13
    - 0750–00935 Review of EE1 summary
  + Session 3:
    - 2100–2300 Review of EE2 summary
  + BoG (Y. Ye):
    - 2320–0120+1 Review of EE2 related (5.3.3) & “non-EE2” (5.3.4)
* Thu. 21 Apr., 2nd day
  + Session 4:
    - 0500–0700 Review of EE1 and related
  + Session 5:
    - 0720–0920 Review of EE1 and related
  + Session 6:
    - 2100–2300 BoG report EE2 related, further review of EE2 summary report
  + BoG (Y. Ye):
    - 2320–0120+1 Review of EE2 related (5.3.3) & “non-EE2” (5.3.4)
* Fri. 22 Apr., 3rd day
  + BoG (E. Alshina):
    - 0500–0700 EE1 results analysis and further planning of EE1
  + Session 7:
    - 0500–0700 Review of 4.2, 4.12, 6.1, 6.2
  + Session 8:
    - 0720–0920 Review of 4.2, 4.12, 6.1, 6.2
  + Session 9:
    - 2100–2300 Review of EE2 summary and remaining EE1 related
  + BoG (Y. Ye):
    - 2320–0120+1 Review of EE2 related (5.3.3) & “non-EE2” (5.3.4)
* Mon. 25 Apr., 4th day
  + 0500–0700 MPEG information sharing session
  + Session 10:
    - 0750–0920 BoG report NN, Review of remaining docs 5.2.4
  + Session 11:
    - 2100–2300 BoG report EE2, Review of docs 5.3.4
  + BoG (Y. Ye):
    - 2320–0120+1 Review of docs 5.3.4
* Tue. 26 Apr., 5th day
  + 0500-0700 BoG on neural networks (E. Alshina)
  + Session 12:
    - 0500–0700 Review of 4.3, 4.9, 4.10
  + Session 13:
    - 0720–0920 Review of 4.10, 4.11, BoG reports, further planning
  + Session 14:
    - 2100–2300 NN BoG report JVET-Z0234, review of remaining docs 5.3.4
  + 2320–0120+1 BoG on EE2 related (Y. Ye)
* Wed. 27 Apr., 6th day
  + 0500–0600 MPEG information sharing session
  + Session 15:
    - 0620–0920 BoG report JVET-Z0210, Review of HLS documents (6.1, 6.2, 6.3)
  + 2100-2300 Joint meeting with AG5: Documents in 4.5 and 4.6
  + 2320-0120+1 BoG on FGS (A. Tourapis)
* Thu. 28 Apr., 7th day
  + Session 16:
    - 0500–0700 BoG report FGS, Review of 6.3, 4.1, revisits, output planning
  + Session 17:
    - 0730-0800 Review remaining 5.3.4 documents
    - 0800-0920 Joint with AG5: Remaining 4.6, 4.5
  + Session 18:
    - 2100–2300 Revisits
* Fri. 29 Apr., 8th day
  + Plenary:
    - 0500–0710 AHG, output document timelines, review/approvals of DoCR&requests
    - 0730–0935 Output document reviews and approvals, draft recommendations, meeting planning, AoB
  + 2100–2300 MPEG information sharing session
  + 00XX+1–00XX+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 (1)
  + Verification testing (0)
  + Test Material (3)
  + Quality assessment (5)
  + Conformance test development (0)
  + Software development (1)
  + Implementation studies and complexity analysis (1)
  + AHG7: Low latency and constrained complexity (5)
  + Encoding algorithm optimization (5)
  + 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 (0) (section 5.1)
  + AHG11 and EE1: Neural network-based video coding (25) (section 5.2)
  + AHG12 and EE2: Enhanced compression beyond VVC capability (56) (section 5.3)
* High-level syntax (HLS) proposals (section 6) with subtopics
  + AHG9: SEI message studies and proposals (7) (section 6.1)
  + Film grain synthesis (3) (section 6.2)
  + Non-SEI HLS aspects (1) (section 6.3)
* Joint meetings, plenary discussions, BoG and viewing reports (0), summary of actions (section 7)
* Project planning (section 8)
* Establishment of AHGs (section 9)
* Output documents (section 10)
* Future meeting plans and concluding remarks (section 11)

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

# AHG reports (13)

These reports were discussed Wednesday 20 Apr. 2022 in session 1 during 0550–0700 and in session 2 0720–0750 UTC (chaired by JRO).

[JVET-Z0001](https://jvet-experts.org/doc_end_user/current_document.php?id=11531) JVET AHG report: Project management (AHG1) [J.-R. Ohm, G. J. Sullivan]

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_01_Y_Virtual/>). It is noted that the previous document site <http://phenix.int-evry.fr/jvet/> is still accessible, but was converted to read-only.

The list of documents produced included the following, particularly:

* The meeting report (JVET-Y1000) [Posted 2022-02-19, also submitted as WG 5 N 96]
* High Efficiency Video Coding (HEVC) Test Model 16 (HM 16) Encoder Description Update 16 (JVET-Y1002) [Posted 2022-04-12, also submitted as WG 5 N 103]
* Errata report items for VVC, HEVC, AVC, Video CICP, and CP usage TR (JVET-Y1004) [Posted 2022-03-22]
* New levels for HEVC (Draft 2) (JVET-Y1005) [Posted 2022-02-18, also submitted as WG 5 CDAM2 N 102]
* Common Test Conditions for HM Video Coding Experiments (JVET-Y1100) [Posted 2022-02-21]
* Algorithm description for Versatile Video Coding and Test Model 16 (VTM 16) (JVET-y2002) [Posted 2022-03-30, also submitted as WG 5 N 106]
* VVC operation range extensions (Draft 6) (JVET-Y2005) [Posted 2022-01-27, last update 2022-03-22, also integrated into VVC version 2 which was submitted as WG 5 N 105 and for ITU consent]
* Additional SEI messages for VSEI (Draft 6) (JVET-Y2006) Posted 2022-01-27, last update 2022-03-21, also integrated into VVC version 2 which was submitted as WG 5 N 100 and for ITU consent]
* Reference software for versatile video coding (Draft 3) (JVET-Y2009) [Posted 2022-04-16, also submitted as WG 5 FDIS N 112 and for ITU consent]
* VTM common test conditions and software reference configurations for SDR video (JVET-Y2010) [Posted 2022-02-21]
* VTM common test conditions and evaluation procedures for HDR/WCG video (JVET-Y2011) [Posted 2022-02-15]
* Common Test Conditions and evaluation procedures for enhanced compression tool testing (JVET-Y2017) [Posted 2022-02-25]
* New level and systems-related supplemental enhancement information for VVC (Draft 1) (JVET-Y2019) [Posted 2022-03-19, also submitted as WG 5 WD N 108]
* Film grain synthesis technology for video applications (Draft 1) (JVET-Y2020) [Posted 2022-04-20, also submitted as WG 5 WD N 120]
* Exploration experiment on Neural Network-based Video Coding (EE1) (JVET-Y2023) [Posted 2022-01-21, last update 2022-02-04, also submitted as WG 5 N 113]
* Exploration experiment on Enhanced Compression beyond VVC capability (EE2) (JVET-Y2024) [Posted 2022-01-21, last update 2022-02-19, also submitted as WG 5 N 114]
* Algorithm description of Enhanced Compression Model 4 (ECM 4) (JVET-Y2025) [Posted 2022-04-12, last update 2022-04-13, also submitted as WG 5 N 115]
* Conformance testing for VVC operation range extensions (Draft 3) (JVET-Y2026) [Posted 2022-03-18, also submitted as WG 5 DAM1 N 110]

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

* Recommendations of the 6th WG 5 meeting (WG 5 N 95)
* Disposition of comments received on ISO/IEC DIS 14496-10:202X Advanced Video Coding (10th edition) (WG 5 N 97)
* Text of ISO/IEC 14496-10:202X Advanced Video Coding (10th edition) (WG 5 N 98)
* Disposition of comments received on ISO/IEC DIS 23002-7:202X Versatile supplemental enhancement information messages for coded video bitstreams (2nd edition) (WG 5 N 99)
* Request for ISO/IEC 23008-2:2000/Amd.2 High-range levels (WG 5 N 101)
* Disposition of comments received on ISO/IEC DIS 23090-3:202X Versatile video coding (2nd edition) (WG 5 N 104)
* Request for ISO/IEC 23090-3:200x Amd.1 New level and systems-related supplemental enhancement information (WG 5 N 107)
* Disposition of comments received on ISO/IEC 23090-15 CDAM1 (WG 5 N 109)
* Disposition of comments received on ISO/IEC DIS 23090-16 (WG 5 N 111)
* Request for ISO/IEC TR 23002-9 Film grain synthesis technology for video applications (WG 5 N 119)
* Liaison statement to ISO/IEC JTC1/SC 29/WG 1 (JPEG) on machine learning-based image and video compression (WG 5 N 116)
* List of AHGs established at the 6th WG 5 meeting (WG 5 N 117)

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 issues 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. Significant activities were also conducted on preparation of subjective tests, and on development of VVC conformance testing.

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 135 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 26th meeting had been made publicly available on the ITU-hosted ftp site <http://wftp3.itu.int/av-arch/jvet-site/2022_04_Z_Virtual/>.

The AHG recommended its continuation.

[JVET-Z0002](https://jvet-experts.org/doc_end_user/current_document.php?id=11535) 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]

**Output documents produced:**

**JVET-Y1002 High Efficiency Video Coding (HEVC) Test Model 16 (HM 16) Encoder Description Update 16**

JVET released the HEVC test model (HM) 16.26 software following its 25th meeting, held online. This document serves as an overview of HEVC Version 1 and the Range Extensions of HEVC Version 2, and also provides an encoder-side description of the HM-16.26 software.

Update 16 incorporated items:

* JVET-Y0077 AHG10: Block importance mapping
* JVET-Y0155 AHG10: Fixes and clean up for temporal prefilter

**JVET-Y1005 New levels for HEVC (Draft 2)**

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 2 incorporated items:

* Addition of Levels 7, 7.1, and 7.2 (JVET-Y0072)
* Some technical corrections from JVET-X1004
* JVET-Y0050 AHG2/AHG9: On the alpha channel information SEI message
* JVET-Y0107 AHG9: Text improvement for the film grain SEI

**JVET-Y2002 Algorithm description for Versatile Video Coding and Test Model 16 (VTM 16)**

The JVET established the VVC Test Model 16 (VTM16) algorithm description and encoding method at its 25th meeting (12 – 21 January 2022, teleconference). This document serves as a source of general tutorial information on the VVC design and also provides an encoder-side description of VTM16. 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 16 (VTM16) algorithm description and encoding method

* Added description of rate control in VTM, incorporated JVET-Y0105: AHG10: An improved VVC rate control scheme
* Incorporated JVET-Y0155: AHG10: Fixes and clean up for temporal prefilter
* Incorporated JVET-Y0077: AHG10: Block importance mapping

**JVET-Y2005 VVC operation range extensions (Draft 6)**

This document contains the draft text for changes to the Versatile Video Coding (VVC) standard (ITU‑T H.266 | ISO/IEC 23090-3), for the support of the operation range extensions, the addition of level 6.3 and some SEI messages, including the SEI manifest SEI message, and the SEI prefix indication SEI message, and the constrained RASL encoding indication SEI message. In addition, the document also contains SEI payload type values and other interfaces for SEI messages added to the VSEI specification, as well as some technical corrections to the content in first edition of VVC.

Draft 6 incorporated items:

* Bug fixes on new GCI flags (JVET-Y0237) and some related editorial changes
* Removal of the defined MinCrBase values for high tier at levels below 4.0 (JVET-Y0056)
* Addition of a NOTE clarifying which constraints do or do not apply for still picture profiles (JVET-Y0057)
* Addition of a requirement that a Main 10 4:4:4 Still Picture profile decoder shall also be capable of decoding the first picture of a Main 10 bitstream (JVET-Y0063)

**JVET-Y2006 Additional SEI messages for VSEI (Draft 6)**

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 annotated regions SEI message, the alpha channel information SEI message, the depth representation information SEI message, the multiview acquisition information SEI message, the scalability dimension information SEI message, the extended dependent random access point (DRAP) indication SEI message, the display orientation SEI message, the colour transform information SEI message, and the multiview view position SEI message. The draft text also includes text changes for some technical corrections and editorial improvements.

Draft 6 incorporated items:

* JVET-Y0050 AHG2/AHG9: On the alpha channel information SEI message
* Editorial changes of the phrases of "shall allow" and "shall ignore"
* JVET-Y0107 AHG9: Text improvement for the film grain SEI

**JVET-Y2019 New level and systems-related supplemental enhancement information for VVC (Draft 1)**

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 1 incorporated items:

* Addition of "hooks" for the green metadata and video decoding interface (VDI envelope) SEI messages (JVET-Y0041)
* Addition of new levels that enable higher resolution applications up to 16384x8640 at 120fps (JVET-Y0072)

**Related input contributions**

The following input contributions were noted as relevant to the work of this ad hoc group:

* JVET-Z0060 On new levels for HEVC
* JVET-Z0119 Editorial status of alpha blending text and ITU-T Last Call for VSEI v2
* JVET-Z0122 Some HEVC text changes

**Remaining bug tickets**

* [#1547](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1547) sb\_coded\_flag non-present inference logic error
* [#1544](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1544) sb\_coded\_flag definition lacks size of a sub-block

**Recommendations**

The AHG recommends to:

* Approve JVET-Y1002, JVET-Y1005, JVET-Y2002, JVET-Y2005, JVET-Y2006, and JVET-Y2019 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.

[JVET-Z0003](https://jvet-experts.org/doc_end_user/current_document.php?id=11536) JVET AHG report: Test model software development (AHG3) [F. Bossen, X. Li, K. Sühring, Y. He, K. Sharman, Y. Seregin, A. Tourapis]

The software model versions prior to the start of the meeting were:

* [VTM 16.0](https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/-/releases/VTM-15.0) (Mar. 2022)
* [HM-16.25](https://vcgit.hhi.fraunhofer.de/jvet/HM/-/releases/HM-16.25) (Apr. 2022)
* [HM-16.21+SCM-8.8](https://vcgit.hhi.fraunhofer.de/jvet/HM/-/tags/HM-16.21+SCM-8.8) (Mar. 2020)
* [SHM 12.4](https://vcgit.hhi.fraunhofer.de/jvet/SHM/-/tags/SHM-12.4) (Jan. 2018)
* [HTM 16.3](https://vcgit.hhi.fraunhofer.de/jvet/HTM/-/tags/HTM-16.3) (Jul. 2018)
* [JM 19.0](https://vcgit.hhi.fraunhofer.de/jvet/JM/-/tags/JM-19.0)
* [JSVM 9.19.15](https://vcgit.hhi.fraunhofer.de/jvet/jsvm/-/tags/JSVM_9_19_15)
* [JMVC 8.5](https://vcgit.hhi.fraunhofer.de/jvet/jmvc/-/tags/JMVC_8_5)
* [3DV ATM 15.0](https://vcgit.hhi.fraunhofer.de/jvet/3dv-atm/-/tags/3DV-ATM_v15.0) (no version history)
* [HDRTools 0.23](https://gitlab.com/standards/HDRTools/-/tags/v0.23) (October 2021)

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

1. **Software development**

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

The server is located at:

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

The registration and development workflow are documented at:

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

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

1. **VTM related activities**

The VTM software can be found at

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

The software development continued on the GitLab server. VTM versions 15.1 and 15.2 were tagged on Jan. 24 and Jan. 25, and VTM version 16.0 was tagged on Mar. 4. VTM 16.1 is expected during the 26th JVET meeting.

VTM 15.1 was tagged on Jan. 24, 2022. Changes include:

* Fix #1525: Don't access sps pointer after potential deletion
* Rename MTS related variables, fix parameters checking and documentation
* Fix compile issue with clang 11 compiler
* JVET-X0073: Decrease profile\_idc for 16 bit profiles by 1
* JVET-X0048-X0103: Implementation of film grain analysis and synthesis
* Add check of reference frame structure to detect low delay configuration

VTM 15.2 was tagged on Jan. 25, 2022. Changes include:

* Remove macros from previous cycle

VTM 16.0 was tagged Mar. 4, 2022. Changes include:

* Fix build of SW manual & update version numbers
* Fix typo in error message
* Fix check for temporal ID in BitstreamExtractor
* Remove macro JVET\_R0107\_BITSTREAM\_EXTACTION
* Fix #355: Avoid undefined behaviour
* Fix multilayer decoding and indentation, and some cleanup
* Fix braces and indentation
* Improve warning message and include end-of-line character
* Fix #1138: Fix range check of sps\_num\_subpics\_minus1
* Clean up SPS extension code
* Fix #1487: Look for slice boundaries, not tile boundaries
* Fix #1536: Fix computation of number of explicit slice heights
* Fix #1532: Use correct parameter name
* Fix #1519: Preserve pel fraction when clipping MV
* Fix #1535: Don't use deprecated std::unary\_function
* Fix m\_videoIOYuvSEIFGSFile may be used without being initialized
* Fix compilation with gcc 11: 'misleading-indentation' error
* Improve code alignment
* Remove T0196\_SELECTIVE\_RDOQ macro
* Remove unused macro
* Fix copyright year
* Fix #1523: Fixes compilation issues in ARM/non-x86 environments
* Revert "Fix #1523: Compilation issue in ARM environment"
* Revert "Add safeguards to the ENABLE\_SIMD\_OPT\* definitions missing those"
* Add safeguards to the ENABLE\_SIMD\_OPT\* definitions missing those
* Fix #1523: Compilation issue in ARM environment
* Fix variable names and indentation
* JVET-Y0060: Add option to configure AffineAmvp
* Use SSE function when constraints for AVX2 not satisfied
* Restore accidentally overwritten code
* JVET-Y0077: Block Importance Mapping
* JVET-Y0105: An improved VVC rate control scheme
* Added check in decoder
* Corrected level naming
* Clear MinCR definition for high tier for levels below 4.0 and add check
* Change copyright year to 2022
* Improve fix: now works again with temporal sublayer extraction (-t n)
* Improve fix: now merge splitting works correctly for independent layers, even those not including layer 0.
* Fix a decoder crash in multi-layer context when using e.g. ALF and -p 0: reset APS NALs for a skipped layer.
* JVET-Y0152: Fast skip method for TT partition. default threshold value is 1.075.
* Update the cfg for RA GOP16
* Fix RA DeblockingFilterOffsetInPPS
* JVET-Y0085: Cfg changes to align deblocking settings to ECM.
* Fix bug in reference picture lists construction for dependent layers, when base layer picture is Intra: was discarding all inactive pictures, causing strong quality degradation.
* Changes to address SW coordinators comments.
* Adding new params in configuration files
* JVET-Y0126: VTM encoder configurations for tests targeting improved coding performance
* Avoid fast skip in case of inter-layer prediction (same POC): return MAX instead of 0
* Print error message when gci\_num\_additional\_bits is in the range [1..5]
* JVET-Y0237:
* Fix #1524: CRA output when NoOutputBeforeRecoveryFlag = 1
* Add missing GDR\_ENABLED macros
* Do a clean reset of gopList POC in case of GDR configuration
* Remove unnecessary interaction between GDR and hash perfect match
* Fix #1529: Fix gdr restriction for affine merge mode
* Don't apply restriction for Intra coded CU in non refreshed frames
* Fix #1528: missing gdrPeriod for computing refresh area
* Fix FG related code
* Implement modifications discussed in meeting
* JVET-Y0155: Fixes and clean up for temporal prefilter
* Fix #1512: Bug-fix to the ticket 1512

VTM 16.1 is expected to be tagged during the 26th JVET meeting. Changes include so far:

* 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

***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 16.0:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **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 16.0** performance compared to **VTM 15.0**:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **All Intra Main10** |  |  |
|  |  |  | **Over VTM-15.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | -0.84% | -0.79% | 0.43% | 90% | 100% |
| Class A2 | -0.72% | -0.62% | -0.61% | 92% | 101% |
| Class B | -0.83% | -0.50% | -0.47% | 93% | 98% |
| Class C | -0.50% | -0.81% | -0.76% | 92% | 94% |
| Class E | -0.44% | -0.01% | 0.16% | 87% | 93% |
| **Overall** | -0.67% | -0.55% | -0.30% | 91% | 97% |
| Class D | -0.42% | -0.58% | -0.56% | 95% | 102% |
| Class F | -0.21% | -0.75% | -1.26% | 90% | 97% |
|  |  |  |  |  |  |
|  |  |  | **Random access Main10** |  |  |
|  |  |  | **Over VTM-15.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | -0.21% | -0.39% | -0.42% | 84% | 97% |
| Class A2 | -0.42% | -0.60% | -0.51% | 90% | 98% |
| Class B | -0.67% | -1.26% | -1.36% | 88% | 94% |
| Class C | -0.10% | -1.27% | -1.06% | 90% | 97% |
| Class E |  |  |  |  |  |
| **Overall** | -0.38% | -0.96% | -0.92% | 88% | 96% |
| Class D | 0.18% | -0.93% | -0.33% | 90% | 99% |
| Class F | 0.01% | -1.04% | -1.20% | 90% | 94% |
|  |  |  |  |  |  |
|  |  |  | **Low delay B Main10** |  |  |
|  |  |  | **Over VTM-15.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -0.58% | -0.45% | -0.70% | 88% | 95% |
| Class C | -0.33% | -0.63% | -0.26% | 91% | 94% |
| Class E | -0.92% | 0.61% | 1.08% | 91% | 98% |
| **Overall** | -0.58% | -0.25% | -0.11% | 90% | 95% |
| Class D | -0.10% | 0.11% | 0.83% | 93% | 97% |
| Class F | -0.40% | -0.87% | -0.71% | 89% | 97% |
|  |  |  |  |  |  |
|  |  |  | **Low delay P Main10** |  |  |
|  |  |  | **Over VTM-15.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -0.54% | -0.52% | -1.18% | 89% | 97% |
| Class C | -0.13% | -0.03% | -0.28% | 92% | 98% |
| Class E | -0.52% | 1.21% | 2.05% | 94% | 95% |
| **Overall** | -0.40% | 0.08% | -0.07% | 91% | 97% |
| Class D | -0.06% | -0.32% | 0.26% | 93% | 97% |
| Class F | -0.11% | -0.58% | -0.58% | 91% | 97% |

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

***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.

1. **HM related activities**

HM 16.25 was tagged on April 13, 2022. Changes include:

* JVET-X0079: Add new level 6.3
* JVET-X0116: Enabled temporal filter for low-delay configurations and also fixed a bug for non 4:2:0
* JVET-T0050: Add ability to detect static objects to encoder
* ARSEI Fix functionality
* Added the ability to detect static objects from the supplied annotated regions SEI config files. If the detected object is static w.r.t previous frame, then no additional information will be sent in the bitstream.
* Update software manual for HM 16.25

HM 16.26 is expected to be tagged during or shortly after the 26th JVET meeting. Changes include so far:

* JVET-Y0155: Fixes for motion-compensated temporal prefilter
* JVET-Y0105: An improved VVC rate control scheme

The following MRs are pending:

* JVET-Y0077: Block Importance Mapping
* [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

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.

The following tables show **HM 16.25** performance compared to **HM 16.24**:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **All Intra Main10** |  |  |
|  |  |  | **Over HM-16.24** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | 0.00% | 0.00% | 0.00% | 101% | 100% |
| Class A2 | 0.00% | 0.00% | 0.00% | 100% | 100% |
| Class B | 0.00% | 0.00% | 0.00% | 101% | 99% |
| Class C | 0.00% | 0.00% | 0.00% | 99% | 100% |
| Class E | 0.00% | 0.00% | 0.00% | 100% | 100% |
| **Overall** | 0.00% | 0.00% | 0.00% | 100% | 100% |
| Class D | 0.00% | 0.00% | 0.00% | 100% | 97% |
| Class F | 0.00% | 0.00% | 0.00% | 100% | 99% |
|  |  |  |  |  |  |
|  |  |  | **Random access Main10** |  |  |
|  |  |  | **Over HM-16.24** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | 0.00% | 0.00% | 0.00% | 98% | 102% |
| Class A2 | 0.00% | 0.00% | 0.00% | 98% | 103% |
| Class B | 0.00% | 0.00% | 0.00% | 103% | 105% |
| Class C | 0.00% | 0.00% | 0.00% | 99% | 104% |
| Class E |  |  |  |  |  |
| **Overall** | 0.00% | 0.00% | 0.00% | 100% | 104% |
| Class D | 0.00% | 0.00% | 0.00% | 101% | 97% |
| Class F | 0.00% | 0.00% | 0.00% | 99% | 103% |
|  |  |  |  |  |  |
|  |  |  | **Low delay B Main10** |  |  |
|  |  |  | **Over HM-16.24** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -1.18% | -0.73% | -0.94% | 100% | 99% |
| Class C | -0.20% | -0.15% | -0.03% | 101% | 112% |
| Class E | -0.68% | -0.82% | -0.79% | 100% | 105% |
| **Overall** | -0.73% | -0.56% | -0.60% | 100% | 105% |
| Class D | 0.12% | 0.01% | -0.07% | 99% | 106% |
| Class F | 0.00% | 0.00% | 0.00% | 99% | 108% |
|  |  |  |  |  |  |
|  |  |  | **Low delay P Main10** |  |  |
|  |  |  | **Over HM-16.24** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -1.87% | -1.50% | -1.48% | 101% | 100% |
| Class C | -0.20% | -0.05% | -0.02% | 101% | 110% |
| Class E | -1.20% | -1.24% | -1.33% | 103% | 108% |
| **Overall** | -1.15% | -0.95% | -0.95% | 101% | 105% |
| Class D | -0.03% | -0.39% | -0.15% | 101% | 109% |
| Class F | 0.00% | 0.00% | 0.00% | 99% | 106% |

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

Help to address these tickets would be appreciated.

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

1. **360Lib related activities**

Development of 360Lib was moved to the GitLab server. 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-16.0 according to 360-degree video CTC (JVET-U2012) compared to that using VTM-15.0 (VTM-15.0 as anchor).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **PERP: VTM-16.0 over VTM-15.0** | | | | | |
|  | **End-to-end**  **WS-PSNR** | | | **End-to-end**  **S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | -0.59% | 0.20% | -0.06% | -0.57% | 0.20% | -0.05% |
| Class S2 | -0.35% | 0.31% | 0.11% | -0.36% | 0.31% | 0.12% |
| **Overall** | -0.50% | 0.24% | 0.01% | -0.49% | 0.24% | 0.02% |

The following table compares generalized cubemap (GCMP) coding and padded equi-rectangular projection (PERP) coding using VTM-16.0 (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 | -11.44% | -5.53% | -6.00% | -11.45% | -5.47% | -5.95% |
| Class S2 | -3.62% | 0.46% | 0.89% | -3.61% | 0.57% | 0.96% |
| **Overall** | -8.31% | -3.14% | -3.24% | -8.31% | -3.05% | -3.18% |

The following tables are for PERP and GCMP coding comparison between VTM-16.0 and HM-16.22 (HM as anchor), respectively.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **VTM-16.0 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 | -30.57% | -37.07% | -39.72% | -30.55% | -37.12% | -39.72% |
| Class S2 | -36.43% | -35.74% | -38.21% | -36.42% | -35.77% | -38.26% |
| **Overall** | -32.92% | -36.53% | -39.12% | -32.90% | -36.58% | -39.14% |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **VTM-15.0 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 | -35.25% | -39.22% | -41.28% | -35.22% | -39.15% | -41.23% |
| Class S2 | -39.38% | -38.27% | -40.30% | -39.39% | -38.27% | -40.34% |
| **Overall** | -36.90% | -38.84% | -40.89% | -36.89% | -38.79% | -40.87% |

1. **SCM related activities**

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

1. **SHM related activities**

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

1. **HTM related activities**

There had not been any updates to the HTM of MV-HEVC and 3D-HEVC.

1. **HDRTools related activities**

There had not been any updates of HDRTools.

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

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

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

1. **Bug tracking**

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

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

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

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

Bug tracking for HDRTools is located at:

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

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

1. **Software repositories**

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

1. **CTC alignment and merging**

JVET-Y2010 was produced as JVET output document containing the merged VTM and HM CTC for SD 4:2:0 10-bit video.

JVET-Z0175 was registered on merging HDR CTC.

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 27th meeting.

1. **Recommendations**

The AHG recommends to:

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

It is noted that the gain of VTM16 in case of RA is mainly due to change in deblocking.

[JVET-Z0004](https://jvet-experts.org/doc_end_user/current_document.php?id=11537) 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]

**Activities**

***Remote experts viewing***

Remote experts viewing (REV) sessions have been prepared and conducted in the context of AhG4 considering new test material and AhG11/EE1 Neural network-based video. The results are reported in JVET-Z0053.

***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).

New test content is proposed in contributions JVET-Z0138 and JVET-Z0156.

**Related contributions**

|  |  |  |
| --- | --- | --- |
| [JVET-Z0045](https://jvet-experts.org/doc_end_user/current_document.php?id=11479) | AhG11/AhG4/EE1 viewing preparation report | [E. Alshina](mailto:elena.alshina@huawei.com), [M. Wien](mailto:mathias.wien@lfb.rwth-aachen.de), [A. Segall](mailto:asegall@amazon.com), [J. Sauer (coordinators)](mailto:johannes.sauer1@huawei.com) |
| [JVET-Z0053](https://jvet-experts.org/doc_end_user/current_document.php?id=11487) | [AHG4] REV Result for AHG11/EE1 and AHG4 new test sequences | [M. Wien (RWTH)](mailto:wien@lfb.rwth-aachen.de) |
| [JVET-Z0057](https://jvet-experts.org/doc_end_user/current_document.php?id=11492) | On subjective evaluation of video quality with the crowdsourcing approach | B. Naderi (TU Berlin), R. Cutler (Microsoft) |
| [JVET-Z0107](https://jvet-experts.org/doc_end_user/current_document.php?id=11554) | Quality evaluation of internal 10-bit versus 8-bit processing with 8-bit source content | [J. Jung](mailto:joeljung@tencent.com), X. Li, S. Liu (Tencent) |
| [JVET-Z0108](https://jvet-experts.org/doc_end_user/current_document.php?id=11555) | Evaluation of 17 objective quality metrics on JVET contents | [J. Jung](mailto:joeljung@tencent.com), J. G. Lopez, X. Li, S. Liu (Tencent) |
| [JVET-Z0109](https://jvet-experts.org/doc_end_user/current_document.php?id=11556) | Evaluation of objective quality metrics on HEVC, VVC and AV1 contents | [J. Jung](mailto:joeljung@tencent.com), J. G. Lopze, X. Li, S. Liu (Tencent) |
| [JVET-Z0138](https://jvet-experts.org/doc_end_user/current_document.php?id=11587) | AHG7: Update on gaming sequences | [T. Poirier](mailto:tangi.poirier@interdigital.com), G. Martin-Cocher, E. Faivre d'Arcier (Interdigital) |
| [JVET-Z0156](https://jvet-experts.org/doc_end_user/current_document.php?id=11605) | AHG4: New test sequences for JVET exploration | [J. Chen](mailto:jiechen.cj@alibaba-inc.com), [Y. Ye (Alibaba)](mailto:yan.ye@alibaba-inc.com), [R. Li](mailto:yichen.lr@alibaba-inc.com), [W. Jiang (Youku)](mailto:wenfei.jwf@taobao.com) |

**Recommendations**

The AHG recommends:

* To consider the results of the remote experts viewing sessions in the discussion of the NN-based video coding tools in AhG11.
* To discuss contributions JVET-Z0057, JVET-Z0108, and JVET-Z0109 in a joint session with SC29/AG5.
* To study the new test sequences proposed in JVET-Z0138 and JVET-Z156.
* 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 noted that the level of voluntarily participating in the expert viewing has decreased compared to previous meetings. It is suggested to identify participants (list of names) before preparing future expert viewing tests.

[JVET-Z0005](https://jvet-experts.org/doc_end_user/current_document.php?id=11538) JVET AHG report: Conformance testing (AHG5) [D. Rusanovskyy, I. Moccagatta, F. Bossen, K. Kawamura, T. Hashimoto, H.-J. Jhu, K. Sühring, Y. Yu]

The progress on the Conformance testing specification is consistent with the preliminary timeline agreed at the 25th JVET meeting, as follows:

* 24th meeting Oct 2021: CDAM v2 conformance
* 25th meeting Jan 2022:
  + FDIS v1 conformance (plus ITU consent)
  + no action pending CDAM v2 ballot in ISO/IEC
* 27th meeting July 2022: FDAM v2
* CDAM v2 ITU consent in October

(some update needed on timeline)

**Status on bitstream submission**

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

* conformance bitstreams for VVC:
  + 104 bitstream categories have been identified
  + At least one bitstream has been submitted in each identified category
  + 283 total bitstreams have been provided, checked, and made available
  + No changes between 25th and 26th meeting.
* conformance bitstreams for VVC operation range extensions:
  + 57 (+6) 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 (was 47) categories
  + All (was 35) bitstream descriptions have been provided
  + 127 bitstreams of 56 identified categories have been cross-checked and uploaded
  + 1 bitstream of 1 identified category is in the process of being re-generated.

**Activities and Discussion**

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

VVC activities:

There are not currently any known issues with the other provided VVC conformance bitstream packages. All provided bitstreams can be decoded using VTM16 w/o and with range extension support.

VVC operation range extensions activities:

* Volunteers to generate and cross-check the bitstreams have been solicited and identified.
* Volunteers and cross-checkers have exchanged bitstreams using their own means, and bitstreams have been uploaded when cross-check has been confirmed.
* Bitstreams have been regenerated and cross-checked following the adoption of JVET-Y0237 option C and filing of VVC conformance tickets #[1538](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1538) (bad value for gci\_num\_additional\_bits) and #[1539](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1539) (bad value for sps\_bitdepth\_minus8). Both tickets have been closed.
* All cross-checked bitstreams can be decoded using VTM16.
* It has been pointed out that some conformance packages generated using RGB data have decoded video checksum files with names ending in .rgb.md5 while Section A.2.2 of JVET-Y2026 stipulates that such filenames should end with .yuv.md5. One suggestion to address this mismatch is to keep yuv.md5 file naming convention for all tests to avoid introducing a new file naming convention and confusion. This may require re-generating conformance bitstreams in one category.
* It has been pointed out that for several bitstreams the VTM decoder command line option “--OutputColourSpaceConvert=GBRtoRGB” has been used to produce yuv.md5 files. The decoding directions are provided in the bitstream description files. It was discussed if this information should be also provided in the conformance specification text.

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

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

Contributions

There are no input contributions.

**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:



**Recommendations**

The AHG recommends the following:

* Close on VVC operation range extensions checksum files extension for RGB packages.
* Discuss if additional clarification on using “--OutputColourSpaceConvert=GBRtoRGB” should be added to the VVC operation range extensions conformance document.
* Port VVC operation range extensions editorial changes made to ISO/IEC version of the document to JVET document.

Issues were asked to be clarified offline.

It was later reported that the issues were resolved as follows from offline consultation among editors:

• keep yuv.md5 file naming convention for all tests to avoid introducing a new file naming convention and confusion. This requires re-generating 3 v2 packages to rename .rgb.md5 as .yuv.md5 and a change to Sec. 6.5.2 of conformance to remove reference to decoded yuv file.

• Add “--OutputColourSpaceConvert=GBRtoRGB” in the conformance text.

It was further reported later that ISO publication staff requested two changes to the v1 FDIS text:

* Remove company names from bit streams
* Remove email addresses of contact persons

It was suggested to inform the contributors of bitstreams about this. The editors are asked to work out a schema for consistent renaming, such that the deviation between ITU and ISO versions is minimized.

[JVET-Z0006](https://jvet-experts.org/doc_end_user/current_document.php?id=11478) JVET AHG report: ECM software development (AHG6) [V. Seregin, J. Chen, F. Le Léannec, K. Zhang]

**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-Y0116 (version 2.1a): Extension of MRL list to 5 lines 1,3,5,7,12

JVET-Y0065 (test 3.1c, not in LP CTC): GPM intra-inter mode

JVET-Y0134 (test 3.6a): MV candidate reordering for TMVP and NAMVP types, and reference picture selection for TMVP

JVET-Y0067 (test 3.9d): TM based reordering for MMVD and affine MMVD and MVD sign prediction

JVET-Y0058 (test 3.13): Modifications of IBC merge/AMVP list construction

JVET-Y0142 (test 4.4a): Adaptive intra MTS with fixed threshold

JVET-Y0106: CCSAO edge offset

JVET-Y0141 (test 3): Sign prediction improvement

JVET-Y0159 (method 1): Inter MTS uses fixed 4 candidates

JVET-Y0089: DMVR with BCW

JVET-Y0128: Fixing issues for RPR enabling and non-CTC configuration in ECM

JVET-Y0129: MVD signalling

JVET-Y0156: Fix for histogram of gradients derivation in DIMD mode

JVET-Y0060: Add AffineAMVP cfg option

JVET-Y0152 (SW/CTC): Fast skip method of TT partition search

JVET-Y0155 (SW/CTC): Fixes and clean up for temporal prefilter

JVET-Y0240 (SW): Block importance mapping

ECM software improvements:

MS-SSIM

log output alignment with VTM

ECM-4.0 was tagged on February 15, 2022.

***CTC Performance***

In this section, ECM-4.0 test results following ECM CTC configuration descried in JVET-X2017 are summarized.

The below tables show ECM-4.0 performance over ECM-3.1 anchor.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | |
|  | **Over ECM-3.1** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -0.45% | -0.94% | -0.86% | 104% | 110% |
| Class A2 | -0.51% | -0.98% | -3.54% | 103% | 110% |
| Class B | -0.49% | -1.35% | 0.08% | 107% | 109% |
| Class C | -0.32% | 0.08% | 0.12% | 111% | 111% |
| Class E | -0.55% | -0.85% | -0.69% | 106% | 110% |
| **Overall** | -0.46% | -0.82% | -0.80% | 107% | 110% |
| Class D | -0.29% | -0.18% | -0.01% | 114% | 115% |
| Class F | -0.40% | -0.70% | -0.35% | 99% | 100% |
| Class TGM | -1.01% | -1.66% | -1.57% | 96% | 95% |
|  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | |
|  | **Over ECM-3.1** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -1.39% | -2.32% | -2.31% | 94% | 104% |
| Class A2 | -1.53% | -2.88% | -4.81% | 96% | 108% |
| Class B | -1.51% | -2.91% | -1.73% | 100% | 113% |
| Class C | -1.12% | -1.25% | -1.47% | 102% | 124% |
| Class E |  |  |  |  |  |
| **Overall** | -1.39% | -2.34% | -2.39% | 99% | 113% |
| Class D | -0.92% | -0.86% | -1.23% | 106% | 127% |
| Class F | -0.86% | -0.61% | -0.69% | 97% | 102% |
| Class TGM | -1.26% | -1.33% | -1.29% | 98% | 98% |
|  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | |
|  | **Over ECM-3.1** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -1.66% | -3.57% | -2.32% | 107% | 122% |
| Class C | -1.80% | -2.60% | -2.03% | 109% | 128% |
| Class E | -0.94% | -2.03% | -2.20% | 110% | 100% |
| **Overall** | -1.53% | -2.86% | -2.19% | 108% | 118% |
| Class D | -1.59% | -3.25% | -2.91% | 111% | 135% |
| Class F | -1.30% | -2.99% | -1.30% | 105% | 99% |
| Class TGM | -1.47% | -1.59% | -1.44% | 102% | 99% |
|  |  |  |  |  |  |
|  | **Low delay P Main 10** | | | | |
|  | **Over ECM-3.1** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -1.24% | -3.20% | -1.95% | 111% | 127% |
| Class C | -1.20% | -1.27% | -1.02% | 115% | 141% |
| Class E | -0.65% | -2.35% | -0.63% | 116% | 112% |
| **Overall** | -1.08% | -2.35% | -1.31% | 114% | 127% |
| Class D | -1.10% | -2.23% | -0.98% | 116% | 146% |
| Class F | -0.94% | -0.81% | -0.89% | 111% | 112% |
| Class TGM | -1.39% | -1.47% | -1.30% | 105% | 105% |

Next tables show ECM-4.0 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-11.0ecm** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -6.60% | -12.43% | -17.68% | 357% | 262% |
| Class A2 | -6.24% | -13.37% | -14.77% | 342% | 254% |
| Class B | -5.70% | -14.30% | -13.65% | 382% | 248% |
| Class C | -6.89% | -9.30% | -9.80% | 373% | 242% |
| Class E | -7.67% | -12.17% | -13.14% | 363% | 264% |
| **Overall** | -6.53% | -12.37% | -13.57% | 366% | 253% |
| Class D | -5.73% | -8.03% | -7.65% | 371% | 241% |
| Class F | -10.93% | -15.45% | -14.70% | 253% | 255% |
| Class TGM | -16.70% | -18.91% | -18.31% | 245% | 276% |
|  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | |
|  | **Over VTM-11.0ecm** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -16.20% | -17.76% | -23.60% | 389% | 471% |
| Class A2 | -16.97% | -22.37% | -24.24% | 375% | 546% |
| Class B | -14.46% | -21.68% | -19.89% | 403% | 505% |
| Class C | -16.40% | -17.23% | -17.19% | 426% | 504% |
| Class E |  |  |  |  |  |
| **Overall** | -15.83% | -19.85% | -20.78% | 400% | 506% |
| Class D | -16.99% | -17.54% | -17.59% | 411% | 523% |
| Class F | -14.23% | -17.29% | -17.00% | 349% | 363% |
| Class TGM | -16.11% | -19.04% | -19.39% | 387% | 299% |
|  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | |
|  | **Over VTM-11.0ecm** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -12.29% | -26.50% | -24.33% | 351% | 397% |
| Class C | -13.71% | -20.74% | -19.80% | 365% | 378% |
| Class E | -11.18% | -18.16% | -19.41% | 344% | 299% |
| **Overall** | -12.48% | -22.49% | -21.59% | 353% | 364% |
| Class D | -15.49% | -22.79% | -21.74% | 348% | 387% |
| Class F | -12.98% | -20.76% | -18.58% | 325% | 296% |
| Class TGM | -15.26% | -22.73% | -22.40% | 370% | 292% |
|  |  |  |  |  |  |
|  | **Low delay P Main 10** | | | | |
|  | **Over VTM-11.0ecm** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -13.10% | -27.56% | -25.62% | 324% | 377% |
| Class C | -14.70% | -21.22% | -20.76% | 308% | 399% |
| Class E | -12.17% | -18.72% | -20.33% | 321% | 298% |
| **Overall** | -13.40% | -23.24% | -22.68% | 318% | 362% |
| Class D | -16.61% | -23.06% | -22.80% | 290% | 398% |
| Class F | -13.44% | -20.52% | -19.30% | 311% | 296% |
| Class TGM | -14.92% | -22.44% | -22.03% | 373% | 291% |

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

**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 the VTM anchor includes the improvements that were adopted in the last meeting to VTM16.

[JVET-Z0007](https://jvet-experts.org/doc_end_user/current_document.php?id=11539) JVET AHG report: Low latency and constrained complexity (AHG7) [A. Duenas, T. Poirier, S. Liu, L. Wang, J. Xu]

**Related contributions**

***Low delay configuration***

**JVET-Z0110: On low delay configuration**

In this contribution, the coding efficiency and coding speed by low delay B (LDB) and low delay P configurations are compared under different conditions.

**JVET-Z0114: AHG7: Low delay configuration for cloud gamin**

This contribution proposes a new low delay configuration (named LDB-1Ref-asymmetric) where the number of reference frames is restricted to 1 frame per list, and reference frame list is modified such that reference list 0 refers to the previous picture in the GOP and reference list 1 refers to the picture at the start of the GOP.

**JVET-Z0116: AHG7: Refined low latency and controlled complexity configuration for cloud gaming**

This contribution proposes a refined low latency and controlled complexity (LLCC) configuration for cloud gaming (JVET-Y0043) by disabling MTS for low latency and controlled complexity video applications such as cloud gaming.

***GDR implementation***

**JVET-Z0118: AHG7: GDR Implementation for ECM 4.0**

This contribution presents a GDR implementation for ECM 4.0 [1]. With this GDR implementation, there is no need to impose constraints on coding tools, which helps to reduce the code size for ECM significantly as compared to VTM-like GDR [3]. In addition, this new GDR implementation provides great flexibility in future tool development for ECM. Specifically, adding of additional new coding tools in ECM code will not affect part of GDR code.

Simulations were conducted under low-delay B configuration specified by LLCC AhG. The results demonstrate that the overall loss of the new GDR over anchor is only x.xx%.

**JVET-Z0141: AHG7: GDR in ECM-4.0**

This contribution presents results of GDR implementation for ECM4.0 [1]. GDR has been implemented in VTM [2] and then ported in ECM-3.0 [3] with new coding tools turned off. This GDR implementation has been tested with all new tools of ECM enabled.

***Sequences***

**JVET-Z0138: AHG7: Update on gaming sequences**

This contribution proposes an update on new gaming sequences proposed in JVET-Y-0041 at the last meeting. It also proposes to create a new class of content for gaming sequences.

**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 define configurations for low latency and controlled complexity, taking in account some of the suggestions made on the reflector.
* and to integrate a GDR implementation in the latest version of ECM.

[JVET-Z0008](https://jvet-experts.org/doc_end_user/current_document.php?id=11540) 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]

The AHG has used the main JVET reflector, jvet@lists.rwth-aachen.de, usually with [AHG8] in message headers, however no correspondence marked as AHG8 was sent between the 25th and 26th meetings.

The major area of work related to the AHG in this meeting cycle was the generation and cross-checking of conformance bitstreams for VVC version 2 in coordination with AHG5. A revision of JVET-Y2026 “Conformance testing for VVC operation range extensions (Draft 3)” was uploaded on 18th March 2022.

No contributions related to the study of high bit depth, high bit rate or high frame rate coding have been registered for the 26th meeting.

**Benchmarks**

The results in this section compare the performance of VTM16.0 against an anchor of VTM15.0. Results comparing VTM16.0 against VTM 14.0 are also included in this contribution as the previous AHG report did not provide benchmarks for VTM 15.0.

***Standard QP range***

The standard QP range covers QPs 22 to 37 for 12 bit coding.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access** | | | | | | | | | |
|  | **Over VTM15.0** | | | | | | | | | |
|  |  |  | **wPSNR** |  |  | **PSNR** |  |  |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | -0.77% | -1.48% | -1.05% | -1.91% | -2.09% | -0.76% | -1.35% | -1.47% | 93% | 100% |
| Class H2 |  |  |  |  |  | -0.71% | -1.54% | -1.32% | 92% | 100% |
| **Overall** | -0.77% | -1.48% | -1.05% | -1.91% | -2.09% | -0.73% | -1.44% | -1.40% | 92% | 100% |
|  |  |  |  |  |  |  |  |  |  |  |
|  | **All Intra** | | | | | | | | | |
|  | **Over VTM15.0** | | | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | -0.65% | -1.24% | -1.19% | -2.40% | -3.24% | -0.94% | -1.55% | -1.64% | 92% | 100% |
| Class H2 |  |  |  |  |  | -0.80% | -1.02% | -0.61% | 92% | 100% |
| **Overall** | -0.65% | -1.24% | -1.19% | -2.40% | -3.24% | -0.87% | -1.29% | -1.13% | 92% | 100% |

The majority of the BD-rate gains reported in these results can be attributed to the adoption of the updated deblocking configuration from ECM in JVET-Y0085 “AHG10: Report of Deblocking filter setting for VTM”.

The encoder time speed up can be attributed to the adoption of the software changes in JVET-Y0152 “AHG10: Fast skip of TT split partitioning on top of ECM reference software” and JVET-Y0126 “AHG10: VTM encoder configurations for tests targeting improved coding performance”.

***Low QP Range***

The low QP range covers QPs -13 to 12 for 12 bit coding, and -33 to -8 for 16 bit coding.

ss

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **HDR PQ** |  |  | **AI** |  |  |  |  |  |
|  |  |  | **Over VTM15.0** |  |  |  |  |  |
|  | wPsnrY | wPsnrU | wPsnrV | psnrY | psnrU | psnrV | EncT | DecT |
| PQ444 | 0.00% | 0.01% | 0.00% | 0.00% | 0.01% | 0.01% | 96% | 100% |
| PQ422 | 0.01% | 0.00% | 0.01% | 0.01% | 0.00% | 0.00% | 97% | 100% |
| **Overall** | 0.01% | 0.00% | 0.01% | 0.00% | 0.01% | 0.00% | 97% | 100% |
|  |  |  |  |  |  |  |  |  |
|  |  |  | **LDB** |  |  |  |  |  |
|  |  |  | **Over VTM15.0** |  |  |  |  |  |
|  | wPsnrY | wPsnrU | wPsnrV | psnrY | psnrU | psnrV | EncT | DecT |
| PQ444 | 0.01% | 0.00% | 0.01% | 0.00% | -0.03% | 0.00% | 97% | 99% |
| PQ422 | 0.01% | -0.03% | 0.00% | 0.01% | -0.04% | 0.00% | 98% | 99% |
| **Overall** | 0.01% | -0.02% | 0.01% | 0.01% | -0.03% | 0.00% | 98% | 99% |
|  |  |  |  |  |  |  |  |  |
|  |  |  | **RA** |  |  |  |  |  |
|  |  |  | **Over VTM15.0** |  |  |  |  |  |
|  | wPsnrY | wPsnrU | wPsnrV | psnrY | psnrU | psnrV | EncT | DecT |
| PQ444 | 0.00% | -0.02% | -0.01% | -0.01% | -0.02% | -0.02% | 97% | 99% |
| PQ422 | 0.00% | 0.00% | -0.02% | 0.00% | 0.00% | -0.02% | 97% | 99% |
| **Overall** | 0.00% | -0.01% | -0.02% | -0.01% | -0.01% | -0.02% | 97% | 99% |
|  |  |  |  |  |  |  |  |  |
| **Overall PQ** | 0.00% | -0.01% | 0.00% | 0.00% | -0.01% | 0.00% | 97% | 99% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **HDR HLG** |  |  | **AI** |  |  |
|  |  |  | **Over VTM15.0** |  |  |
|  | psnrY | psnrU | psnrV | EncT | DecT |
| HLG444 | 0.00% | 0.00% | 0.00% | 99% | 99% |
| HLG422 | 0.00% | 0.00% | 0.00% | 100% | 100% |
| **Overall** | 0.00% | 0.00% | 0.00% | 100% | 100% |
|  |  |  |  |  |  |
|  |  |  | **LDB** |  |  |
|  |  |  | **Over VTM15.0** |  |  |
|  | psnrY | psnrU | psnrV | EncT | DecT |
| HLG444 | -0.02% | -0.01% | 0.03% | 98% | 98% |
| HLG422 | -0.01% | 0.00% | 0.02% | 99% | 100% |
| **Overall** | -0.01% | 0.00% | 0.03% | 98% | 99% |
|  |  |  |  |  |  |
|  |  |  | **RA** |  |  |
|  |  |  | **Over VTM15.0** |  |  |
|  | psnrY | psnrU | psnrV | EncT | DecT |
| HLG444 | -0.02% | 0.00% | 0.00% | 97% | 99% |
| HLG422 | -0.01% | -0.01% | 0.01% | 99% | 100% |
| **Overall** | -0.02% | -0.01% | 0.00% | 98% | 100% |
|  |  |  |  |  |  |
| **Overall HLG** | -0.01% | 0.00% | 0.01% | 99% | 100% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **SVT RGB** |  |  | **AI** |  |  |
|  |  |  | **Over VTM15.0** |  |  |
|  | psnrG | psnrB | psnrR | EncT | DecT |
| SVT16 | 0.00% | 0.00% | 0.00% | 99% | 100% |
| SVT12 | 0.00% | 0.00% | 0.00% | 100% | 101% |
| **Overall** | 0.00% | 0.00% | 0.00% | 100% | 101% |
|  |  |  |  |  |  |
|  |  |  | **LDB** |  |  |
|  |  |  | **Over VTM15.0** |  |  |
|  | psnrG | psnrB | psnrR | EncT | DecT |
| SVT16 | 0.00% | 0.00% | 0.00% | 100% | 100% |
| SVT12 | 0.00% | -0.12% | -0.12% | 101% | 101% |
| **Overall** | 0.00% | -0.06% | -0.06% | 100% | 101% |
|  |  |  |  |  |  |
|  |  |  | **RA** |  |  |
|  |  |  | **Over VTM15.0** |  |  |
|  | psnrG | psnrB | psnrR | EncT | DecT |
| SVT16 | 0.00% | 0.00% | 0.00% | 99% | 100% |
| SVT12 | 0.00% | -0.06% | -0.06% | 99% | 100% |
| **Overall** | 0.00% | -0.03% | -0.03% | 99% | 100% |
|  |  |  |  |  |  |
| **Overall RGB** | 0.00% | -0.03% | -0.03% | 100% | 100% |

**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-Z0009](https://jvet-experts.org/doc_end_user/current_document.php?id=11541) JVET AHG report: SEI message studies (AHG9) S. McCarthy, Y.-K. Wang, T. Chujoh, C. Fogg, P. de Lagrange, G. J. Sullivan, A. Tourapis, S. Wenger]

A total of 9 contributions, not including cross-checks, are identified relating to the mandates of AHG9. Some contributions relate to more than one mandate. The number of contributions relating to each mandate is as follows

* 2 contributions relate to the mandate to study the SEI messages in VSEI, VVC, HEVC, and AVC;
* 3 contributions relate to the mandate to study signalling of essential resampling phase indication and neural network-based post filtering, and prepare draft text for such signalling;
  + 0 contributions relate to signalling of essential resampling phase indication
  + 3 contributions relate to signalling of neural network-based post filtering
* 3 contributions relate to the mandate to collect software and showcase information for SEI messages, including encoder and decoder implementations and bitstreams for demonstration and testing;
* 1 contribution relates 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
* 1 contribution relates 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 AHG9.

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

[JVET-Z0047](https://jvet-experts.org/doc_end_user/current_document.php?id=11481) AHG13: Improvements of film grain analysis [P. de Lagrange, E. Francois, Z. Ameur (InterDigital), M. Radosavljević (Xiaomi)]

(JVET-Z0047 also relates to the mandate to collect software and showcase information)

[JVET-Z0115](https://jvet-experts.org/doc_end_user/current_document.php?id=11562) AHG13: Proposed film grain synthesis reference model [P. de Lagrange, F. Urban, E. François (InterDigital)]

***Study signalling of essential resampling phase indication and neural network-based post filtering, and prepare draft text for such signalling***

[JVET-Z0052](https://jvet-experts.org/doc_end_user/current_document.php?id=11486) AHG9: NNR post-filter SEI message [M. M. Hannuksela, M. Santamaria, F. Cricri, E. B. Aksu, H. R. Tavakoli (Nokia), T. Chujoh, Y. Yasugi, T. Ikai (Sharp)]

[JVET-Z0121](https://jvet-experts.org/doc_end_user/current_document.php?id=11569) AHG9: Neural-network post filtering SEI message [S. McCarthy, A. Arora, T. Shao, P. Yin, T. Lu, F. Pu, W. Husak (Dolby)]

[JVET-Z0151](https://jvet-experts.org/doc_end_user/current_document.php?id=11600) AHG9: On NNR post-filter SEI message [ Y. Yasugi, T. Chujoh, T. Ikai (Sharp)]

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

[JVET-Z0046](https://jvet-experts.org/doc_end_user/current_document.php?id=11480) VTM Software Implementation for GREEN-MPEG SEI Messaging [C. Herglotz, M. Kränzler, A. Kaup (FAU)]

[JVET-Z0047](https://jvet-experts.org/doc_end_user/current_document.php?id=11481) AHG13: Improvements of film grain analysis [P. de Lagrange, E. Francois, Z. Ameur (InterDigital), M. Radosavljević (Xiaomi)]

(JVET-Z0047 also relates to the mandate to study the SEI messages in VSEI, VVC, HEVC and AVC)

[JVET-Z0120](https://jvet-experts.org/doc_end_user/current_document.php?id=11568) AHG9: Shutter interval information SEI message for VSEI [S. McCarthy, F. Pu, W. Husak, P. Yin, T. Lu, A. Arora, T. Shao (Dolby), J.R. Arumugam, S. Agrawal, K. Patankar (Ittiam)]

(JVET-Z0120 also relates to the mandate to study SEI messages in HEVC and AVC for use in the context of VVC.)

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

[JVET-Z0129](https://jvet-experts.org/doc_end_user/current_document.php?id=11577) AHG9: SEI for transparency information as dedicated layer for transparent screens [E.Thomas, P. Andrivon, F. Le Leannec, M. Radosavljević, M-L. Champel (Xiaomi)]

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

[JVET-Z0132](https://jvet-experts.org/doc_end_user/current_document.php?id=11581) AHG13: On film grain synthesis [Y. He, M. Coban, M. Karczewicz (Qualcomm), P. de Lagrange, E. François (InterDigital), M. Radosavljević (Xiaomi), S. McCarthy, W. Husak (Dolby)]

***Study SEI messages defined in HEVC and AVC for potential use in the VVC context***

[JVET-Z0120](https://jvet-experts.org/doc_end_user/current_document.php?id=11568) AHG9: Shutter interval information SEI message for VSEI [S. McCarthy, F. Pu, W. Husak, P. Yin, T. Lu, A. Arora, T. Shao (Dolby), J.R. Arumugam, S. Agrawal, K. Patankar (Ittiam)]

(JVET-Z0120 also relates to the mandate to collect software and showcase information)

**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-Z0010](https://jvet-experts.org/doc_end_user/current_document.php?id=11542) JVET AHG report: Encoding algorithm optimization (AHG10) [P. de Lagrange, A. Duenas, R. Sjöberg, A. Tourapis]

A total of 7 contributions, not including cross-checks, are identified relating to AHG10, and summarized in the following sections.

***Tool adaptation and configuration***

**JVET-Z0072 AHG10/AHG12: Enhanced reference picture structures for ECM and VTM**

This contribution proposes to update CTCs for VTM and ECM with a modified reference picture list and QP settings for both RA and LDB configurations:

* For RA, it suggests adding references pictures while removing duplicates, like in the alternative configuration of VTM (high performance / high complexity), and listing the active reference picture in order of POC distance. This is said to improve VTM BD-rate by 0.5% with about 5% encoding time increase. ECM results are similar.
* For LDB, it suggests storing more pictures in the DBP (while keeping the same number of active reference pictures), and to adjust QP offsets by +1 for every second picture in the GOP as a compromise to lower encoding time. More specifically, some pictures in L1 already present in L0 are replaced by additional reference pictures. This is said to improve VTM BD-rate by 1.6%, with mostly unchanged encoding time. ECM results are similar.

**JVET-Z0111 Adaptively bypass affine ME in VTM**

This contribution reports about trading off performance related to affine motion compensation for faster encoding time. Two options are tested, with a different tradeoff (the major difference being on low-delay). Affine motion estimation is skipped at encoder side, depending on local context. The loss is around 0.2% for 5% faster encoding.

***Encoding techniques of optimization for objective quality metrics and their relationship to subjective quality***

**JVET-Z0099 AHG10: Deblocking in RDO and beta offset minus 2 for VTM**

This contribution proposes a slightly stronger deblocking setting for VTM, together with enabling deblocking in RDO, as in the alternative configuration of VTM (high performance / high complexity).

Enabling deblocking in RDO is reported to improve BD-rate by 0.51% / 0.27% for RA / LDB respectively, while increasing encoding time by about 5%. Stronger settings bring an additional 0.05% for RA and 0.14% for LDB.

**JVET-Z0108 Evaluation of 17 objective quality metrics on JVET contents**

While a related contribution (JVET-Z0109) investigates the correlation between several objective quality metrics and MOS scores, this one explores in what extent this correlation actually has an impact on the agreement of decisions taken based on these metrics or visual observation.. It reports about various experiments where SSIM and VMAF are often the most reliable metrics for taking decisions, with PSNR not so far behind, while neural networks-based metrics do not stand out clearly.

***Optimized encoding for reference picture resampling and scalability modes in VTM***

**JVET-Z0065 EE1-2.1: RPR encoder with multiple scale factors**

This contribution reports the test results of adaptive selection of picture resolution for VVC RPR functionality without multi-pass coding at the encoder side: it allows encoder to adaptively choose one of three scale factors, {x2.0 (half), x1.5 (2/3) and x1.0 (origin)}, at GOP level based on the PSNR and initial QP for the first picture of each GOP. Another test further includes QP adjustments for both luma and chroma components. The BD-rate improvement in random-access from the first test is around 1% for 10% less encoding time, and the second test further improve that by 0.4% and again 10% less encoding time.

***Investigate other methods of improving objective and/or subjective quality***

**JVET-Z0107 Quality evaluation of internal 10-bit versus 8-bit processing with 8-bit source content**

This contribution evaluates the impact of 10-bit coding an 8-bit content, both in term of objective metrics (of several types) and visual assessment by expert viewers. 10-bit processing is reported to bring some benefit, both visually and objectively; the gain is higher in low-delay, and no negative impact was observed.

***Study the effect of varying configuration parameters depending on temporal layer***

**JVET-Z0104 AHG10: Report of temporal layer dependent deblocking filter setting for HM**

This contribution reports about the objective results when applying the deblocking filter setting of VTM-16 to HM, more specifically adjusting BetaOffset\_div2 and TcOffset\_div2 based on temporal layer.

A BD-rate performance drop is observed (more than 0.6% in RA, up to 1.6% in LDP).

**Recommendation**

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

[JVET-Z0011](https://jvet-experts.org/doc_end_user/current_document.php?id=11543) 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]

***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 Februrary 4, 2022.

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

|  |  |  |
| --- | --- | --- |
| JVET-Z0023 | 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 the previous meeting and 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.

***Coordination with SC29/AG5***

The AHG and representatives of SC29/AG5 held a teleconference on February 21, 2022, to prepare a viewing procedure for EE contributions. A report on the teleconference is provided in JVET-Z0045. Results from the viewing procedure are provided in JVET-Z0053.

***Study and Maintain SADL***

The SADL library was updated during the AHG period. A first update was announced on the reflector on January 25, 2022. The software is available at https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/sadl.

Change logs include:

*Commit 6216f7cf*

*General in both converter and SADL:*

*- added shape and expand layers*

*- dump pads information in conv2d and maxpool in order to handle more stride/kernel size combination*

*- update to version v2 file format (because of above change)*

*- draft of documentation in LaTeX format added*

*\* converter:*

*- add a tranpose layer after some layers because of the memory layout change*

*- add support for prelu -> map to leakyrelu in sadl*

*\* sadl:*

*- allow up to dimension 6 for tensor (was 4), but only for internal data layout manipulation*

*- increase the maximum number of layers to 512*

*- conv2d, maxpool: rely on onnx padding information to compute kernel offset because TF and pytorch have different policies.*

*- added generic naive algorithm for transpose*

*Commit 8f6ed324*

*Fix for MACOS*

*Commit 5bd13239*

*Added parameters to control data layout layers during conversion. make transpose generic in all cases*

to provide support for the module being studied in EE1-1.2. The update was announced on the reflector on January 25, 2022, and made available

***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 26th meeting is provided below:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Title** | **Common Test Conditions** | **Results** | | | **Training Data** | |
| **RA** | **LDB** | **AI** | **CTC** | **Additional** |
| **Loop Filter** | | | | | | | |
| JVET-Z0089 | AHG11: An Improved Unet-Based In-Loop Filter Method | Yes | No | No | Yes | BVI | DIV2K |
| **Post Filtering** | | | | | | | |
| JVET-Z0052 | AHG9: NNR post-filter SEI message | No | No | No | No | - | - |
| JVET-Z0082 | AHG11: Content-adaptive neural network post-filter | Yes | Yes | No | No | BVI | DIV2K |
| JVET-Z0101 | AHG11: Post-process filter based on fusion of CNN and transformer | Yes | No | No | No | BVI | - |
| JVET-Z0121 | AHG9: Neural-network post filtering SEI message | No | No | No | No | - | - |
| JVET-Z0144 | AHG11: CNN-Based Post-Processing Filter for Video Compression with Multi-Scale Feature Representation | Yes | No | No | No | - | DIV2K |
| JVET-Z0151 | AHG9: On NNR post-filter SEI message | No | No | No | No | - | - |
| **Super-Resolution** | | | | | | | |
| JVET-Z0088 | AHG11: A CNN-based Super Resolution Method Combined with Existing RPR Functionality | No | No | No | No | BVI-DVC | - |
| **Inter-Prediction** | | | | | | | |
| JVET-Z0074 | AHG11: Neural Network Based Motion Compensation Enhancement for Video Coding | Yes | Yes | No | No | BVI-DVC |  |
| **End-to-End** | | | | | | | |
| JVET-Z0077 | AHG11: Extension of DOVC to Regular 2D Videos | No | - | - | - | BVI | DIV2K |

**Input contributions**

There are 34 input contriubtions related to the AHG mandates. Twenty of the contributions are part of the EE activity, while the remaining 14 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-Z0023 | 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-Z0065 | EE1-2.1: RPR encoder with multiple scale factors | J. Nam, S. Yoo, J. Lim, S. Kim (LGE) |
| JVET-Z0070 | EE1-1.3: Combination of deblocking and NN | K. Andersson, J. Ström, D. Liu, R. Sjöberg (Ericsson) |
| JVET-Z0073 | EE1-1.5: Test on NN-based filter as proposed in JVET-Y0052 | H. Zhang, C. Jung (Xidian Univ.), D. Zou, M. Li (OPPO) |
| JVET-Z0086 | EE1-1.4: ALF improvement for NNVC | W. Zou, Y. Zhou (Xidian Univ.), C. Huang, Y. X. Bai (ZTE) |
| JVET-Z0091 | EE1-1.2: Neural network based in-loop filter with a single model | L. Wang, X. Xu, S. Liu (Tencent), F. Galpin (InterDigital) |
| JVET-Z0094 | EE1-1.1: Neural network based in-loop filter with 2 models | L. Wang, X. Xu, S. Liu (Tencent), F. Galpin (InterDigital) |
| JVET-Z0096 | EE1-2.2: CNN-based Super Resolution for Video Coding Using Decoded Information | C. Lin, Y. Li, K. Zhang, L. Zhang (Bytedance) |
| JVET-Z0097 | EE1-2.3: CNN-based Super Resolution for Video Coding Using Separate Networks for Chroma Components | C. Lin, Y. Li, K. Zhang, L. Zhang (Bytedance) |
| JVET-Z0098 | EE1-2.4: CNN-based Super Resolution for Video Coding with GOP Level Adaptive Resolution | C. Lin, Y. Li, K. Zhang, L. Zhang (Bytedance), J. Nam, S. Yoo, J. Lim, S. Kim (LGE) |
| JVET-Z0112 | EE1-1.6: Test on Deep In-Loop Filter with Adaptive Parameter Selection and Residual Scaling based on SADL implementation | Y. Li, K. Zhang, L. Zhang (Bytedance), H. Wang, M. Coban, A. M. Kotra, M. Karczewicz (Qualcomm), F. Galpin (InterDigital) |
| JVET-Z0113 | EE1-1.7: Combined Test of EE1-1.6 and EE1-1.3 | Y. Li, K. Zhang, 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) |
| **EE Related** | | |
| JVET-Z0087 | EE1-related: ALF-SPLIT based on JVET-Z0070 and JVET-Y0078 | W. Zou, Y. Zhou, C. M. Gu (Xidian Univ.), C. Huang, Y. X. Bai, C.G. Liu, X.C. Zhu (ZTE) |
| JVET-Z0092 | EE1-related: Additional results on Test 1.1 and Test 1.2 with 8-bit quantization | L. Wang, X. Xu, S. Liu (Tencent), F. Galpin (InterDigital) |
| JVET-Z0093 | EE1-related: The performance of EE1 Test 1.1 and Test 1.2 on ECM-4.0 | R. Chang, L. Wang, X. Xu, S. Liu (Tencent), F. Galpin (InterDigital) |
| JVET-Z0106 | EE1-related: Reduced complexity NN loop filter and ablation study | J. Ström, D. Liu, M. Damghanian, K. Andersson, Y. Li, P. Wennersten, R. Yu (Ericsson) |
| JVET-Z0128 | EE1-related: on BaseQP parameter in EE1-1.1 | Z. Xie, Y. Yu, H. Yu, D. Wang (OPPO) |
| JVET-Z0154 | EE1-1.6-related: Improved RDO Considering Deep In-Loop Filter | J. Li, Y. Li, K. Zhang, L. Zhang (Bytedance) |
| JVET-Z0155 | EE1-1.7-related: Improved RDO Considering Deep In-Loop Filter and Deblocking | J. Li, Y. Li, K. Zhang, L. Zhang (Bytedance) |
| **Cross Checks** | | |
| JVET-Z0071 | Cross-check of JVET-Z0065: EE1-2.1: RPR encoder with multiple scale factors | K. Andersson (Ericsson) |

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

|  |  |  |
| --- | --- | --- |
| **Reporting** | | |
| JVET-Z0011 | 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-Z0045 | AhG11/AhG4/EE1 viewing preparation report | E. Alshina, M. Wien, A. Segall, J. Sauer (coordinators) |
| JVET-Z0053 | [AHG4] REV Result for AHG11/EE1 and AHG4 new test sequences | M. Wien (RWTH) |
| **Loop Filtering** | | |
| JVET-Z0089 | AHG11: An Improved Unet-Based In-Loop Filter Method | C. Fang, J. Lin, D. Jiang, X. Zhang, S. Peng (Dahua) |
| **Post Filtering** | | |
| JVET-Z0052 | AHG9: NNR post-filter SEI message | M. M. Hannuksela, M. Santamaria, F. Cricri, E. B. Aksu, H. R. Tavakoli (Nokia), T. Chujoh, Y. Yasugi, T. Ikai (Sharp) |
| JVET-Z0082 | AHG11: Content-adaptive neural network post-filter | M. Santamaria, R. Yang, F. Cricri, J. Lainema, R. G. Youvalari, H. Zhang, G. Rangu, H. R. Tavakoli, H. Afrabandpey, M. M. Hannuksela (Nokia) |
| JVET-Z0101 | AHG11: Post-process filter based on fusion of CNN and transformer | T. Liu, W. Cui, C. Hui, F. Jiang (Harbin Inst. Tech.), Y. Gao, S. Xie, P. Wu (ZTE) |
| JVET-Z0121 | AHG9: Neural-network post filtering SEI message | S. McCarthy, A. Arora, T. Shao, P. Yin, T. Lu, F. Pu, W. Husak (Dolby) |
| JVET-Z0144 | AHG11: CNN-Based Post-Processing Filter for Video Compression with Multi-Scale Feature Representation | Z. Qi, C. Jung (Xidian Univ.), Y. Liu, M. Li (OPPO) |
| JVET-Z0151 | AHG9: On NNR post-filter SEI message | Y. Yasugi, T. Chujoh, T. Ikai (Sharp) |
| **Super Resolution** | | |
| JVET-Z0088 | AHG11: A CNN-based Super Resolution Method Combined with Existing RPR Functionality | S. Peng, D. Jiang, J. Lin, C. Fang, X. Zhang (Dahua) |
| **Inter Prediction** | | |
| JVET-Z0074 | AHG11: Neural Network Based Motion Compensation Enhancement for Video Coding | C. Ma, R.-L. Liao, Y. Ye (Alibaba) |
| **End-to-End** | | |
| JVET-Z0077 | AHG11: Extension of DOVC to Regular 2D Videos | Q. Qin, C. Jung (Xidian Univ.), D. Zou, M. Li (OPPO) |
| **Software** | | |
| JVET-Z0161 | AhG11: SADL update | F. Galpin, T. Dumas, P. Bordes, E. François (InterDigital) |
| **Cross Checks** | | |
|  |  |  |

**Recommendations**

The AHG recommends:

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

[JVET-Z0012](https://jvet-experts.org/doc_end_user/current_document.php?id=11544) JVET AHG report: Enhanced compression beyond VVC capability (AHG12) [M. Karczewicz, Y. Ye, L. Zhang, B. Bross, X. Li, K. Naser, H. Yang]

The Common Test Conditions were updated (JVET-Y2017). The test conditions for HDR were added and the checksums for some of the test sequences were corrected. The primary activity of the AHG was the “Exploration experiment on enhanced compression beyond VVC capability” (JVET-Y2024). The combined improvements of the ECM-4.0 over VTM-11.0ecm anchorfor AI, RA and LB configurations are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | All Intra Main10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -6.60% | -12.43% | -17.68% | 357% | 262% |
| Class A2 | -6.24% | -13.37% | -14.77% | 342% | 254% |
| Class B | -5.70% | -14.30% | -13.65% | 382% | 248% |
| Class C | -6.89% | -9.30% | -9.80% | 373% | 242% |
| Class E | -7.67% | -12.17% | -13.14% | 363% | 264% |
| Overall | -6.53% | -12.37% | -13.57% | 366% | 253% |
| Class D | -5.73% | -8.03% | -7.65% | 371% | 241% |
| Class F | -10.93% | -15.45% | -14.70% | 253% | 255% |
| Class TGM | -16.70% | -18.91% | -18.31% | 245% | 276% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Random Access Main 10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -16.20% | -17.76% | -23.60% | 389% | 471% |
| Class A2 | -16.97% | -22.37% | -24.24% | 375% | 546% |
| Class B | -14.46% | -21.68% | -19.89% | 403% | 505% |
| Class C | -16.40% | -17.23% | -17.19% | 426% | 504% |
| Class E |  |  |  |  |  |
| Overall | -15.83% | -19.85% | -20.78% | 400% | 506% |
| Class D | -16.99% | -17.54% | -17.59% | 411% | 523% |
| Class F | -14.23% | -17.29% | -17.00% | 349% | 363% |
| Class TGM | -16.11% | -19.04% | -19.39% | 387% | 299% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Low Delay B Main 10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -12.29% | -26.50% | -24.33% | 351% | 397% |
| Class C | -13.71% | -20.74% | -19.80% | 365% | 378% |
| Class E | -11.18% | -18.16% | -19.41% | 344% | 299% |
| Overall | -12.48% | -22.49% | -21.59% | 353% | 364% |
| Class D | -15.49% | -22.79% | -21.74% | 348% | 387% |
| Class F | -12.98% | -20.76% | -18.58% | 325% | 296% |
| Class TGM | -15.26% | -22.73% | -22.40% | 370% | 292% |

The rate reduction over VTM for RA configuration for {Y, U, V} increased from {-14.77%, -18.61%, -19.42%} to {-15.83%, -19.85%, -20.78%}.

**Contributions**

In addition to 21 EE2 contributions, 38 (comparing to 41 last meeting) related contributions were received which can be subdivided as follows:

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

JVET-Z0105, "AHG12: Virtual boundary processing for the new In-Loop filter tools in ECM", A. M. Kotra, N. Hu, V. Seregin, M. Karczewicz (Qualcomm)

JVET-Z0146, "AHG12: Using samples before deblocking filter for adaptive loop filter", N. Hu, V. Seregin, M. Karczewicz (Qualcomm)

JVET-Z0149, "Non-EE2: Adaptive Filter Shape Switch for ALF", W. Yin, K. Zhang, L. Zhang (Bytedance)

***Intra (5)***

JVET-Z0064, "AHG12: Convolutional cross-component model (CCCM) for intra prediction", P. Astola, J. Lainema, R. G. Youvalari, A. Aminlou, K. Panusopone (Nokia)

JVET-Z0100, "Non-EE2: Weighted Chroma Prediction (WCP)", J.-Y. Huo, H.-Q. Du, Z.-Y. Zhang, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.), M. Li, Y. Liu (OPPO)

JVET-Z0124, "Non-EE2: Spatial GPM", F. Wang, Y. Yu, H. Yu, D. Wang (OPPO)

JVET-Z0140, "AHG12: Enhanced CCLM", C.-W. Kuo, X. Xiu, N. Yan, H.-J. Jhu, W. Chen, H. Gao, X. Wang (Kwai)

JVET-Z0143, "Non-EE2: Chroma intra modes derived from collocated luma blocks and neighboring chroma blocks", Y.-J. Chang, K. Cao, B. Ray, V. Seregin, M. Karczewicz (Qualcomm) Rufitskiy, E. Dinan (Ofinno)

***Inter (20)***

JVET-Z0059, "Non-EE2: Adaptive width for GPM blending area", Y. Kidani, H. Kato, K. Unno, K. Kawamura (KDDI)

JVET-Z0063, "AHG12: Fix for parsing of MHP information in ECM", R. Yu (Ericsson)

JVET-Z0066, "EE2-related: Template matching using extended MVP candidate list", K. Kim, D. Kim, J.-H. Son, J.-S. Kwak (WILUS)

JVET-Z0068, "EE2-2.6 related: Longer chroma interpolation filters", K. Andersson, R. Yu (Ericsson)

JVET-Z0069, "AHG12: Longer chroma filters for RPR in ECM", K. Andersson, R. Yu (Ericsson)

JVET-Z0078, "Non-EE2: Support MMVD and Affine MMVD without template matching process", H. Jang, J. Nam, N. Park, J. Lim, S. Kim (LGE)

JVET-Z0083, "Non-EE2: Fix on MHP parsing condition", N. Park, J. Nam, J. Lim, S. Kim (LGE)

JVET-Z0085, "Non-EE2: AmvpMerge without decoder side mv refinement process", H. Jang, J. Nam, N. Park, J. Lim, S. Kim (LGE)

JVET-Z0102, "Non-EE2: Correction of the ARMC re-ordering step regarding the zero candidates", G. Laroche, P. Onno, R. Bellessort (Canon)

JVET-Z0103, "Non-EE2: On ARMC improvements", G. Laroche, P. Onno (Canon)

JVET-Z0123, "EE2-related: On chroma interpolation filters for motion compensation", J. Gan, Y. Yu, H. Yu (OPPO)

JVET-Z0125, "EE2-related: Regression based affine candidate derivation", Y. Zhang, H. Huang, V. Seregin, M. Coban, M. Karczewicz (Qualcomm)

JVET-Z0126, "Non-EE2: Improvement on Local Illumination Compensation", Z. Xie, Y. Yu, H. Yu, D. Wang (OPPO)

JVET-Z0127, "Non-EE2: On the maximum number of MHP merge candidates", H. Huang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-Z0130, "EE2-related: Motion compensation boundary padding", Z. Zhang, H. Huang, C.-C. Chen, Y.-J. Chang, V. Seregin, M. Coban, M. Karczewicz (Qualcomm), F. Le Léannec, P. Andrivon, M. Radosavljevć, E.Thomas (Xiaomi)

JVET-Z0137, "Non-EE2: Adaptive Blending for GPM", H. Gao, X. Xiu, W. Chen, H.-J. Jhu, C.-W. Kuo, N. Yan, X. Wang (Kwai)

JVET-Z0142, "EE2-related: On MMVD and Affine MMVD Extension", M. Salehifar, Y. He, K. Zhang, N. Zhang, L. Zhang (Bytedance)

JVET-Z0145, "EE2-2.5 related: More test results for ARMC with refined motion", Y. Wang, K. Zhang, N. Zhang, Z. Deng, L. Zhang (Bytedance)

JVET-Z0148, "EE2-2.5 related: AMVP with MVP refinement and ARMC with refined motion", C. Zhou, Z. Lv, J. Zhang (vivo), Y. Wang, K. Zhang, N. Zhang, Z. Deng, L. Zhang (Bytedance)

JVET-Z0184, "EE2-3.5-related: Additional results of zero-vector candidates replacement in the IBC Merge/AMVP list", D. Ruiz Coll, A. Filippov, V. Rufitskiy (Ofinno)

***RPR (3)***

JVET-Z0062, "AHG12: RPR luma filter modifications and RPR enabling fixes in ECM", R. Yu, K. Andersson (Ericsson)

JVET-Z0067, "Non-EE2: RPR bug fixes and new results of RPR with luma-only re-scaling", P. Bordes, F. Galpin, H. Guermoud, E. Francois (InterDigital)

JVET-Z0079, "Non-EE2: Support RPR on ECM-4.0 and handling method for template matching based coding tools", H. Jang, J. Nam, N. Park, J. Lim, S. Kim (LGE)

***Transform Coding (2)***

JVET-Z0048, "EE2-related: Modification of LFNST for MIP coded block", J.-Y. Huo, W.-H. Qiao, X. Hao, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.), M. Li, Y. Liu (OPPO)

JVET-Z0147, "Non-EE2: Improvement of Inter-MTS in ECM", B. Ray, V. Seregin, M. Karczewicz (Qualcomm)

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

JVET-Z0131, "EE2-related: Block Vector Difference Binarization", K. Cao, V. Seregin, M. Karczewicz (Qualcomm)

JVET-Z0152, "EE2-related: IBC Merge Mode with Block Vector Differences", N. Zhang, J. Xu, K. Zhang, M. Salehifar, L. Zhang (Bytedance)

JVET-Z0159, "Non-EE2: Reconstruction-Reordered IBC for screen content coding", Z. Deng, K. Zhang, L. Zhang (Bytedance)

***ECM Encoder and Test Conditions (2)***

JVET-Z0072, "AHG10/AHG12: Enhanced reference picture structures for ECM and VTM", K. Andersson, R. Sjöberg, R. Yu, L. Litwic (Ericsson)

JVET-Z0150, "Memory usage report on VTM / ECM", T. Hashimoto, Y. Yasugi, T. Ikai (Sharp)

**Recommendations**

The AHG recommends to:

* To review all the related contributions.

It is noted that the perspective of this exploration should be discussed at some time during the meeting.

[JVET-Z0013](https://jvet-experts.org/doc_end_user/current_document.php?id=11545) JVET AHG report: Film grain technologies (AHG13) [W. Husak, M. Radosavljević, W. Wan, D. Grois, A. Tourapis]

Four contributions related to AHG13 were identified as of 04/19/2022:

* One was the AHG report:
  + JVET-Z0013 JVET AHG report: Film grain technologies (AHG13)
* Three are related to proposing improvements to the existing film grain blending process:
  + JVET-Z0047 AHG13: Improvements of film grain analysis
  + JVET-Z0115 AHG13: Proposed film grain synthesis reference model
  + JVET-Z0132 AHG13: On film grain synthesis

***Contributions***

**JVET-Z0047 AHG13: Improvements of film grain analysis**

This contribution proposes improvements for the film grain analysis part of the VTM. It introduces improved film grain scaling factor estimation and introduces two new options in the VTM code.

**JVET-Z0115 AHG13: Proposed film grain synthesis reference model**

This contribution describes a film grain synthesis model that supports both modes of FGS SEI (frequency-filtering and autoregressive), as well as other parameter formats, with attention paid to hardware complexity. It is proposed to study this model in ad-hoc group on film grain technologies, as a candidate for the reference model of a potential “soft conformance” definition.

**JVET-Z0132 AHG13: On film grain synthesis**

This contribution proposes study of film grain synthesis and relevant conformance requirements in AHG13 with the following items:

1. Conformance requirements

2. Auto-regressive model reference software development

3. Reference film grain synthesis process

**Recommendation**

The AHG recommends:

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

# Project development (25)

## Deployment and advertisement of standards (2)

Contributions in this area were discussed in session 16 at 0600–XXXX UTC on Thursday 28 April 2022 (chaired by JRO).

[JVET-Z0020](https://jvet-experts.org/doc_end_user/current_document.php?id=11490) 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-Y0020-v1 of January 2022.(include changes relative to JVET-Y0020 from v2)

[JVET-Z0021](https://jvet-experts.org/doc_end_user/current_document.php?id=11566) Deployment status of the VVC standard [G. J. Sullivan (Microsoft)]

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

Following changes relative to JVET-Y0021-v4 of January 2022:(include changes relative to JVET-Y0020 from v2)

## Text development and errata reporting (2)

Contributions in this area were discussed in session 7 at 0500–0600 UTC on Friday 22 April 2022 (chaired by JRO).

At the beginning of the session, it was discussed that the two tickets mentioned in AHG2 report:

* [#1547](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1547) sb\_coded\_flag non-present inference logic error
* [#1544](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1544) sb\_coded\_flag definition lacks size of a sub-block

as well as the new ticket #1548 need to be resolved.

The following documets were provided after offline consideration of experts with knowledge on these items with editors.

[JVET-Z0249](https://jvet-experts.org/doc_end_user/current_document.php?id=11700) Proposed Fix for Ticket #1547: sb\_coded\_flag non-present inference [T. Nguyen (HHI)] [late]

In ticket #1547 (https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1547), it was reported that the semantic for the sb\_coded\_flag syntax element in the current specification is different from the implementation in the VTM. Investigations indicate that the semantics for the sb\_coded\_flag syntax element in the regular residual coding mode was probably accidentally removed. Thus, the behavior in the VTM software is intended and correct, and only the specification needs to be revised. This document provides the modifications to the specification to fix the incorrect semantics for the sb\_coded\_flag syntax element.

[JVET-Z0250](https://jvet-experts.org/doc_end_user/current_document.php?id=11702) Alternate proposed fix for ticket #1547 and crosscheck of JVET-Z0249 [J. Gan, Y. Yu (OPPO)] [late]

VVC bug ticket #1547 reports mismatch between the VVC specification and the VTM reference software for the semantics of the sb\_coded\_flag syntax element. The mismatch was confirmed on the bug tracker by multiple experts, and it was agreed preferable for the specification to be aligned with the software.

JVET-Z0249 was uploaded as a late contribution to address this issue. However, this contribution asserts that there are some errors in the proposed text of JVET-Z0249. This contribution proposes alternative text which is asserted to fix the issue of ticket #1547 without introducing new inconsistencies.

Was presented in session 18 2110 UTC

There is an obvious deviation between software and spec related to sb\_coded\_flag. Its values may be inferred, and also zero inferred values can be used as context in CABAC.

It is generally agreed that the text shall be aligned with the software, as at some point in text editing the text deviated from what it was intended to be.

Based on the text versions delivered in JVET-Z0249 and JVET-Z0250, editors (B. Bross et al.) should take an appropriate action in coordination with the proponents of these documents and produce a text to be included into JVET-Z2019.

Providing a solution to tickets #1544 and #1548 appears straightforward and is left to discretion of editors.

Decision(BF/text): Editors’ action items as described above. For ticket #1547, text should be aligned with the software, in coordination with the proponents of JVET-Z0249 and JVET-Z0250. For ticket #1544, the missing definition of sub-block size needs to be added. For ticket #1548, the constraint that the payload must be identical for multiple SEI messages shall be relaxed, as they may relate to different layers or different subpictures.

Decision: During the plenary Friday 29 April, 0915, an expert raised an issue about a new bug report (ticket #1549) about the IsAvailable variable setting for luma component. The corresponding fix is asserted to be obvious and shall be done by editors’ discretion in draft 2 of the VVC extension.

[JVET-Z0119](https://jvet-experts.org/doc_end_user/current_document.php?id=11567) Editorial status of alpha blending text and ITU-T Last Call for VSEI v2 [G. J. Sullivan (Microsoft), Y.-K. Wang (Bytedance)]

This contribution discusses the editorial status of alpha blending text description, particularly for VVC v2 text preparation. It reports that further refinement f the editorial description of this part of the text for VVC v2 was conducted after text was submitted for ITU-T Last Call but before corresponding text was submitted for ISO/IEC FDIS, and reports a comment may be submitted in the ITU-T Last Call comment period to request such further refinement to be included in the ITU-T version. Change-marked text is provided to show the differences.

It was agreed that alignment of the text between ISO and ITU has higher importance than the publication delay of a few weeks that will occur when a last call comment is made.

Experts are asked to carefully study the text and contact Gary Sullivan in case of inconsistencies found.

Corresponding changes should also be applied in the next versions of HEVC and AVC. It is noted that the clarification of the clipping process for opaque and transparent indicators (which is a more technical change rather than a clarification) is not relevant for AVC.

[JVET-Z0122](https://jvet-experts.org/doc_end_user/current_document.php?id=11570) Some HEVC text changes [Y.-K. Wang (Bytedance), G. J. Sullivan (Microsoft)]

This contribution provides some HEVC text changes, for the unlimited level that was agreed at the previous JVET meeting but the text changes were missing from the output documents, as well as for some minor typo corrections.

The changes are provided in attachment, based on the HEVC attachment in JVET-Y1004, with change marks to show the differences. The new changes are marked by the user ID "Ye-Kui Wang (yk1)".

It was agreed that this is an obvious action item to implement the last meeting’s decision.

It was asked of the limitation in 7.4.3.10 about constraining the number of bins per CU should still apply in the unlimited level. It was agreed that relaxing this constraint is desirable for the unlimited level. Experts are also asked to carefully check the text for other places where constraints are built in text/semantics rather than syntax.

It was agreed to rephrase the proposed note in annex A to avoid that a note imposes normative behaviour.

The same modification also needs to be applied to the VVC text of unlimited level.

## Test conditions (1)

Contributions in this area were discussed in session 12 at 04510–0540 UTC on Tuesday 26 April 2022 (chaired by JRO).

JVET-Z0157 (already presented in BoG JVET-Z0210) is also related to test conditions. It requests to enable IBC also for camera captured content in the CTC for ECM. It was reported that this gives approximately 0.5% gain for AI.

It was suggested that, if such a change is done, it would be appropriate to enable IBC for the VTM anchor as well, to properly reflect the gain that is due to the modifications in ECM. However, the proposal JVET-Z0157 also modifies the search such that a larger local range is searched after the hash-based initial search. Proponents are asked to implement a corresponding modification in the VTM encoder and, if possible, report back results comparing ECM (including the adopted elements from EE2-3.6l) vs. VTM (both with IBC enabled) during the meeting. This would be relevant for the AI settings only.

This aspect was again discussed in session 18. It was reported that turning on IBC in VTM gives approximately 0.4% gain, which is comparable to the gain enabling it in ECM. From this point of view, turning on IBC in CTC for both VTM and ECM would not give additional improvement of ECM over VTM, while the encoder runtime would increase. No action was taken.

[JVET-Z0175](https://jvet-experts.org/doc_end_user/current_document.php?id=11625) AHG3: On merging the HM and VTM HDR CTCs [A. Segall, E. François, W. Husak, S. Iwamura, D. Rusanovskyy] [late]

AHG3 has the mandate to investigate merging the HM and VTM common test conditions. In support of that mandate, the starting point for merging the HDR HM and VTM common test conditions was updated that was originally provided in JVET-X0007. The starting point document is provided with this contribution.

This document defines four test conditions for HDR/WCG content, reflecting intra-only, random-access, and low-delay settings:

• Intra, 10 bit

• Random access, 10 bit

• Low delay, 10 bit

• Low delay, P slices only, 10 bit

A subset of these test conditions might be used for a particular experiment. For example, when testing an intra coding tool, only intra configurations might be used.

This is specifically defined for HM and VTM.

It was decided to convert this into output doc JVET-ZXXXX

## Verification testing (0)

Contributions in this area were discussed in session X at XXXX–XXXX UTC on XXday 2X April 2022 (chaired by JRO).

## Test material (3)

Contributions in this area were discussed in session X at XXXX–XXXX UTC on XXday 2X April 2022 (chaired by JRO).

[JVET-Z0053](https://jvet-experts.org/doc_end_user/current_document.php?id=11487) [AHG4] REV Result for AHG11/EE1 and AHG4 new test sequences [M. Wien (RWTH)]

Remote expert viewing tests were conducted before the 26th JVET meeting for the exploration activity on EE1/DNN-based coding tools and for studying new gaming-type test material proposed in JVET-Y0041. The tests were performed using mp4 files provided by the proponents with a total of 68 A-B-comparisons for EE1 and 14 comparisons for the gaming-type content. Sequences of 720p, 1080p, and 2160p resolution were evaluated for EE1. For the gaming-type content, 1080p and 2160p were evaluated. Calls for participation in the subjective viewing were issued on the JVET reflector. Overall, 18 experts volunteered in participating in the tests. The testing procedure followed the draft SC29/AG5 guidelines for remote experts viewing, providing an A/B comparison of sequences under test.

The number of participants was only 18 which was indicated as the lowest number to be accepted for running the tests. Due to the uncontrolled and variable viewing conditions the resulting data may be of limited reliability and should be treated with care.

Based on the reported MOS values, the following observations are made: For the EE1-2 experiment, a significant visual benefit is reported for 14 out of 24 test cases. For the remaining cases, MOS = 0 is included in the confidence interval such that comparable quality is indicated. In the EE1-1.7 experiment, 9 test cases out of indicate a significant visual benefit while one case showed a significant loss. The remaining 14 cases indicate comparable quality. In the super-resolution category (EE1-2.x), visual benefit is observed for three cases for the NN-based method (2.4) and one for the RPR-based approach. The other cases indicate no benefit by the confidence interval with two cases for the RPR-based approach which show a tendency towards loss.

It is recommended to determine a sufficiently large group of volunteers well before the planned test date in order to avoid difficulties in acquisition at test time. This may be done by the end of the current meeting or (latest) by the AhG meeting in preparation of the viewing tests. The test preparation would be completed if the requested number of participants is reached.

It was commented that the Darktree sequence has rapid motion which makes visual assessment challenging for viewers, esp. for the 5sec evaluation time used in the REV. It was further commented that the Fountainbleau sequences are easier to watch, with FontainebleauCinematicS showing a good mix of features that can be used to distinguish between different coding configurations. The Racing sequence was considered to be most easy to watch from the given set of sequences.

[JVET-Z0138](https://jvet-experts.org/doc_end_user/current_document.php?id=11587) AHG7: Update on gaming sequences [T. Poirier, G. Martin-Cocher, E. Faivre d'Arcier (InterDigital)]

This contribution proposes an update on new gaming sequences proposed in JVET-Y-0041 at the last meeting. It also proposes to create a new class of content for gaming sequences.

The proposed update provides the 3D-GameKit and the RacingCar sequences at improved quality.

It was suggested to study these new sequences w.r.t suitability for visual assessment. It was further suggested to develop criteria for the composition of a gaming-type CTC class. This should include the number of sequences to be included, the characteristics with respect to resolution, compression, motion, color range. It was noted that considering both, 4K and HD resolution, is necessary. The different resolution could be regarded e.g. by a suffix to the class character.

It was commented that other resolutions might be of interest as well (e.g. portrait mode). It was suggested to study the impact of such resolutions from a coding perspective. It was noted that visual assessment of such content might be done differently than for landscape mode. The question was raised if the orientation of the video would have an impact on the tool development process. It was suggested to further investigate on this topic.

It was suggested to develop a proposal for a gaming-type CTC class as input to the next meeting as part of an JVET AHG4 activity.

- 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 (e.g. availability of ground truth motion for the sequences proposed here).

[JVET-Z0156](https://jvet-experts.org/doc_end_user/current_document.php?id=11605) AHG4: New test sequences for JVET exploration [J. Chen, Y. Ye (Alibaba), R. Li, W. Jiang (Youku)]

This contribution proposes 2 new 4K test contents (4 sequences in total in 8- and 10-bit) for future video coding exploration. The proposed sequences are from TV drama content on Youku. The coding results of VTM-11.0-anc and ECM-4.0 on these sequences and the copyright information are also provided in this contribution.

It was commented that the sequences are very easy to watch and thereby may be considered suitable for subjective evaluated.

It was noted that the sequences may be relatively easy to code. For the CTC QP range, the rate range is reported to be roughly from 300 kbps to 5000 kbps with the PSNR ranging from about 40 to 50 dB for luma. It was noted that the lower bitrates partially be attributed to the frame rate of 25 Hz.

It was commented that it should be discussed what composition of content should be represented in the CTC.

It was commented that very low bitrate sequences may not trigger challenges in the application since the rate budget might not be challenged. It was noted that the proposed content has been taken from an actual application.

It is suggested to consider an update of the test set for the CTC since the sequences currently present have been used for a very long time and overfitting may occur.

It was asked to investigate on the types of artifacts that are observed in the range of moderate to high QP as this may trigger insight on potential improvements for coding tools. Further study on this was encouraged.

## Quality assessment (5)

Contributions in this area were discussed in a joint meeting with AG 5 and VCEG at 2100–2300 UTC on Wednesday 27 April 2022 (chaired by Mathias Wien, GJS and JRO).

[JVET-Z0045](https://jvet-experts.org/doc_end_user/current_document.php?id=11479) AhG11/AhG4/EE1 viewing preparation report [E. Alshina, M. Wien, A. Segall, J. Sauer]

This document reports results of discussion on preparation for viewing test to evaluate Neural Network-based Video Coding (NNVC) technologies under study in JVET.

No need for presentation – included in AHG report.

[JVET-Z0057](https://jvet-experts.org/doc_end_user/current_document.php?id=11492) On subjective evaluation of video quality with the crowdsourcing approach [B. Naderi (TU Berlin), R. Cutler (Microsoft)]

This contribution presents an open-sourced framework for conducting subjective video quality assessment tests using crowdsourcing. The platform is under study in ITU-T SG12/Q7 work item P.CrowdV. The aim of this contribution is to seek feedback from JVET experts and to suggest considering to use this framework for subjective testing in JVET work.

The authors were thanked for their contribution. It was pointed out that the implemented approaches for calibration testing, viewing distance, setup etc. are very interesting also in the context of the REV.

It is stated that the processing of the results is done automatically by the results parser. The parser is configurable.

It was noted by an expert that among a variety of testing platforms, this approach is most constructive in communication with the test subjects. It was reported that the framework was successfully applied locally without link to AMT.

It was indicated adaptation of the framework is possible and that modifications for the application with the REV.

It was asked how the participants would score during the session and how the data is transferred to the data base. The data is entered during the session, the session can be interrupted, also replay is possible.

It was noted that entering the data during the session gives the opportunity to measure reaction time which may help to analyze the action of the viewers and the complexity of the task.

It requested if test data from JVET tests could be used for a replication in the crowdsourcing scenario in order to study the method in this context. It was suggested to ask the proponents who produced the sequences for permission.

During the plenary on Friday April 29, 0630, this question was brought up again, and no complaints were raised against a usage of the material from tests for studies in AG 5.

[JVET-Z0107](https://jvet-experts.org/doc_end_user/current_document.php?id=11554) Quality evaluation of internal 10-bit versus 8-bit processing with 8-bit source content [J. Jung, X. Li, S. Li (Tencent)]

This contribution reports the quality evaluation of internal 10-bit processing versus 8-bit processing, considering 8-bit input content. A selection of JVET sequences, encoded with VTM 16.0rc1, is subjectively evaluated through a CCR viewing methodology, by expert viewers. The detailed analysis of the subjective scores reveals that the difference between the two approaches is more visible at low bitrate, for class E and for low delay configuration. The reported difference remains however very slight on average, suggesting that it is safe to use 10-bit internal bit depth processing, even with 8-bit source content.

It was noted that RP1 for the B-S05 sequence (BQTerrace) had some larger difference in bitrate which do not seem to translate to visual benefit. Some of the objective metrics captured a difference though. It was pointed out that RP1 corresponds to QP22. Since this sequence is known to have specific noise characteristics, it is suggested that the rate difference might be hardly visible in this very high bitrate range.

The contribution used the 7-grade scale for evaluation. It is suggested to consider using this scale in the context of the JVET REV as well in order to increase the discriminatory precision of the expert assessments. It was noted that supporting multiple scales in the REV guidelines is deemed very useful.

[JVET-Z0108](https://jvet-experts.org/doc_end_user/current_document.php?id=11555) Evaluation of 17 objective quality metrics on JVET contents [J. Jung, J. G. Lopez, X. Li, S. Liu (Tencent)]

Objective metrics are usually considered to correlate insufficiently with Mean Opinion Scores (MOS) to provide a reliable overall quality score. This contribution updates the contribution m58816. It evaluates 13 full-reference and 4 no-reference objective metrics using subjective scores obtained on JVET content, and during JVET remote expert viewing sessions. Instead of computing classical correlations (Pearson, Spearman, etc) with average MOS, and trying to assess how much the metric reflects the overall quality level, the contribution investigates the ability of the metrics to answer the question: “Is the proposal better than the anchor?”. A Receiver Operating Characteristics approach with a binary classifier is used. The confidence intervals and the bias resulting from the usage of a forced-choice scale are taken into account. It is reported that a group of metrics can be considered as reliable for most of the decisions, with correct decision rates above 85%.

It was asked if other metrics were considered (e.g. SSIM+, ..).

It was noted that the focus of the evaluation here was mostly on the clear cases. It was asked how the metrics perform for the more difficult cases. It was stated that the current results only rely on the available CMOS data. The data close to CMOS=0 has some uncertainty which makes the analysis difficult. It was suggested that the REV method should be further developed to enable improved capability of fine distinctions.

It was noted that the collective of viewers provides a stable result and that a similar statement might be made by a collective of metrics.

It was noted that the investigated metrics would focus on luma and that chroma effects needs to be regarded.

It was asked if the concept of CIs would be applicable to the objective metrics. It was stated that the dimensions of such CIs are not known or very difficult to achieve for the metrics.

It was asked if resolution dependent features of metrics would have been taken into account. It was stated that for VMAF, the 4K/1080p versions were applied for the respective content.

It was suggested to work on the design of the REV to improve the reliability of its output, also at the more complicated cases. It was stated that the 4-grade scale might be replaced by a different one to support this type of assessment. It was mentioned that the choice of the scale is depending of the question under investigation.

[JVET-Z0109](https://jvet-experts.org/doc_end_user/current_document.php?id=11556) Evaluation of objective quality metrics on HEVC, VVC and AV1 contents [J. Jung, J. G. Lopze, X. Li, S. Liu (Tencent)]

This contribution evaluates objective quality metrics on a recent database provided by Bristol University, with both HD and UHD content, encoded with recent encoders (HEVC, AV1, VVC) and with mean opinion scores (MOS) obtained from subjective experiments made in laboratory. The results are reported as typical correlations and average errors, compared to the MOS. Classical and learning-based metrics have been evaluated. It is reported that VMAF and AVQT are the best performing metrics, far above other usual metrics, especially the PSNR. More detailed results are provided, analyzing metrics’ efficiency according to the spatial resolution, to the encoder used, to the spatial and temporal information level, and to the bitrate range.

It is suggested to add additional metrics to the JVET template in order to consistently collect more information. It was suggested that contradicting results of multiple metrics should be taken as an indicator for the need of subjective assessment. It was argued that a suitable process for this aspect needs to be developed.

It was mentioned that the investigation might be strongly depending on the applied configuration of the specific encoder software packages and that different application settings may require different treatment.

It was mentioned that in the development w.r.t HDR, multiple metrics were used and this situation was not showing to be helpful at that point in time. Further, if steps are incremental (i.e., at low deltas), diverging indications may not help the process.

It was noted that the current state of standardization work might be the right time to work on the process and the metrics.

## Conformance test development (0)

Contributions in this area were discussed in session X at XXXX–XXXX UTC on XXday 2X April 2022 (chaired by JRO).

## Software development (1)

Contributions in this area were discussed in session X at XXXX–XXXX UTC on XXday 2X April 2022 (chaired by JRO).

[JVET-Z0161](https://jvet-experts.org/doc_end_user/current_document.php?id=11610) AhG11: SADL update [F. Galpin, T. Dumas, P. Bordes, E. François (InterDigital)]

Was reviewed in BoG JVET-Z0234

## Implementation studies and complexity analysis (1)

Contributions in this area were discussed in session 12 at 0540–0555 UTC on Tuesday 26 April 2022 (chaired by JRO).

[JVET-Z0150](https://jvet-experts.org/doc_end_user/current_document.php?id=11599) Memory usage report on VTM / ECM [T. Hashimoto, Y. Yasugi, T. Ikai (Sharp)]

This contribution proposes to support memory usage log output in VTM / ECM software. With the ongoing development of ECM, the encoder requires more and more memory. This could heavily reduce the number of parallelism in cluster machines with limited memory, e.g. 256 GB machine only run 14 simulations even if it may have 20 or more processing units (cores) and it would interfere with development and experimentation of ECM. From this point of view, it is asserted that developers should be aware of the amount of memory usage in development.

This contribution donates a print out code for VTM / ECM which supports Linux environment, i.e. the changes less than 50 lines.

In v2, results are updated.

Table 1: Maximum amount of memory in RA condition (GiB)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| RA | VTM16 | ECM1 | ECM2 | ECM3 | ECM4 |
| classA | 8.8 | 14.9 | 15.0 | 16.1 | 18.1 |
| classB | 2.8 | 6.1 | 6.2 | 7.0 | 8.1 |
| classC | 1.0 | 3.3 | 3.5 | 3.6 | 4.4 |
| classD | 0.6 | 2.8 | 2.9 | 3.0 | 3.7 |
| classE |  |  |  |  |  |
| classF | 2.6 | 5.7 | 5.9 | 6.8 | 7.8 |

The proposal also includes a macro “REUSE\_CU\_RESULTS” that reduces memory usage with marginal change in performance. It however increases the encoder run time.

It was pointed out that memory usage can also be retrieved using the “time” command of Linux, and that Linux clusters typically also report memory usage of a job.

It is generally deemed to be helpful allowing experts such an analysis (for information purposes).

Decision(SW): Integrate into ECM (not mandatory to use)

[JVET-Z0224](https://jvet-experts.org/doc_end_user/current_document.php?id=11675) Crosscheck of JVET-Z0150 (Memory usage report on VTM / ECM) [Y. Kidani, K. Kawamura (KDDI)] [late]

## AHG7: Low latency and constrained complexity (5)

Contributions in this area were discussed in sessions 12 and 13 at 0605–0715 UTC and 0740-0810 UTC on Tuesday 26 April 2022 (chaired by JRO).

[JVET-Z0110](https://jvet-experts.org/doc_end_user/current_document.php?id=11557) On low delay configuration [X. Li, G. Li, S. Liu (Tencent)]

In this contribution, the coding efficiency and coding speed by low delay B (LDB) and low delay P (LDP) are compared and reported:

• LDB is 6+% more efficient (luma) than LDP with 29% longer runtime under VTM CTC. When only 1 reference picture is used for LDB and LDP, the average luma efficiency and runtime gaps are reduced to 3.61% and 21%, respectively. If a coding tool (GPM) which cannot be applied in P picture is disabled, the efficiency and runtime gap are further reduced to 2.06% and 15%, respectively.

• LDB and LDP with multiple reference pictures are much more efficient than those with single reference picture. The luma efficiency gaps of 20.36% and 16.37% by LDB and LDP are reported for 4 reference pictures vs 1 reference picture.

• LDP with 4 reference pictures is averagely 10.54% (luma) better than LDB with 1 reference picture with 11% less runtime.

It was mentioned by one expert that LDB with two references is a setting that is actually used in existing low-delay encoders. This might be a better comparison versus LDP with 4 references. I case of LDB with 1 reference, gain only comes from GPM and BCW.

Contribution for information and discussion on this subject – no specific action suggested.

[JVET-Z0114](https://jvet-experts.org/doc_end_user/current_document.php?id=11561) AHG7: Low delay configuration for cloud gaming [S. Puri, P. Le Guyadec, K. Naser, G. Martin-Cocher, T. Poirier (InterDigital)] [late]

This contribution proposes a new low delay configuration (named LDB-1Ref-asymmetric) where the number of reference frames is restricted to 1 frame per list, and reference frame list is modified such that reference list 0 refers to the previous picture in the GOP and reference list 1 refers to the picture at the start of the GOP.

This configuration is proposed as a refinement of the LLCC baseline configuration proposed during the last meeting ([JVET-Y0043](https://jvet-experts.org/doc_end_user/current_document.php?id=11235)) where it was agreed to restrict the number of reference frames to 2 in LD-B configuration.

Following set of tests are performed on one intra period (1IP). Test 1 and 2 compares the LDB-2Ref and LDB-1Ref (named LDB-1Ref-symmetric) configuration with the proposed LDB-1Ref-asymmetric configuration for class F, TGM and Gaming. Test 3 and 4 compares the proposed LDB-1Ref-asymmetric with LDP-4Ref and LDB-2Ref respectively.

Test 1: (LDB-2Ref (anchor) vs LDB-1Ref-asymmetric (tested)) 0.71%, 0.85%, 1.06%, 94%, 103%  
Test 2: (LDB-1Ref (anchor) vs LDB-1Ref-asymmetric (tested)) -7.19%, -8.84%, -8.77%, 102%, 110%  
Test 3: (LDP-4Ref (anchor) vs LDB-1Ref-asymmetric (tested)) 1.95%, 1.50%, 1.53%, 110%, 103%  
Test 4: (LDP-2Ref (anchor) vs LDB-1Ref-asymmetric (tested)) -2.00%, -3.11%, -3.04%, 120%, 106%  
Results on the full frame are added to the additional results under section 4. Section 4 also provides results on performance of LDB-2Ref-asymmetric compared to LDB-2Ref for completion.

It is asserted that the proposed LDB-1Ref-asymmetric configuration provides a better operating point due reduction in both encoder time and motion search operations when compared to LDB-2Ref configuration.

It is proposed to keep the LDB-2Ref configuration for the class B to E and to use the proposed LDB-1Ref-asymmetric configuration for at least class F, TGM and the newly proposed Gaming Class of sequences.

LDB-1ref symmetric has the previous picture in both list0 and list1. This takes potential benefit from GPM and BCW. The first picture of the GOP which is the second reference has higher quality – kind of “golden frame”.

It is pointed out that 1-ref asymmetric is worse than symmetric in terms of error resilience (similar as 2-ref, where also two pictures have to be retransmitted).

Why is the decoding time for 1-ref asymmetric increasing? Might be unreliable measurement.

It is pointed out that in practical encoders using more reference pictures is not a problem nowadays.

Goal should be primarily finding better conditions for cloud gaming. Changing CTC for other content not relevant.

Due to its better performance, it is agreed that the 1-ref asymmetric LDB configuration shall replace the 2-ref LDB as default configuration in the AHG7 study, for the case of investigating performance in cloud gaming sequences.

[JVET-Z0116](https://jvet-experts.org/doc_end_user/current_document.php?id=11563) AHG7: Refined low latency and controlled complexity configuration for cloud gaming [S. Puri, P. Le Guyadec, K. Naser, G. Martin-Cocher, T. Poirier (InterDigital)] [late]

This contribution proposes a refined low latency and controlled complexity (LLCC) configuration for cloud gaming (JVET-Y0043) by disabling MTS for low latency and controlled complexity video applications such as cloud gaming.

In JVET-Y0043, it was agreed to modify LDB configuration to restrict the number of reference frames to 2. This contribution proposes to disable Multiple Transform Set (MTS) tool in LDB configuration for class F, TGM and Gaming.

Following results are reported by disabling both MTS and MTSImplicit for Class Gaming on top of VTM-15.0 for one intra period (1IP):

Test 1: LD-B 4Ref: 0.66%, 0.59%, 0.13%, 95%, 96%

Test 2: LD-B 2Ref: 0.69%, 0.47%, 0.21%, 94%, 95%

Following results are reported by disabling both MTS and MTSImplicit for Class TGM on top of VTM-15.0 for one intra period (1IP):

Test 3: LD-B 4Ref: -0.25%, -0.14%, 0.00%, 100%, 101%

Test 4: LD-B 2Ref: -0.08%, -0.14%, 0.00%, 100%, 100%

Results with MTSImplicit tool enabled are provided in the additional results section.

Gaming sequences are “class DMV” (the new sequences currently under investigation) and “SA4 gaming” (some external sequence data set).

For gaming sequences, some loss occurs by disabling MTS, and the reduction in encoder/decoder runtime is not large.

Results are not based on entire sequences, only one GOP.

It is pointed out that for a judgement which tools are hardware-friendly or not may be highly dependent on architecture and business scenario.

No agreement could be reached on disabling MTS in the default AHG7 configuration.

[JVET-Z0118](https://jvet-experts.org/doc_end_user/current_document.php?id=11565) AHG7: GDR Implementation for ECM 4.0 [S. Hong, L. Wang, K. Panusopone (Nokia)]

This contribution presents a GDR implementation for ECM 4.0 [1]. With this GDR implementation, there is no need to impose constraints on coding tools, which helps to reduce the code size for ECM significantly as compared to VTM-like GDR [3]. In addition, this new GDR implementation provides great flexibility in future tool development for ECM. Specifically, adding of additional new coding tools in ECM code will not affect part of GDR code.

Simulations were conducted under low-delay B configuration specified by LLCC AhG [4]. The partial results (excluding class B Sequence) demonstrate that the overall loss of the new GDR over anchor is only 3.68%

Would it also be running with other configurations than LDB with 2 reference pictures? Yes.

A cross-checker points out that the current implementation only supports vertical virtual boundary as the boundary between “clean” and “dirty” area. Consistency with JVET-Z0105 (a new modification of VB processing) needs to be considered, as JVET-Z0105 handles VB also horizontally.

It was noted that the VB processing of JVET-0118 is a subset of JVET-Z0105, using the same padding concept.

Decision: Adopt JVET-Z0118 (software, and potentially ECM description, as far as there are deviations from the VVC GDR)

[JVET-Z0241](https://jvet-experts.org/doc_end_user/current_document.php?id=11692) Cross check of JVET-Z0118 (AHG7: GDR implementation for ECM 4.0) [A. M. Kotra (Qualcomm)] [late]

[JVET-Z0245](https://jvet-experts.org/doc_end_user/current_document.php?id=11696) Cross check of JVET-Z0118 (AHG7: GDR implementation for ECM 4.0) [T. Poirier (Interdigital)] [late]

[JVET-Z0141](https://jvet-experts.org/doc_end_user/current_document.php?id=11590) AHG7: GDR in ECM-4.0 [T. Poirier, F. Aumont, G. Martin-Cocher (InterDigital)] [late]

Initial version rejected as “placeholder”.

This contribution presents results of GDR implementation for ECM4.0 [1]. GDR has been implemented in VTM [2] and then ported in ECM-3.0 [3. This GDR implementation has been tested with all new tools of ECM except ALF enabled and no leaks are observed.

It is proposed to add this GDR implementation to the ECM reference software.

Simulations were conducted under LDP configuration, the results shows 14.3% loss for class D compared to an anchor with 1 Intra picture every second.

In the version 3 of the contribution a combination with JVET-Z0105 [4] is proposed to handle the Virtual boundary processing. Hence ALF is enabled for this combination, the results shows 11.9% loss for class D.

Implementation only available for LDP.

JVET-Z0118 seems more mature currently. Further study recommended on JVET-Z0141.

[JVET-Z0242](https://jvet-experts.org/doc_end_user/current_document.php?id=11693) Cross check of JVET-Z0141 (AHG7: GDR in ECM-4.0) [[S. Hong (Nokia)](mailto:seungwook.hong@nokia.com)] [late]

## AHG10: Encoding algorithm optimization (5)

Contributions in this area were discussed in session 13 at 0810–0905 UTC on Tuesday 26 April 2022 (chaired by JRO).

Note that JVET-Z0065 (from EE1) could also be used as encoder optimization for RPR.

[JVET-Z0072](https://jvet-experts.org/doc_end_user/current_document.php?id=11507) AHG10/AHG12: Enhanced reference picture structures for ECM and VTM [K. Andersson, R. Sjöberg, R. Yu, L. Litwic (Ericsson)]

This contribution suggests a modified reference picture lists setting for ECM/VTM random access to store more pictures in the allowed Decoded Picture Buffer (DPB). The proposal also includes a change to the QP value for two pictures in the GOP. This is claimed to be almost identical to the settings proposed for VTM in JVET-Y0126 with the addition to order active entries in L0 and L1 such that pictures closer to the current picture in terms of POC are earlier in the lists.

This contribution also suggests a modified reference picture lists setting for ECM/VTM low-delay B configuration. Comparing to the current ECM/VTM LDB common test configuration, the proposal stores more pictures in the Decoded Picture Buffer (DPB) but maintains the same number of active reference pictures in L0 and L1. The proposal also suggests an adjustment of the QP offset settings for every second picture in the GOP.

The objective benefit in terms of BDR is reported as follows:

RA: VTM: -0.52% Y/-0.56%U/-0.55%V ECM: -0.43%Y/-0.68%U/-0.64%V

LDB: VTM: -1.63%Y/-3.23%U/-2.87%V ECM: -1.62%Y/-2.76%U/-2.94%V

We propose to update the CTC with the proposed changes.

In v2 accurate encoding and decoding numbers for VTM is added. Those indicate an increase in encoding time by 5% for RA and a decrease by 2% for LDB.

Gains by tendency more towards the lower resolution classes.

No impact on subjective quality, according to proponents which did some subjective inspection.

What is the impact on changing QP for each second picture in LDB? Around 0.5%, and also helps reducing encoding time.

Decision(CTC): Adopt JVET-Z0072

[JVET-Z0247](https://jvet-experts.org/doc_end_user/current_document.php?id=11698) Crosscheck of JVET-Z0072 (AHG10/AHG12: Enhanced reference picture structures for ECM and VTM) [Christian Helmrich, Christian Bartnik (Fraunhofer HHI)] [late]

[JVET-Z0099](https://jvet-experts.org/doc_end_user/current_document.php?id=11546) AHG10: Deblocking in RDO and beta offset minus 2 for VTM [K. Andersson, J. Enhorn, R. Sjöberg, J. Ström, L. Litwic (Ericsson)]

Deblocking can provide a significant subjective benefit for VVC. At the last meeting the deblocking parameters were changed to be aligned with ECM since they provided an objective benefit while keeping subjective quality. To further improve objective benefits with maintained subjective quality, this contribution proposes to enable deblocking in RDO in CTC for RA and LD in VTM. It is also proposed to allow for more deblocking in VTM than ECM by a less restrictive setting. The proponents claim that this provides an additional objective benefit in VTM together with deblocking in RDO. This is also claimed to be identical with the deblocking settings in the alternative configurations for VTM (high complexity setting).

The reported results from enabling deblocking in RDO for VTM are as follows:

RA: -0.51%Y, -0.87%U, -0.84%V

LDB: -0.28%Y, -0.90%U, -1.00%V

Enabling deblocking in RDO and with the less restrictive settings (tC offset set to 0 and beta offset set to -2) for VTM:

RA: -0.56%Y, -0.99%U, -0.89%V

LDB: -0.41%Y, -0.73%U, -0.96%V

The objective benefit is reported to result in 5% encoding time increase.

In v2 we also add AI results for VTM, but we do not think the gains are sufficient here:

DB in RDO AI: -0,26%Y/-0,76%U/-0,75%V

DB in RDO and less restrictive settings AI: -0,09%Y/-0,73%U/-0,98%V

In v2 we also include results for ECM-4.0:

RA: -0,37% Y/-0,75%U/-0,68%V

LDB: -0,23%Y/ -0,22%U/-0,78%V

DB-RDO gives less gain in ECM than in VTM. Cross-checker reports that modifying the beta offset does not provide additional gain in ECM (other than VTM).

This could indicate that other loop filter tools of ECM are competing with this encoder optimization option.

In both cases ECM/VTM, by tendency better performance for high resolution (classes A).

Run time increase in VTM RA roughly 5%, in ECM almost negligible.

In AI higher increase, but AI has less gain, not attractive.

Decision(CTC): Adopt JVET-Z0099, enable RDO-DBF for both VTM and ECM in RA, LDB, LDP, for VTM also change beta offset -2, and tc offset 0.

[JVET-Z0213](https://jvet-experts.org/doc_end_user/current_document.php?id=11664) Crosscheck of JVET-Z0099 (AHG10: Deblocking in RDO and beta offset minus 2 for VTM) [N. Hu (Qualcomm)] [late]

[JVET-Z0104](https://jvet-experts.org/doc_end_user/current_document.php?id=11551) AHG10: Report of temporal layer dependent deblocking filter setting for HM [H. Zhang, X. Li, S. Liu (Tencent)]

In response to the request in 25th JVET meeting, this document reports the objective quality of applying the same deblocking filter parameter setting of VTM-16 to HM. In VTM, the deblocking filter control parameters BetaOffset\_div2 and TcOffset\_div2 are adjusted based on temporal layer. The same strategy and parameter are also introduced to HM. It is reported that under CTC, compared with HM-16.24, the overall performance drop is observed as follows:

• AI: 0.49%/1.07%/1.59% with 100%EncT/101%DecT

• RA: 0.64%/0.19%/0.31% with 100%EncT/100%DecT

• LDB: 0.26%/-0.01%/0.11% with 100%EncT/100%DecT

• LDP: 1.61%/0.42%/0.78% with 100%EncT/99%DecT

Contribution for information – results indicate that the corresponding changes of DBF parameters are not beneficial in HM (unlike the findings made for VTM and ECM). This may be due to the additional loop filters (in particular, ALF) that are present in VTM and ECM, not in HM.

[JVET-Z0111](https://jvet-experts.org/doc_end_user/current_document.php?id=11558) Adaptively bypass affine ME in VTM [W. Kuang, X. Li, G. Li, S. Liu (Tencent)]

In this contribution, an encoder only method is proposed to adaptively bypass affine ME. Following results are reported

• Option A: 0.24%/0.15%/0.14% loss with 94% encoding time for RA, 0.25%/0.29%/0.11% loss with 91% encoding time LDB, 0.21%/0.26%/0.30% loss with 94% encoding time LDP

• Option B: 0.21%/0.14%/0/14% loss with 95% encoding time for RA, 0.19%/0.22%/-0.06% loss with 94% encoding time LDB, 0.13%/0.12%/0.05% loss with 97% encoding time LDP

Interesting approach for reducing encoder run time, option B more attractive.

It is suggested to include this as an option in VTM.

Decision(SW): Adopt JVET-Z0111 option B (optional, not CTC)

[JVET-Z0193](https://jvet-experts.org/doc_end_user/current_document.php?id=11643) Crosscheck of JVET-Z0111 (Adaptively bypass affine ME in VTM) [H.-J. Jhu, X. Xiu (Kwai)] [late]

[JVET-Z0209](https://jvet-experts.org/doc_end_user/current_document.php?id=11660) AHG10：Encoding algorithm optimization for HTM [J.-Y. Huo, X.-L. Zhou, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.), M. Li, Y. Liu (OPPO)] [late]

In HTM 16.3[1], in depth coding, the rendering-based synthesized view distortion is used to calculate the RD cost, bringing sufficient coding gain which is time-consuming. In this proposal, a fast Rate-Distortion Optimization (RDO) method for depth maps is proposed, which checks the RD cost during its calculation process. In HTM, given a coding mode, the RD cost is composed of several cumulative items. If the accumulated RD cost is equal to or exceeds the minimum RD cost of previously coded modes, it will not be necessary to continue the RD cost calculation for the mode. For RA configuration, compared with HTM16.3, -0.04% BD-rate performance is reported by calculating the synthesized view PSNR and total bitrate, with 88.5% EncT, 99.3% DecT, 99.9% RenT. For depth encoding time, 71.2% EncT is reported.

Reduction of overall encoding time for HTM by more than 10% appears valuable. Proponents should get in contact with the software coordinator Karsten Sühring to inspect the maturity of the code. In this context, several suggestions were made, and proponents already made efforts to modify their software. It is expected that it might be possible to finish the integration and release a new version of HTM until the next meeting.

Decision(SW): Adopt JVET-Z0209

## Profile/tier/level specification (1)

Contributions in this area were discussed in session 7 at 0605–0615 UTC on Friday 22 April 2022 (chaired by JRO).

JVET-Z0122 is also related.

[JVET-Z0060](https://jvet-experts.org/doc_end_user/current_document.php?id=11495) On new levels for HEVC [M. Ikeda, T. Suzuki (Sony)]

This contribution proposes to correct Max bit rate for level 7.1 and 7.2 of HEVC. Max CPB size and Max bit rate for level 6.3 and higher were added in JVET-Y1005. Max CPB size is increased according to increase of max luma sample rate. However, Max bit rate for level 7.1 is not increased. This contribution proposes to increase the value of Max bit rate for level 7.1 and 7.2 for high tier according to the increase of max luma sample rate.

Decision: Adopt JVET-Z0060 (increase max bit rate values for levels 7.1 and 7.2)

## 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 (0)

Contributions in this area were discussed in session X at XXXX–XXXX UTC on XXday 2X April 2022 (chaired by JRO).

## AHG11: Neural network-based video coding (25)

### Summary and BoG reports

See also contributions on NN-related HLS in section 6.1, on software implementation in section 4.8, and on subjective quality testing related to EE1 in section 4.6. Documents in this section were presented and discussed in session 2 at 0750–0935 UTC on Wednesday 20 April 2022 (chaired by JRO).

[JVET-Z0023](https://jvet-experts.org/doc_end_user/current_document.php?id=11616) 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-Y and JVET-Z meetings to evaluate **Neural Network-based Video Coding (**NNVC) technologies, analyze their performance, and analyze their complexity aspects. In NN-based filter category BD-rate gain over AhG11 anchor is **8.2...11.5%** (depending on complexity) is demonstrated. In Super Resolution category **1.1** % BD-rate average gain over AhG11 anchor can be achieved (among them 3% for 4K content).

1. **Introduction**

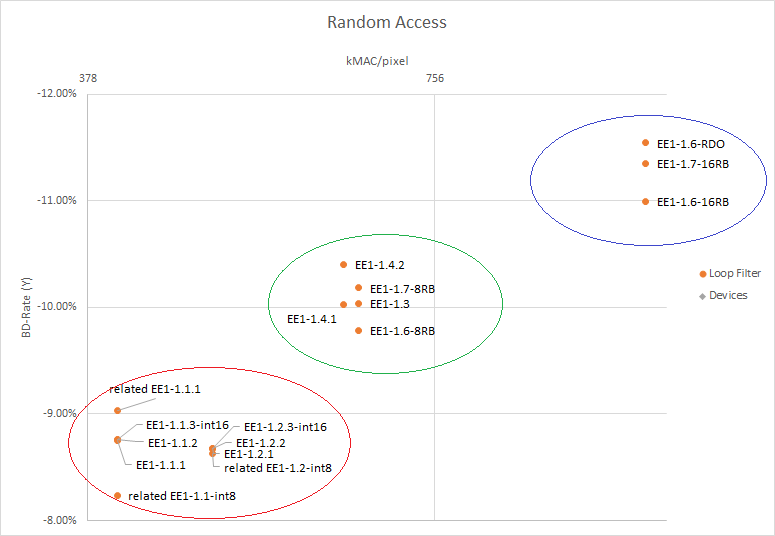
In total 7 tests studying NN-based filter technologies and 4 tests for Super Resolution technologies have been conducted. Several improvements as well as simplifications on top of technologies studied in EE1 have be proposed in EE1-related contributions. An effect of NN parameters quantizing to fixed point 16 and even 8 bits precision was studied.

Viewing was prepared and conducted together with SC 29 AG 5 for the proposals with highest BD-rate gain over AhG11 anchor.

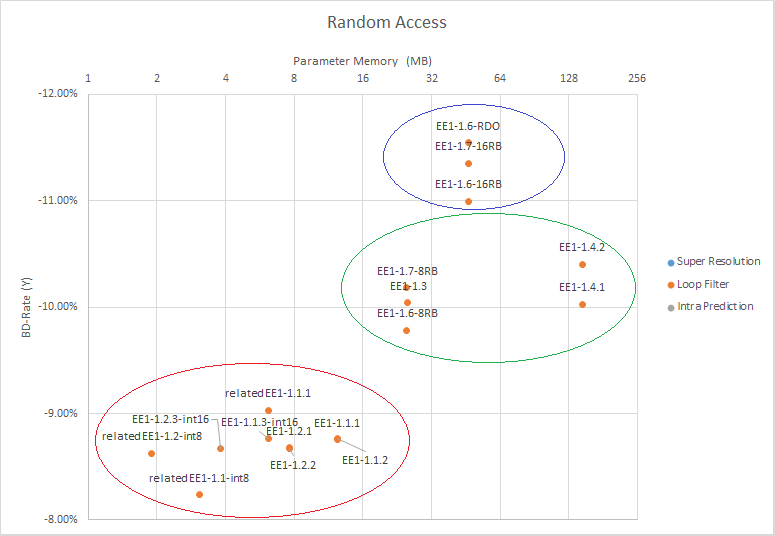
Anchor and test conditions in EE1 remains un-changed for the long time. Ahg11 conducts testing for 5 QPs (22, 27, 32, 37, 42) [1]. Common reference is VTM11.0 with improved GOP-based temporal filter [2]. SADL [3] implementation for two major NN-based filters architectures have been tested in EE1.

1. **NN-based filters category**

Three “clusters” NN-filters providing BD-rate saving relatively to AhG11 anchor can be identified. They are marked by red, green and blue circles on Fig. 1 and 2.



*Fig. 1*. BD-rate relatively to anchor in RA test for some NN-based filter depending on kMAC/pxl.



*Fig. 2*. BD-rate relatively to anchor in RA test for some NN-based filter depending on total memory size for all NN parameters.

Test EE1-1.2 with NN parameters quantized to Int8 likely the least complex among variants depicted on Fig. 1-2. But still computation complexity in kMAC/pxl is higher than 400 (while 64 is reference number for RTX 3080, 4K@60).

Three major NN-based filters architectures have been used in EE1 test.

**NN-based filter architecture originated by Bytedance and Qualcomm.**

The most popular among them was proposed by Bytedance [4] in JVET-X0065 (Fig. 3) enforced by residual scaling and some other elements by Qualcomm [5] JVET-X0140, later improved by Ericsson [7] and finally InterDigital provided SALD [3] implementation for inference.

In this filter design Y is processed separately, U and V components re processed jointly.



*Table 1*. *Performance results for NN-filters based on architectures from Bytedance and Qualcomm.*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | Random Access (CTC) | | | All Intra (CTC) | | | Inference |
| Contribution | Source | Total Number of Parameters (Millions) | Worst Case Complexity (kMAC  /pixel) | Y | Cb | Cr | Y | Cb | Cr |  |
| Split ALF in combination with JVET-X0065 | | | | | | | | | | |
| EE1-1.4.1 | [JVET-Z0086](https://jvet-experts.org/doc_end_user/current_document.php?id=11521) | 36.7 | 630.0 | -10.0% | -23.9% | -23.5% | -7.4% | -20.7% | -20.9% | PyTorch v1.6 |
| EE1-1.4.2 | [JVET-Z0086](https://jvet-experts.org/doc_end_user/current_document.php?id=11521) | 36.7 | 630.0 | -10.4% | -13.7% | -13.3% | -8.7% | -10.9% | -11.3% | PyTorch v1.6 |
| Solution for blocking artifacts ( JVET-Y0143) | | | | | | | | | | |
| EE1-1.3 | [JVET-Z0070](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0070-v1.zip) | 6.3 | 649.0 | -10.0% | -21.6% | -21.4% | -7.3% | -19.8% | -20.1% | PyTorch v1.6 |
| EE1-1.6-8RB | [JVET-Z0112](https://jvet-experts.org/doc_end_user/current_document.php?id=11559) | **6.2** | 649.0 | -9.8% | -21.2% | -21.1% | -7.4% | -19.3% | -19.4% | *SADL* |
| EE1-1.6-16RB | [JVET-Z0112](https://jvet-experts.org/doc_end_user/current_document.php?id=11559) | **11.6** | 1151.0 | -11.0% | -23.0% | -24.5% | -8.4% | -21.5% | -22.5% | SADL |
| *SADL implementation* | | | | | | | | | | |
| EE1-1.7-8RB | JVET-Z0113 | **6.2** | 649.0 | -10.2% | -22.2% | -22.1% | -7.2% | -19.7% | -20.0% | SADL |
| EE1-1.7-16RB | JVET-Z0113 | **11.6** | 1151.0 | -11.4% | -23.9% | -25.4% | -8.2% | -22.2% | -23.3% | SADL |
| Advanced RDO for filter selection | | | | | | | | | | |
| EE1-1.6-RDO | [JVET-Z0154](https://jvet-experts.org/doc_end_user/current_document.php?id=11603) | **11.6** | 1151.0 | -11.5% | -23.1% | -24.6% | -8.9% | -21.6% | -22.7% | SADL |
| EE1-1.7-RDO | [JVET-Z0155](https://jvet-experts.org/doc_end_user/current_document.php?id=11604) | **11.6** | 1151.0 |  |  |  | -8.5% | -22.4% | -23.5% | *SADL* |
| Ablations study | | | | | | | | | | |
| EE1-1.6-8RB-retrain | [JVET-Z0106](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0106-v1.zip) | **6.2** | 649.0 |  |  |  | -7.4% | -18.6% | -19.1% | PyTorch v1.9 |
| EE1-1.6-8RB-retrain-long | [JVET-Z0106](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0106-v1.zip) | **6.2** | 649.0 |  |  |  | -7.6% | -18.5% | -19.1% | PyTorch v1.9 |
| EE1-1.6-8RB-w/o part | [JVET-Z0106](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0106-v1.zip) | **6.2** | 638.0 |  |  |  | -7.6% | -18.4% | -19.3% | PyTorch v1.9 |
| EE1-1.6-8RB-w/o BS | [JVET-Z0106](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0106-v1.zip) | **6.2** | 638.0 |  |  |  | -7.6% | -18.4% | -19.0% | PyTorch v1.9 |
| EE1-1.6-8RB-w/o part &BS | [JVET-Z0106](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0106-v1.zip) | **6.2** | 628.0 |  |  |  | -7.4% | -18.6% | -19.2% | PyTorch v1.9 |
| EE1-1.6-8RB-w/o pred | [JVET-Z0106](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0106-v1.zip) | **6.2** | 638.0 |  |  |  | -7.3% | -18.5% | -19.1% | PyTorch v1.9 |
| EE1-1.6-4RB | [JVET-Z0106](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0106-v1.zip) | **5.6** | 481.0 |  |  |  | -6.8% | -19.3% | -19.6% | PyTorch v1.9 |
| EE1-1.6-16RB | [JVET-Z0106](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0106-v1.zip) | **7.6** | 985.0 |  |  |  | -8.3% | -17.7% | -18.5% | PyTorch v1.9 |
| EE1-1.6-8RB att.last | [JVET-Z0106](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0106-v1.zip) | **6.2** | 646.0 |  |  |  | -7.6% | -18.7% | -19.6% | PyTorch v1.9 |
| EE1-1.6-8RB w/o att. | [JVET-Z0106](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0106-v1.zip) | **6.2** | 646.0 |  |  |  | -7.6% | -18.6% | -19.2% | PyTorch v1.9 |

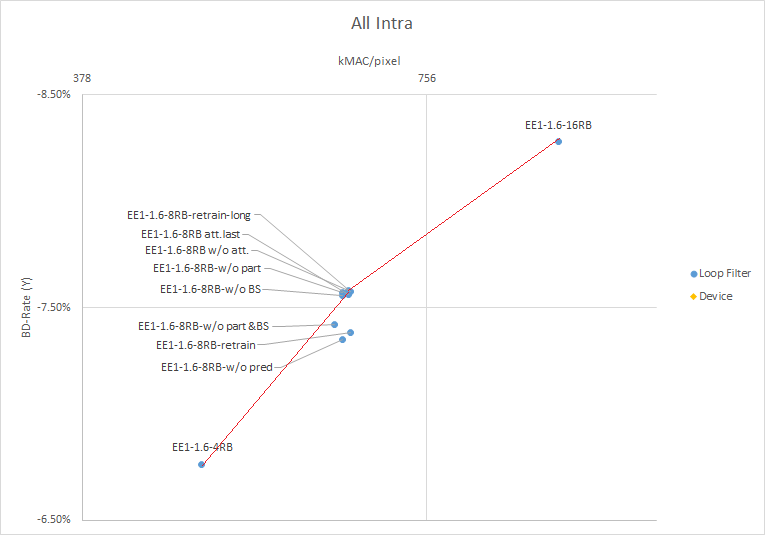
In EE1 related contribution ablation study for this NN-based filter architecture as performed. Ericsson was able to re-train model from the scratch, and so verify training procedure. As expected after re-training performance is slightly different, but difference is 0.0% BD-rate Y and up to 0.7% BD-rate Chroma.

Another useful finding from Ericson is longer training (378 epochs instead 155) improves performance by 0.2% BD-rate (in all intra cfg).

Results of ablation study by Ericsson (JVET-Z0106) are visualized on Fig. 4. All test results (except just retrain) corresponds long training.

The change for number of residual blocks (default is 8, 4 and 16 have been tested) has major effect on performance and complexity. Red line of Fig. 4 connects RB4, RB8 and RB16 variants of NN filter. The curvature of this curve shows that number of residual blocks 8 is optimal performance/complexity trade-off.

Removal from design prediction information provides the biggest performance drop. Removal of BS or partitioning information, as well as disabling attention mechanism provides almost not performance / complexity change. Even removal both partitioning and BS from design is still close to the optimal, but reduces tensors size in the memory.



*Fig. 4. Ablation study by Ericsson (JVET-Z0106).*

It should be noticed that ablation test have been provide in all-intra cfg only. In Inter configuration an effect of various elements of NN-filter design might be very different. Such kind of study can be recommended.

1. **NN-based filter architecture originated by Tencent.**

In this filter design Y, U and V components are processed jointly. Two variants tested in EE1 are sown on Fig. 5. They were proposed by Tencent in

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

*Fig. 5*. *Basic NN-filter architecture by Tencent: with two models EE1-1.1 (left) and with single model models EE1-1.2 (right).*

Performance results for both variants is very close. The number of parameters in single model variant is ×1.5 smaller. Smaller number of parameters is compensated by extra input (Slice type). It is reported that the training time ~150 hours for both variants.

Test results in format recommended by AhG11 are summarized in Table 2. Implementation change from Torch to SALD w/o change for NN parameters precision (32 bits) is nearly lossless. Encoding and decoding run-time surprisingly higher SADL variant: Enc. ×1.2 of anchor (libtorch) 🡪 ×2.0 of anchor (SADL), Dec. ×300 of anchor (libtorch) 🡪 ×1700 of anchor (SADL).

*Table 2*. *Performance results for NN-filters based on architectures from Tencent.*

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | |  | Random Access (CTC) | | | All Intra (CTC) | | | Notes |
| Tets | Bits | Total Number of Parameters (Milli) | | Worst Case Complexity (kMAC /pixel) | Y | Cb | Cr | Y | Cb | Cr |  |
| EE1-1.1.1 | F32 | **3.1** | 401.0 | | -8.8% | -19.9% | -19.4% | -7.4% | -16.8% | -17.3% | Torch |
| EE1-1.1.2 | F32 | **3.1** | 401.0 | | -8.8% | -19.9% | -19.4% | -7.4% | -16.8% | -17.3% | *SADL* |
| EE1-1.1.3-int16 | Int16 | **3.1** | 401.0 | | -8.8% | -19.9% | -19.4% | -7.4% | -16.8% | -17.3% | *SADL* |
| EE1-1.1-int8 | **Int8** | **3.1** | 401.0 | | -8.2% | -19.0% | -18.1% | -6.3% | -15.0% | -16.1% | *SADL* |
| Related  EE1-1.1.1 | F 32 | 1.5 | 401.0 | | -9.0% | -20.0% | -19.5% |  |  |  |  |
| EE1-1.2.1 | F 32 | 1.9 | 485.0 | | -8.7% | -18.6% | -19.0% | -6.5% | -14.9% | -16.0% | Torch |
| EE1-1.2.2 | F 32 | 1.9 | 485.0 | | -8.7% | -18.6% | -19.0% | -6.5% | -14.9% | -16.0% | *SADL* |
| EE1-1.2.3-int16 | Int16 | 1.9 | 485.0 | | -8.7% | -18.7% | -19.0% | -6.5% | -14.9% | -16.0% | *SADL* |
| EE1-1.2-int8 | **Int 8** | 1.9 | 485.0 | | -8.6% | -18.1% | -18.4% | -6.3% | -14.4% | -15.4% | *SADL* |

Quantizing NN parameters to 16 bits comes almost w/o performance change. With quantized to 8 bits precision parameters NN filter performance slightly degrades (0.1~0.2% BD-rate) for variant with single model (EE1-1.2). For the variant with 2 models degradation due to NN parameter quantization to 8 bits is higher (~0.6% BD-rate).

It is expected that additional speed up is possible if not only NN weights but also tensor values after activation will be quantized to limited (16, or even better 8 bits) precision.

**CNN post-processing filter based on depthwise separable convolution and attention mechanism.**

Filter design in this test used WCDANN architecture (Fig. 6). This design was proposed in [1] by OPPO and Xidian Uni.





1. **Super Resolution category**

In this round of EE eventually apple-to-apple comparison for super-resolution with RPR filer up-sampling and NN-based super-resolution becomes possible. Tables 3 and 4 summarized performance test results and complexity.

It tests EE1-2.1.1 LGE combines existing in VVC standard RPR re-sampling filter with adaptive selection of coded picture size [11].

It tests EE1-2.1.2 additionally Chroma QP is reduced in order to balance Luma / Chroma performance. Still some average gain (0.4%) relatively to AhG11 anchor is preserved in RA configuration. In all intra after Chroma QP adjustment there is no gain in average.

*Table 3*. *Performance results for test in Super Resolution category. Average BD-rate (all classes) is reported.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Random Access (CTC) | | | All Intra (CTC) | | |
| Contribution | Total Number of Parameters (Millions) | Worst Case Complexity (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) | 0.0 | 0.0 | **-1.0%** | 2.6% | 2.5% | **-0.6%** | 3.1% | 2.7% | ***RPR filter & adaptive size*** |
| EE1-2.1.2 | 0.0 | 0.0 | -0.4% | -1.4% | -1.4% | 0.0% | 0.5% | 0.0% | *RPR filter & adaptive size* |
| EE1-2.2 | 2.4 | 244.0 | 52.7% | 11.1% | 11.6% | 14.8% | 1.5% | -4.7% | *NN-filter1 w/****o*** *adaptive size* |
| EE1-2.3 | 7.0 | 854.0 | 51.8% | -7.6% | -10.6% | 12.4% | -12.5% | -20.2% | *NN-filter1 w/****o*** *adaptive size* |
| EE1-2.4 | 7.0 | 854.0 | **-1.3%** | -0.6% | -0.8% | **-1.1%** | -0.4% | -0.7% | ***NN-filter2 & adaptive size*** |

Two variants of NN-based filters tested in this EE1 round. Both are initiated by Bytedance [12], [13]. Luma is up-sampled separately. Chorma up-sampling NN receives Luma as extra input.



(a)



(b)

*Fig. 7* (a) Architecture of Luma (a) and Chroma (b) up-sampling NN.

In test EE1-2.3 U and V use dedicated network, while in test EE1-2.2 same network is used for U and V. This explains difference in complexity.

If NN-based up-sampling filter from EE1-2.3 replaces RPR filter in test EE1-2.1 (adaptive coded picture size). Then extra 0.3% gain in RA cfg amf 0.5% in all-intra cfg can be obtained. It shall be noticed that adaptive coded picture size selection is more beneficial for high resolution content. Table 4 shows results for only 4K classes (A1 and A2).

*Table 4.* *Performance results for test in Super Resolution category. Average BD-rate (only 4K).*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Random Access (CTC) | | | All Intra (CTC) | | |
| Contribution | Total Number of Parameters (Millions) | Worst Case Complexity (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) | 0.0 | 0.0 | **-2.5%** | 5.0% | 5.2% | **-2.2%** | 7.5% | 6.1% | ***RPR filter & adaptive size*** |
| EE1-2.1.2 | 0.0 | 0.0 | -1.3% | -2.9% | -2.3% | -0.9% | 2.0% | 0.8% | *RPR filter & adaptive size* |
| EE1-2.2 | 2.4 | 244.0 | -5.9% | 18.5% | 7.8% | -8.1% | 33.2% | 0.8% | *NN-filter1 w/****o*** *adaptive size* |
| EE1-2.3 | 7.0 | 854.0 | -6.5% | 0.3% | -12.7% | -9.2% | 13.8% | -13.8% | *NN-filter1 w/****o*** *adaptive size* |
| EE1-2.4 | 7.0 | 854.0 | **-3.0%** | -1.4% | -1.7% | **-2.7%** | -0.8% | -1.5% | ***NN-filter2 & adaptive size*** |

Replacement of RPR filter by NN-based one gives more or less same performance improvement. Average gain achieved for 4K content is almost 3%.

NN based approaches are using RPR downsampling.

For 4K, upsampling is optimum for almost all cases, therefore no adaptation needed. Here, the NN1 shows better performance than NN2 (the latter close to RPR)

1. **Visual test**

The best performing NN-based filters (tests EE1-1.2 and EE1-1.7) were selected for viewing. Also super resolution with RPR filter (EE1-2.1) and NN based filters (EE1.2.4) went to the viewing. Viewing results are discussed in JVET-Z0053. It appears that blocking artifacts problem of EE1.7 still not resolved.

1. **Cross-checks**

During preparing EE1 summary some complexity numbers were corrected by cross-checkers. NN-based filters performance was verified by OPPO (cross-checked tests EE1-1.1, 1.2, 1.6 and 1.7) and Bytedance (cross-checked test EE1-1.4).

1. **Recommendations**

Continue EE study, considering ablation study. Pay more attention during cross-check not only for verification performance results but also complexity numbers. Study inference implementation for limited precision for NN parameters and for activations to provide more possibilities for optimization. Revise summary template [14] used for NN-based proposals: 1) add encoder/decoder run-time information, 2) remove only 4K results.

BoG (E. Alshina) for summary template revision and double-check performance/complexity numbers for non-EE1 proposals, also discuss experiences with SADL from EE, and how to improve its usage.

SADL has been used in many parts, gives better insight into impact of architectures and training, and also quantized parameters. Has significantly larger runtime than Pytorch, though. Should be further clarified with contributors to check if the settings were correct e.g. for the integer conversion.

Related contribution (Z0106) investigates impact of training and performs ablation study, useful to define next EE.

[JVET-Z0053](https://jvet-experts.org/doc_end_user/current_document.php?id=11487) [AHG4] REV Result for AHG11/EE1 and AHG4 new test sequences [M. Wien (RWTH)]

Remote expert viewing tests were conducted before the 26th JVET meeting for the exploration activity on EE1/DNN-based coding tools and for studying new gaming-type test material proposed in JVET-Y0041. The tests were performed using mp4 files provided by the proponents with a total of 68 A-B-comparisons for EE1 and 14 comparisons for the gaming-type content. Sequences of 720p, 1080p, and 2160p resolution were evaluated for EE1. For the gaming-type content, 1080p and 2160p were evaluated. Calls for participation in the subjective viewing were issued on the JVET reflector. Overall, 18 experts volunteered in participating in the tests. The testing procedure followed the draft SC29/AG5 guidelines for remote experts viewing, providing an A/B comparison of sequences under test.

The number of participants was only 18 which was indicated as the lowest number to be accepted for running the tests. Due to the uncontrolled and variable viewing conditions the resulting data may be of limited reliability and should be treated with care.

Based on the reported MOS values, the following observations are made: For the EE1-2 experiment, a significant visual benefit is reported for 14 out of 24 test cases. For the remaining cases, MOS = 0 is included in the confidence interval such that comparable quality is indicated. In the EE1-1.7 experiment, 9 test cases out of indicate a significant visual benefit while one case showed a significant loss. The remaining 14 cases indicate comparable quality. In the super-resolution category (EE1-2.x), visual benefit is observed for three cases for the NN-based method (2.4) and one for the RPR-based approach. The other cases indicate no benefit by the confidence interval with two cases for the RPR-based approach which show a tendency towards loss.

It is recommended to determine a sufficiently large group of volunteers well before the planned test date in order to avoid difficulties in acquisition at test time. This may be done by the end of the current meeting or (latest) by the AhG meeting in preparation of the viewing tests. The test preparation would be completed if the requested number of participants is reached.

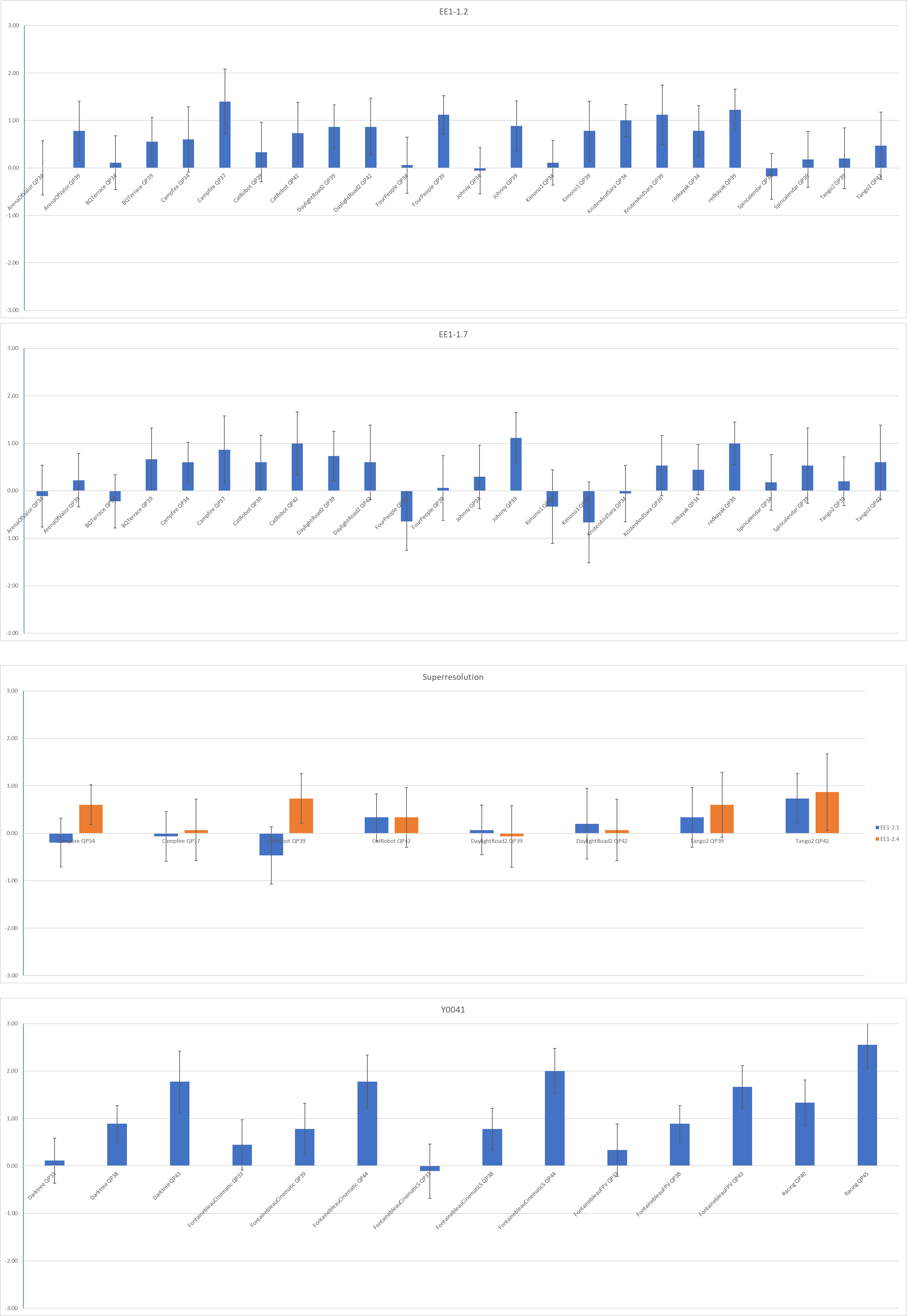
The MOS values and the corresponding confidence intervals (CI) resulting from the evaluation are reported in Table 4. According to the assignment of voting scores, a positive number indicates a visual benefit for the proposal over the anchor while a negative number indicates the proposal to be inferior to the anchor.

The tests are marked in the table. The ‘check’ test represents a test cell where the assessment of the original against a coded version or the assessment of high and low QP for the same source have been evaluated. The existence or location of these cells were not revealed to the participants before the experiments.

The scores of the 18 participants have not been further processed with the following exemption: The votes of participants voting the original or the lower QP inferior to the higher QP in the ‘check’ test were not regarded in the evaluation of the corresponding session.

**Table 4: MOS results for the A/B comparisons under test**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Test** | **Name** | **MOS** | **CI** |  | **Test** | **Name** | **MOS** | **CI** |  | **Test** | **Name** | **MOS** | **CI** |
| EE1-1.2 | ArenaOfValor QP34 | 0.00 | 0.57 |  | EE1-1.7 | ArenaOfValor QP34 | -0.11 | 0.65 |  | Y0041 | Darktree QP33 | 0.11 | 0.47 |
| EE1-1.2 | ArenaOfValor QP39 | 0.78 | 0.62 |  | EE1-1.7 | ArenaOfValor QP39 | 0.22 | 0.56 |  | Y0041 | Darktree QP38 | 0.89 | 0.38 |
| EE1-1.2 | BQTerrace QP34 | 0.11 | 0.57 |  | EE1-1.7 | BQTerrace QP34 | -0.22 | 0.56 |  | Y0041 | Darktree QP43 | 1.78 | 0.64 |
| EE1-1.2 | BQTerrace QP39 | 0.56 | 0.51 |  | EE1-1.7 | BQTerrace QP39 | 0.67 | 0.65 |  | Y0041 | FontainebleauCinematic QP33 | 0.44 | 0.53 |
| EE1-1.2 | Campfire QP34 | 0.60 | 0.68 |  | EE1-1.7 | Campfire QP34 | 0.60 | 0.42 |  | Y0041 | FontainebleauCinematic QP39 | 0.78 | 0.54 |
| EE1-1.2 | Campfire QP37 | 1.40 | 0.68 |  | EE1-1.7 | Campfire QP37 | 0.87 | 0.71 |  | Y0041 | FontainebleauCinematic QP44 | 1.78 | 0.56 |
| EE1-1.2 | CatRobot QP39 | 0.33 | 0.62 |  | EE1-1.7 | CatRobot QP39 | 0.60 | 0.57 |  | Y0041 | FontainebleauCinematicS QP33 | -0.11 | 0.57 |
| EE1-1.2 | CatRobot QP42 | 0.73 | 0.65 |  | EE1-1.7 | CatRobot QP42 | 1.00 | 0.66 |  | Y0041 | FontainebleauCinematicS QP38 | 0.78 | 0.44 |
| EE1-1.2 | DaylightRoad2 QP39 | 0.87 | 0.46 |  | EE1-1.7 | DaylightRoad2 QP39 | 0.73 | 0.52 |  | Y0041 | FontainebleauCinematicS QP44 | 2.00 | 0.48 |
| EE1-1.2 | DaylightRoad2 QP42 | 0.87 | 0.60 |  | EE1-1.7 | DaylightRoad2 QP42 | 0.60 | 0.78 |  | Y0041 | FontainebleauFPV QP33 | 0.33 | 0.55 |
| EE1-1.2 | FourPeople QP34 | 0.06 | 0.59 |  | EE1-1.7 | FourPeople QP34 | -0.65 | 0.60 |  | Y0041 | FontainebleauFPV QP38 | 0.89 | 0.38 |
| EE1-1.2 | FourPeople QP39 | 1.12 | 0.41 |  | EE1-1.7 | FourPeople QP39 | 0.06 | 0.68 |  | Y0041 | FontainebleauFPV QP43 | 1.67 | 0.45 |
| EE1-1.2 | Johnny QP34 | -0.06 | 0.49 |  | EE1-1.7 | Johnny QP34 | 0.29 | 0.67 |  | Y0041 | Racing QP40 | 1.33 | 0.48 |
| EE1-1.2 | Johnny QP39 | 0.88 | 0.53 |  | EE1-1.7 | Johnny QP39 | 1.12 | 0.53 |  | Y0041 | Racing QP45 | 2.56 | 0.51 |
| EE1-1.2 | Kimono1 QP34 | 0.11 | 0.47 |  | EE1-1.7 | Kimono1 QP34 | -0.33 | 0.78 |  |  |  |  |  |
| EE1-1.2 | Kimono1 QP39 | 0.78 | 0.62 |  | EE1-1.7 | Kimono1 QP39 | -0.67 | 0.85 |  | Check | CatRobot QP39 | 1.13 | 0.26 |
| EE1-1.2 | KristenAndSara QP34 | 1.00 | 0.34 |  | EE1-1.7 | KristenAndSara QP34 | -0.06 | 0.59 |  | Check | FontainebleauFPV QP32 | 1.33 | 0.35 |
| EE1-1.2 | KristenAndSara QP39 | 1.12 | 0.63 |  | EE1-1.7 | KristenAndSara QP39 | 0.53 | 0.63 |  | Check | Kimono1 QP39 | 2.56 | 0.40 |
| EE1-1.2 | redkayak QP34 | 0.78 | 0.54 |  | EE1-1.7 | redkayak QP34 | 0.44 | 0.53 |  | Check | KristenAndSara QP34 | 3.00 | 0.00 |
| EE1-1.2 | redkayak QP39 | 1.22 | 0.44 |  | EE1-1.7 | redkayak QP39 | 1.00 | 0.45 |  | Check | Racing QP37 | 1.56 | 0.43 |
| EE1-1.2 | Spincalendar QP34 | -0.18 | 0.48 |  | EE1-1.7 | Spincalendar QP34 | 0.18 | 0.59 |  | Check | Tango2 QP42 | 1.40 | 0.42 |
| EE1-1.2 | Spincalendar QP39 | 0.18 | 0.59 |  | EE1-1.7 | Spincalendar QP39 | 0.53 | 0.79 |  |  |  |  |  |
| EE1-1.2 | Tango2 QP39 | 0.20 | 0.64 |  | EE1-1.7 | Tango2 QP39 | 0.20 | 0.51 |  |  |  |  |  |
| EE1-1.2 | Tango2 QP42 | 0.47 | 0.71 |  | EE1-1.7 | Tango2 QP42 | 0.60 | 0.78 |  |  |  |  |  |



**Figure 2: MOS plots**

In the above figure, the results for the different test categories are aggregated. The different experiments are marked in the figure. It is noted that in the super-resolution category, EE1-2.1 refers to the VTM RPR-based approach while EE1-2.4 refers to NN-based super-resolution.

Some participants raised the fact that sequence pairs might look very similar. The overall duration of the experiment was mentioned. The following specific comments were recorded.

DNN:

* (720p) FourPeople and Spin Calendar quality is too low, better to use QP-1 next time.
* (720p) NN LF looks a bit blurry and anchor is more sharp. NN LF sometimes give artificial edges.
* (1080p) BQTerrace quality too high.
* (1080p) NN LF can give artificial edges and colors.
* (2160p) Tango quality too high.

Probably experts who gave specific comments on NN properties knew the characteristics of such technology and might have been able to identify A/B.

The number of participants was only 18 which was indicated as the lowest number to be acceptable for running the tests. Due to the uncontrolled and variable viewing conditions the resulting data may be of limited reliability and should be treated with care (also as some participants might have been biased).

Based on the reported MOS values, the following observations are made:

For the EE1-2 experiment, a significant visual benefit is reported for 14 out of 24 test cases. For the remaining cases, MOS=0 is included in the confidence interval such that comparable quality is indicated.

In the EE1-1.7 experiment, 9 test cases out of indicate a significant visual benefit while one case showed a significant loss. The remaining 14 cases indicate comparable quality.

In the super-resolution category (EE1-2.x), visual benefit is observed for three cases for the NN-based method (2.4) and one for the RPR-based approach. The other cases indicate no benefit by the confidence interval with two cases for the RPR-based approach which show a tendency towards loss.

It is recommended to determine a sufficiently large group of volunteers well before the planned test date in order to avoid difficulties in acquisition at test time. This may be done by the end of the current meeting or (latest) by the AhG meeting in preparation of the viewing tests. The test preparation would be completed if the requested number of participants is reached.

Results for NN based loop filter show more clear tendency of improvements, and also higher number of test cases than last time (though results be considered with care, see above). For NN based SR, only one case with non-overlapping CI improvement over RPR.

It is asked which amount of cases was actually coded with lower resolution in the SR case? Same decision used for RPR and NN, and a participant confirms that not all GOPs were coded with low resolution. Should be further investigated if that is the reason that in some cases not benefit was found, neither for RPR than for NN. .Such an analysis was reported in session 18. For example, in daylight road and campfire, full resolution coding is used in almost all cases. This matches the visual results, where no benefit was shown. The corresponding slide deck should be provided as a new version of JVET-Z0098.

It was commented that also comparison versus results from last meeting (in case of same sequences used) would be beneficial, to see if there are improvements, or known issues with visual quality had been resolved.

[JVET-Z0234](https://jvet-experts.org/doc_end_user/current_document.php?id=11685) BoG on Neural Networks Video Coding Results Analysis and further planning of EE1 [E. Alshina]

This is a report of activities from the BoG Neural Networks Video Coding Results Analysis and further planning of EE1. The BoG held meeting with about 120 participants during the 26th JVET meeting on April 22, 05:00-7:20 UTC

The BoG was established with the following mandates:

* Verify summary if NNVC (EE1, EE1-related and AhG11 contributions) results analysis
* Update NNVC summary template if needed
* Discuss SADL usage and up-dates
* Create list of new EE1 experiments
* Possibilities of establishing common SW base, requirements, recommendation for further development: better performance or reasonable complexity?
* **NNVC Results analysis**
* ***Discussion***

Initial version of NNVC results summary was sent to the JVET reflector at April, 21 (after JVET session 5). Proponents were invited to correct summary if needed. One correction (by JVET-Z0082) proponents received.

The group has identified that for several proposals not AhG11 anchor was used. It was requested to up-date contributions with proper anchor. It was decided not to include into the summary test results of proposals which do not use AhG 11 anchor.

* Proposals with wrong anchor
  + JVET-Z0073 ,
  + JVET-Z0077,
  + JVET-Z0144
* **Update NNVC reporting templates**
* ***Discussion***

Two document describe complexity and results reporting procedure: ***NNVC CTC and*** ***results reporting template*** [1] ([JVET-X2016](https://jvet-experts.org/doc_end_user/current_document.php?id=11229)) and ***summary of NNVC contribution*** [2] ([JVET-X0188](https://jvet-experts.org/doc_end_user/current_document.php?id=11198)). Both require minor changes. Some complexity aspects are mandatory, another are optional in [1].

It was discussed:

* It is desired to report more details about training strategy in document, among them:
  + Change of training data set at some point
  + Learning rate up-date strategy
* Several experts suggested to remove only 4K results [2] from summary of all tests 🡪 Agreed
* It was agreed to keep run time in [1] only, but not in [2] 🡪 no changes needed
* Is it useful to include training time for the biggest model?
  + Concern: reliable time measurement for training process is difficult (due to GPU loading with process, for example)
  + Possible option to measure 1 epoch (not first) and multiply by number of epochs
  + Number of epochs and training time are optional in [1], should be made mandatory to report
  + Time is not enough , we need number of samples in training, GPU loading
  + Time provided only as rough estimation ~ not useful all technologies summary [2]
  + Agreement: number epochs and training time mandatory to report in results reporting template [1].
* kMAC/pxl computation needs to be clarified:
  + If NN technology operates for 128×128 or for whole picture due to the extension on boundary kMAC/pxl numbers are different
  + SADL estimation for kMAC/pxl is more accurate
  + Boundary handling was not discussed in EE1 so far, but it could lead to non-homogeneous processing on boundary 🡪 significant slow down
  + Option1: indicate if kMAC/pxl assumes implementation on block basis (and size of block) or frame basis (++)
  + Option2: ask all to use kMAC/pxl for 128×128 (--)

Suggestion. If use SADL

* + - Compile w/o SIMD 🡪 accurate kMAC/pxl
    - Compile w/ SIMD 🡪 fast

Agreed: update results reporting template [1] requesting information about kMAC/pxl computation (block based on picture based), include this information to [2].

* Results for Post Filters do not appear on a plots in [2]. Agreement: fix graphs, do not differentiate between in-loop and post-filters on plots.
* **Discuss SADL usage and up-dates**
* ***Discussion***

Franck Galpin presented

[**JVET-Z0161**](https://jvet-experts.org/doc_end_user/current_document.php?id=11610) **AhG11: SADL update [F. Galpin, T. Dumas, P. Bordes, E. François (InterDigital)]**

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

It was explained that PyTorch and TensorFlow depending on version for the speed benefits due to the different memory layout can reorder data (for example, transpose) and so change order of operations, this is one of the reasons for not-bit exact behaviour.

In most of proposals NN diagram doesn’t not contain sufficient information to implement it in SADL or/and describe operation accurate enough for standard specification,

Eventually in order to ensure bit-exact behaviour of NN algorithms some restrictions (de-scaling shifts, clipping, fixed order of operation similar to HEVC/VVC interpolation filter or/and transform) will be required. This I would be “normative” part of the algorithm.

Floating-point arithmetic is non-associative, which cause unpredictable result in case of usage of computational systems with massive parallelism

Consideration of integer implementation and dynamic range control (in order to avoid overflow) is important to ensure bit-exact behaviour of NN algorithms.

Changes in SADL:

* Missing layers: usually easy to fix. 5 new layers were added to support proponents' models:
  + shape, expand, preLU, flatten, transpose
* Some layers are maybe **not recommended** as they are difficult to implement in integer arithmetic (e.g. **division, sigmoid**)
* Some issue with some “complex” layout: sometimes the memory layout in the graph is difficult to assess. A new option has been added to fine control the memory layout.
  + Example: transpose can be part of algorithm or be implementation issue due to memory layout handling

Performance:

* Initialization of the network should be done once to avoid overhead: as in VTM/ECM, buffers are prepared beforehand.
* NN related code should be isolated inside a translation unit and build with correct SIMD options
* Float version is less optimized than integer version. An update is planned.
* It is recommended to prepare the input data and output data directly in the C++ code instead of relying on data manipulation layers in SADL (reshape, concat, expand, etc.).

Quantization:

* Currently SADL can handle int32/int16/int8 data type. Data type is used for both the weights and the intermediate latents. Computation is done in using twice the number of bits.

Recommendations:

* Provide example c++ code for an inference.
* Sample program for NN parameters quantizing planned to be provided

Speed issue of SADL:

Redundant initialization of models 🡪 removing saves 10% run time (already shared to proponents)

Some SIMD have been missed 🡪 added shortly (implementation ready)

Expectation: SADL expected to be 3 ~4 times slower than libtorch, but transparent.

* **Create list of new EE1 experiments**
* ***Discussion***

Based on [JVET-Z\_Notes\_d2.docx](https://www.itu.int/wftp3/av-arch/jvet-site/2022_04_Z_Virtual/JVET-Z_Notes_d2.docx) following contribution recommended to be included into EE1.

[**JVET-Z0093**](https://jvet-experts.org/doc_end_user/current_document.php?id=11528) **EE1-related: The performance of EE1 Test 1.1 and Test 1.2 on ECM-4.0 [R. Chang, L. Wang, X. Xu, S. Liu (Tencent), F. Galpin (InterDigital)]**

[**JVET-Z0106**](https://jvet-experts.org/doc_end_user/current_document.php?id=11553) **EE1-related: Reduced complexity NN loop filter and ablation study [J. Ström, D. Liu, M. Damghanian, K. Andersson, Y. Li, P. Wennersten, R. Yu (Ericsson)]**

[**JVET-Z0128**](https://jvet-experts.org/doc_end_user/current_document.php?id=11576) **EE1-related: on BaseQP parameter in EE1-1.1 [Z. Xie, Y. Yu, H. Yu, D. Wang (OPPO)]**

[**JVET-Z0077**](https://jvet-experts.org/doc_end_user/current_document.php?id=11512) **AHG11: Extension of DOVC to Regular 2D Videos [Q. Qin, C. Jung (Xidian Univ.), D. Zou, M. Li (OPPO)]**

This list will be updated based on further discussion of NNV proposals in JVET (by time of BoG 6 proposals have been not yet presented).

* **Recommendations**

Recommendations from the BoG are summarized as follows:

* + - NNVC Results and analysis
      * + Results for AhG11, EE1related contributions added, confirmed (attached to the BoG report)
        + Results of proposals which do not sue AhG11 anchor should not be listed in the summary.
    - NNVC summary template modification:
      * Up-date results reporting template for each proposals (JVET-X2016)
        + The number epochs and training time **mandatory** to report in results reporting template,
        + Add information about assumption kMAC/pxl was computed (block based on picture based),
        + Request reporting more details about training process such as

Additional training data at some stage of training

Learning rate up-date strategy

* + - * Up-date summary of NNVC proposal template (JVET-X0188):
        + Remove “only 4K” results,
        + Fix graphs, do not differentiate between in-loop and post-filters on plots,
      * Add information about assumption kMAC/pxl was computed (block based on picture based).New EE1 Planning
    - Proponents of proposals included to the next EE1 round are requested to provide test description (send to EE1 coordinators):
      * [JVET-Z0093](https://jvet-experts.org/doc_end_user/current_document.php?id=11528) ,
      * [JVET-Z0106](https://jvet-experts.org/doc_end_user/current_document.php?id=11553) ,
      * [JVET-Z0128](https://jvet-experts.org/doc_end_user/current_document.php?id=11576),
      * [JVET-Z0077](https://jvet-experts.org/doc_end_user/current_document.php?id=11512)
      * List to be up-dated after completion NNVC proposals review
    - SADL up-date
      * Reasons for some speed issues identified, way to fix communicated to proponents
      * Expected speed difference between libtorch and SADL is between ×2 and ~×4
        + SADL is slower, but controllable and transparent
      * SADL authors believe that for video coding applications precision of weights 16 bits, activation 16 bits (w16a16) is the most suitable, so SADL optimized the best for this combination.
      * Recommended (among all other steps on SADL improvements) to provide an examples of
        + C++ code for correct MAC evaluation
        + Sample to dump a NN in quantized SADL format

Presented in session 10. Discussion on common code base – generally agreed that this would be advantageous. It was pointed out to have been agreed earlier only to include tools in such more “official” codebase when training would have been fully cross-checked which was not the case yet for tools in current EE.

(Add some more notes on requirements for such a code base from v3 of the report)

Training scripts should also provide the tools for extracting patches or other information that is fed into the training.

BoG met again Tuesday 26 April 0500-0700. The outcome is as follows, as reported in session 14, Tuesday 26 April 2100-2150:

**Common SW for NNVC technologies study**

* **R&D directions**
  + Different operation points for single NN-tool to later choose the best trade-off
    - Position (post/in-loop)
    - Number of residual blocks
  + Direction “top-left”
  + Ultimate target: NN tools with realistic complexity in few years
* **Code base**
  + NNVC (AhG11 anchor)

Requirements for proposals which go to common SW base:

* **Encoder**
* **Decoder** 
  + Cross-platform capability to be checked during integration to common SW base (not before making decision on algorithm)
* **Precise algorithm description** (pseudo spec or c++ code) 🡪 agreed to be **mandatory** for addition to common SW base
  + How much details?
    - Goal: sufficient enough for understanding proposal with enough detail to verify complexity aspects (for example, MAC), implementation details for different platforms....
    - Transpose, Reshape, Padding, ... (typically not shown in diagrams provided in contributions)
  + Option 1:
    - Take your NN
    - Convert to Onnx (procedure is described in SADL documentation)
    - Dump (instruction to be added to SADL documentation)
  + Option 2:
    - SADL format (examples of code JVET-Y-EE1-1.1, 1.2, 1.6, 1.7)
* **Training scripts**
  + Agreement:
    - Upload **training code** with all details sufficient for **reproducing training + cross-check for training**
    - “Training scripts” = training data 🡪 (*processing, learning algorithm, quant*, ...) 🡪models
  + Time line? When?
    - Conditional adoption (was used for training based VVC technologies)
    - [**ask JVET to add phrase**]: group recommends to adopt technologies to the common SW basis (subject for the cross-check of training)
  + Goal: close enough performance after re-training with same number of epochs
    - “Close performance” **Th** BD-rate difference between model trained by proponent and cross-checker
      1. **Th** = **0.1%** (exact value for this threshold can be reconsidered after we’ll get actual data)
      2. **Th** depends on overall gain tools shows
    - Performance difference higher than the threshold requires an explanation (discussion with proponent or in group)
    - No agreement on the meaning of “close performance”
* Quantized Network to ensure device interoperability

Additional recommendations were as follows:

* + - Common SW for NNVC study
      * Goal: NN technologies with realistic complexity in few years
      * Multiple operation points (even for single tool) to choose from
      * Procedure for adding NN technologies to the common SW base: 1) cross-check of performance in EE1, 2) transparent implementation and description (SADL), 3) training procedure disclosed and cross-checked.
      * BoG recommends to add following technologies to the common SW basis, subject to providing training scripts (see definition above), and successful verification of the training procedure and comparable results, cross-checked successfully:
        + JVET-Z0091 (Tencent, InterDigital filter with single model, JVET-Y-EE1-1.2.2, SADL)
        + JVET-Z0113 (Bytedance, Qualcomm, Ericsson, InterDigital filter with 8 RBs model, JVET-Y-EE1-1.7, SADL)

In the JVET plenary, the recommendations of the BoG are confirmed.

The process of “conditional adoption” is understood as follows:

* Promising technology is identified from the EE, compression performance of provided model successfully crosschecked
* Subsequently, the cross-check of the training is performed (as described above)
* If successfully verified, the technology will be adopted to the software code basis

Codebasis to be defined by next meeting, the current view expressed in the BoG was that it would be based on the AHG11 anchor. It is confirmed that it will be a single codebase with multiple configurable tools.

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

Contributions in this area were discussed in the context of the summary report JVET-Z0023 (see section 5.2.1), and in session 4 on Thursday 21 April 0510-0640 (chaired by JRO).

[JVET-Z0065](https://jvet-experts.org/doc_end_user/current_document.php?id=11500) EE1-2.1: RPR encoder with multiple scale factors [J. Nam, S. Yoo, J. Lim, S. Kim (LGE)]

(abstract)

Is sufficiently reflected in EE summary, no need for presentation. May be interesting to be used in NN-based super-resolution upsampling.

[JVET-Z0071](https://jvet-experts.org/doc_end_user/current_document.php?id=11506) Cross-check of JVET-Z0065: EE1-2.1: RPR encoder with multiple scale factors [K. Andersson (Ericsson)]

[JVET-Z0070](https://jvet-experts.org/doc_end_user/current_document.php?id=11505) EE1-1.3: Combination of deblocking and NN [K. Andersson, J. Ström, D. Liu, R. Sjöberg (Ericsson)]

(abstract)

No change relative JVET-Y0098 (make sure that either NN-based or conventional deblockling is always invoked). No need for detailed presentation.

It was reported that the method investigated in the test was not exactly using the original approach, as another method of NN based deblocking (from JVET-Y0143) was combined with it (EE test 1.7). It appeared that some artefacts were still visible in that combination. It was suggested that it might be useful to also compare subjectively with the original approach from JVET-Y0098.

[JVET-Z0073](https://jvet-experts.org/doc_end_user/current_document.php?id=11508) EE1-1.5: Test on NN-based filter as proposed in JVET-Y0052 [H. Zhang, C. Jung (Xidian Univ.), D. Zou, M. Li (OPPO)]

This contribution presents the EE results of JVET-Y0052. It mainly includes the test results of the CNN-based post-processing filter proposed in JVET-Y0052 and some ablation studies. The ablation studies include comparing the contributions of standard convolution and depthwise separable convolution to the proposed filter as well as evaluating the performance of the proposed filter as an in-loop filter in VTM. The ablation experiments reveal that depthwise separable convolution can save nearly four times the number of parameters with comparable performance under the same configuration compared to standard convolution.

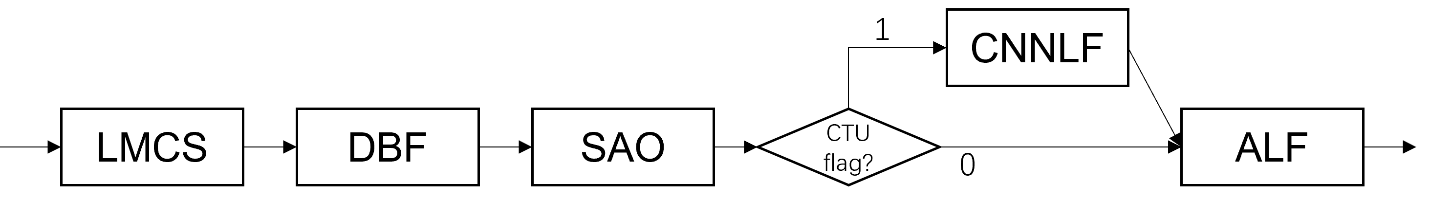


Fig. 5 Pipeline of LDSCA. L: LMCS. D: DBF. S: SAO. C: CNNLF. A: ALF.

After compare five embedding methods, it was concluded that the best embedding method is "LDSCA“ best embedding method is "LDSCA“

• The slice level flag and the CTU are enabled

• Turning off DBF and SAO for I slice can improve the performance

• For P and B slice, DBF SAO need to be enabled

Performance comparison of different embedding methods   
(C class in RA, more detailed results in document)

|  |  |  |
| --- | --- | --- |
| Embedding methods | Y-PSNR | Y-MSIM |
| LCDSA | -2.54% | -2.31% |
| LDCSA | -2.81% | -2.62% |
| LDSCA | -2.78% | -2.61% |
| LDSAC | -2.67% | -2.78% |
| LCA | -2.24% | -2.00% |

Filters were not retrained.

Main advantage of separable filters is lower number of parameters, number of kMAC/pixel is not reduced. However, it should be possible to reduce in an optimum implementation.

[JVET-Z0086](https://jvet-experts.org/doc_end_user/current_document.php?id=11521) EE1-1.4: ALF improvement for NNVC [W. Zou, Y. Zhou (Xidian Univ.), C. Huang, Y. X. Bai (ZTE)]

This contribution presents an ALF modification in NNVC. When a neural network-based filter is used as in-loop filter, two ALF flags are introduced to indicate ALF enabled/disable for luma and chroma components, respectively.

It is reported that, compared with JVET-X0065, the overall performance is as follows:

Test1.4.1: AI: 0.03%/-0.96%/-1.04% with 100%EncT/100%DecT; RA: 0.05%/-0.30%/-0.41% with 100%EncT/100%DecT.

Test1.4.2(By setting chroma QP offset to 1): AI: 0.03%/-1.07%/-1.09% with 100%EncT/100%DecT; RA: 0.xx%/-0.xx%%/-0.xx%% with xx%EncT/xx%DecT.

Compared with VTM-11.0 + NewMCTF, the overall performance is as follows:

Test1.4.1: AI: -7.38%/-20.74%/-20.95% with 124%EncT/27071%DecT; RA: -10.03%/-23.91%/-23.50% with 165%EncT/62196%DecT.

Test1.4.2(By setting chroma QP offset to 1): AI: -8.71%/-11.85%/-12.24% with 122%EncT/27092%DecT; RA: -10.41%/-13.73%%/-13.27%% with 163%EncT/63064%DecT.

No change relative to the previous proposal. JVET-Y0046, no need for detailed presentation.

[JVET-Z0186](https://jvet-experts.org/doc_end_user/current_document.php?id=11636) Crosscheck of JVET-Z0086 (EE1-1.4: ALF improvement for NNVC) [C. Lin (Bytedance)] [late]

[JVET-Z0091](https://jvet-experts.org/doc_end_user/current_document.php?id=11526) EE1-1.2: Neural network based in-loop filter with a single model [L. Wang, X. Xu, S. Liu (Tencent), F. Galpin (InterDigital)]

(abstract)

Was sufficiently discussed in EE1 summary. Some minor changes relative to the previous proposals JVET-Y0078 and JVET-Y0080, no need for detailed presentation. It is noted that the same method is used in JVET-Z0094.

Which of the two variants (1.1 from Z0094, 1.2 from Z0091) would be preferable? The proponents do not have a preference, but it is pointed out that 1.2 has less number of parameters (using only one model for all slice types). On the other hand, 1.2 has 1% less coding gain in AI.

It was said that JVET-Z0092 is an extension of this method (using network quantization) which appears more relevant for further study. However, later the proponents of JVET-Z0092 expressed that they are not interested in joining the EE.

Further investigate JVET-Z0091 and JVET-Z0094 in EE.

[JVET-Z0178](https://jvet-experts.org/doc_end_user/current_document.php?id=11628) Crosscheck of JVET-Z0091 (EE1-1.2: Neural network based in-loop filter with a single model) [Z. Xie, Z. Dai (OPPO)] [late]

[JVET-Z0094](https://jvet-experts.org/doc_end_user/current_document.php?id=11529) EE1-1.1: Neural network based in-loop filter with 2 models [L. Wang, X. Xu, S. Liu (Tencent), F. Galpin (InterDigital)]

(abstract)

See notes above under JVET-Z0091.

[JVET-Z0177](https://jvet-experts.org/doc_end_user/current_document.php?id=11627) Crosscheck of JVET-Z0094 (EE1-1.1: Neural network based in-loop filter with 2 models) [Z. Xie, Z. Dai (OPPO)] [late]

[JVET-Z0096](https://jvet-experts.org/doc_end_user/current_document.php?id=11532) EE1-2.2: CNN-based Super Resolution for Video Coding Using Decoded Information [C. Lin, Y. Li, K. Zhang, L. Zhang (Bytedance)]

(abstract)

Same as previous proposal JVET-Y0069, no need for detailed presentation.

[JVET-Z0097](https://jvet-experts.org/doc_end_user/current_document.php?id=11533) EE1-2.3: CNN-based Super Resolution for Video Coding Using Separate Networks for Chroma Components [C. Lin, Y. Li, K. Zhang, L. Zhang (Bytedance)]

(abstract)

Same as previous proposal JVET-Y0070, no need for detailed presentation.

[JVET-Z0098](https://jvet-experts.org/doc_end_user/current_document.php?id=11534) EE1-2.4: CNN-based Super Resolution for Video Coding with GOP Level Adaptive Resolution [C. Lin, Y. Li, K. Zhang, L. Zhang (Bytedance), J. Nam, S. Yoo, J. Lim, S. Kim (LGE)]

This contribution reports the EE1-2.4 test results, which is a combination of JVET-Y0070 and JVET-Y0068. 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 is the latter case. Compared with VTM-11.0 with new MCTF enabled, the simulation results show {-2.95%, -1.42%, -1.68%} and {-2.71%, -0.80%, -1.50%%} BD-rate savings for {Y, Cb, Cr} under RA and AI configurations, respectively.

It is reported that gain is also observed for class B, which had not been the case before (sequences Market Place and Ritual Dance). The proponents also mention that the comparion against RPR is not completely fair, as the latter also uses an additional scaling factor 1.5x. It was noted that implementing 1.5x on NN technology may not be possible with the current architecture, which uses the stride for upscaling by factor 2. It was pointed out that JVET-X0074 was another proposal that had used 1.5x.

Further study of using more flexible scaling factors would be desirable.

How large is the number of cases with low resolution coding? Needs more analysis.

It is pointed out that using NN technology also for downsampling might be an option to improve performance, but that has not been studied so far.

[JVET-Z0112](https://jvet-experts.org/doc_end_user/current_document.php?id=11559) EE1-1.6: Test on Deep In-Loop Filter with Adaptive Parameter Selection and Residual Scaling based on SADL implementation [Y. Li, K. Zhang, L. Zhang (Bytedance), [H. Wang](mailto:hongtaow@qti.qualcomm.com), M. Coban, A. M. Kotra, M. Karczewicz (Qualcomm), F. Galpin (InterDigital)]

(abstract)

Almost identical with JVET-Y0143, only difference in floating point calculation pytorch vs. SADL. No need for detailed presentation. Proponents would rather prefer further study of 1.7 (see notes under JVET-Z0113).

[JVET-Z0170](https://jvet-experts.org/doc_end_user/current_document.php?id=11620) Cross-check of JVET-Z0112 (EE1-1.6: Test on Deep In-Loop Filter with Adaptive Parameter Selection and Residual Scaling based on SADL implementation) [Z. Dai (OPPO)] [late]

[JVET-Z0113](https://jvet-experts.org/doc_end_user/current_document.php?id=11560) EE1-1.7: Combined Test of EE1-1.6 and EE1-1.3 [Y. Li, K. Zhang, 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 experimental results of the EE1-1.7, i.e., the combined test of EE1-1.6 and EE1-1.3. In this combined test, convolutional neural network-based models are utilized for in-loop filtering and implemented with SADL. Two tests with CNN models of 8 and 16 residual blocks are presented to provide two different tradeoffs.

Compared with VTM-11.0 + NewMCTF, CNN models of 8 residual blocks reportedly show on average {10.18%, 22.16%, 22.11%} and {7.24%, 19.68%, 20.02%} BD-rate reductions for {Y, Cb, Cr} under RA, and AI configurations while CNN models of 16 residual blocks reportedly show on average {11.35%, 23.92%, 25.36%} and {8.18%, 22.22%, 23.30%} BD-rate reductions for {Y, Cb, Cr} under RA, and AI configurations.

A scaling factor for the residual signal generated by the neural network is signaled per color component and per picture. The criterion is minimization of MSE. The factor is quantized.

There is also a block-level control for disabling NN filtering. In that case, conventional DBF is invoked.

Further study in EE on SADL implementation (fixed-point), and better adaptation e.g. of scaling factor optimization to reduced artifacts which were observed in subjective viewing.

[JVET-Z0171](https://jvet-experts.org/doc_end_user/current_document.php?id=11621) Cross-check of JVET-Z0113 (EE1-1.7: Combined Test of EE1-1.6 and EE1-1.3) [Z. Dai (OPPO)] [late]

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

Contributions in this area were discussed in session 4 at 0640–0710 UTC and session 5 at 0735–0845 on Thursday 21 April 2022 (chaired by JRO).

[JVET-Z0087](https://jvet-experts.org/doc_end_user/current_document.php?id=11522) EE1-related: ALF-SPLIT based on JVET-Z0070 and JVET-Y0078 [W. Zou, Y. Zhou, C. M. Gu (Xidian Univ.), C. Huang, Y. X. Bai (ZTE)]

An ALF split scheme was proposed in JVET-Y0046 which indicates ALF enabled/disable for luma and chroma components, respectively. To verify the coding performance, ALF-SPLIT was integrated on top of JVET-Z0070 and JVET-Y0078, respectively. It is reported that, when ALF-SPLT is integrated into JVET-Z0070, compared with JVET-Z0070, the overall coding performance impact for {Y, U, V, EncT, DecT} is { 0.02%, -0.62%, -0.79%, 100%, 100%} for AI and {xx} for RA; compared with VTM-11.0 + NewMCTF, the overall coding performance impact for {Y, U, V, EncT, DecT} is { -7.24%, -20.29%, -20.77%, 147%, 34360% } for AI and {xx} for RA. When ALF-SPLIT is integrated into JVET-Y0078, compared with JVET-Y0078, the overall coding performance impact for {Y, U, V, EncT, DecT} is { 0.02%, -0.69%, -0.43%, 100%, 100% } for AI and {xx} for RA; compared with VTM-11.0 + NewMCTF, the overall coding performance impact for {Y, U, V, EncT, DecT} is { -7.32%, -19.13%, -18.73%, 129%, 34209% } for AI and {xx} for RA.

“ALF split” method was also investigated to be beneficial in combination with other proposals such as JVET-Z0070.

No impact on encoder/decoder run time. Beneficial for chroma only, mainly for classes C (highest for Party Scene) and D. Very slight BD rate increase on luma, due to additional signalling.

No intent to investigate in EE, the information brought in the proposal is interesting.

[JVET-Z0092](https://jvet-experts.org/doc_end_user/current_document.php?id=11527) EE1-related: Additional results on Test 1.1 and Test 1.2 with 8-bit quantization [L. Wang, X. Xu, S. Liu (Tencent), F. Galpin (InterDigital)]

In this contribution, model quantization is further studied based on EE1-1.1 and EE1-1.2 using 8-bit quantization. Two neural network based in-loop filters with the smaller memory size are proposed in this contribution, compared with the implementations in EE1 tests. In all, only two or even one single model is used in the proposed filter design. In addition to the reconstruction image, other side information, such as the prediction image, partition image, slice QP and based QP, is fed into the network. Based on the EE1 anchor, the related results are shown in order of RA, LB and AI configurations as follows.

Filter 1 (based on EE1-1.1, 2 models, 8-bit, SADL implementation):

{-8.24%, -18.96%, -18.14%}, {-6.94%, -16.33%, -15.25%}, {-6.28%, -14.98%, -16.14%}

Filter 2 (based on EE1-1.2, 1 model, 8-bit, SADL implementation):

{-8.63%, -18.09%, -18.43%}, {-7.45%, -15.46%, -16.50%}, {-6.34%, -14.36%, -15.41%}

The main difference of EE1 1.1 and 1.2 with the previous proposals is shown as follows:

• The convolutional block attention module is removed for the simplicity. It seems to have a tiny influence on the performance.

• A simpler Nearest method is used to replace Lanczos method in the resampling process chroma components.

• Multi-scaling refinement method, like the one in JVET Y0098, is used to further improve the performance.

• SADL deployment is studied.

Implementation

• For I slices and B slices, Deblock SAO are both enabled. The optimal filtered result is decided between the two outputs from proposed filter and SAO.

• A scaling operation is carried out to refine the result of NN filter. The factors are 3 fixed weights to blend These fixed weights are 1, 0.75 and 0.5, which are similar as those in JVET-Y0098.

• The proposed filter can be turned on/off at the CTU level and slice level.

In the discussion, it was pointed out that in inference stage still floating point had been used. It was recommended that true integer implementation should be used overall.

It was suggested that a more optimized method of network parameter quantization (e.g., 16 bits instead of 8 bits) might be beneficial to reduce the losses. It was also pointed out that 8 bit quantization is currently not optimized in SADL.

Proponents are not interested in joining EE.

[JVET-Z0093](https://jvet-experts.org/doc_end_user/current_document.php?id=11528) EE1-related: The performance of EE1 Test 1.1 and Test 1.2 on ECM-4.0 [R. Chang, L. Wang, X. Xu, S. Liu (Tencent), F. Galpin (InterDigital)]

This contribution presents the test results of EE1-1.1[[[1]](#endnote-2)] and EE1-1.2 [[[2]](#endnote-3)] when implemented on ECM-4.0, which is also the EE2 anchor. The proposed filter 1 from EE1-1.1 [1] and filter 2 from EE1-1.2 [2] are integrated into ECM-4.0 respectively. Further, results with the corresponding SADL implementations are provided and compared.

Based on the EE2 anchor, the related results are shown in order of RA, LB and AI configurations as follows.

Filter 1 (2 models, Libtorch implementation):

{x.xx%, x.xx%, x.xx%}, { x.xx%, x.xx%, x.xx%}, { x.xx%, x.xx%, x.xx%}

Filter 1 (2 models, SADL implementation):

{x.xx%, x.xx%, x.xx%}, { x.xx%, x.xx%, x.xx%}, { x.xx%, x.xx%, x.xx%}

Filter 2 (1 model, Libtorch implementation):

{x.xx%, x.xx%, x.xx%}, { x.xx%, x.xx%, x.xx%}, { x.xx%, x.xx%, x.xx%}

Filter 2 (1 model, SADL implementation):

{x.xx%, x.xx%, x.xx%}, { x.xx%, x.xx%, x.xx%}, { x.xx%, x.xx%, x.xx%}

Based on the EE1 anchor, the related results are shown in order of RA, LB and AI configurations as follows.

Filter 1 (2 models, Libtorch implementation):

{x.xx%, x.xx%, x.xx%}, { x.xx%, x.xx%, x.xx%}, { x.xx%, x.xx%, x.xx%}

Filter 2 (1 model, Libtorch implementation):

{x.xx%, x.xx%, x.xx%}, { x.xx%, x.xx%, x.xx%}, { x.xx%, x.xx%, x.xx%}

Presentation deck not uploaded.

Numbers above need update from a new version.

Models were retrained for the case of ECM. It was suggested that it would be interesting to see results without retraining (i.e. trained on VTM) as well. Also, fine-tuning starting from the VTM model could be used rather than training from scratch.

Integer implementation (SADL) has only minor drop in performance.

Overall, up to 24% gain compared to VTM in RA.

It was suggested to investigate the aspects of training/retraining in EE.

[JVET-Z0106](https://jvet-experts.org/doc_end_user/current_document.php?id=11553) EE1-related: Reduced complexity NN loop filter and ablation study [J. Ström, D. Liu, M. Damghanian, K. Andersson, Y. Li, P. Wennersten, R. Yu (Ericsson)]

This contribution presents an ablation study of the intra model in the neural network-based loop filter from JVET-Y0143. It is stated that the network is trained several times from scratch, each time changing a single variable, such as removing one of the inputs to the filter. Some changes are claimed to hurt BD-rate performance, such as removing the prediction input. It is further claimed that other changes do not seem to impact the BD-rate performance. As an example, the contribution states that removing the partitioning input lowers the complexity of the intra luma-filter without harming the BD-rate, giving BD-rate figures over the VTM-11.0 + newMCTF anchor of BDR-Y: -7.57% (AI) -9.82% (RA), -8.54% (LDB) while keeping the intra chroma model and the inter models the same as in JVET-Y0143. It is further reported that removing the attention branch in all residual blocks except the last one decreases BDR-Y performance by 0.01% (AI). In another test, it is reported that removing all attention branches has no negative impact on BDR performance (a gain of -0.01% (AI)). The contribution proposes to further study the best inputs and attention mechanisms for NN loop filters in an exploration experiment.

So far, only intra luma was investigated.

Somewhat different training approach was used than in original proposal JVET-Y0143, other GPU, other QP range (lower by 2), other number of epochs, learning rate decreased.

Proponents would not be willing to provide training script (as long as others don’t do this as well).

Further, attention input (partitioning, local DBF boundary strength, prediction, residual) was partially removed. Partitioning and BS should not both be removed.

Investigate in EE, in particular the need for particular attention mechanisms, and impact of training.

[JVET-Z0203](https://jvet-experts.org/doc_end_user/current_document.php?id=11653) Crosscheck of JVET-Z0106 (EE1-related: Reduced complexity NN loop filter and ablation study) [Y. Li (Bytedance)] [late]

[JVET-Z0128](https://jvet-experts.org/doc_end_user/current_document.php?id=11576) EE1-related: on BaseQP parameter in EE1-1.1 [Z. Xie, Y. Yu, H. Yu, D. Wang (OPPO)]

This contribution proposes a QP-adjustment method based on EE1-1.1 to improve the coding performance of EE1-1.1. In EE1-1.1, the QP parameters are part of the input to the network at the inference stage and these QP parameters include BaseQP and SliceQP. It is proposed to add an offset to the BaseQP at the frame-level and use the final adjusted BaseQP as the input of the network. Compared with the EE1-1.1, the proposed method can further improve the coding efficiency shown as below:

RA: Y -0.30%, U -0.11%, V -0.13%;

LDB: Y -0.48%, U -0.76%, V -0.62%.

Could also be combined with EE1-1.1.2.

Encoder run time is increased significantly (40-50% compared to EE1-1.1.1).

Investigate in EE (in combination with both 1.1.1 and 1.1.2). Also identify possibility of reducing encoder run time.

[JVET-Z0154](https://jvet-experts.org/doc_end_user/current_document.php?id=11603) EE1-1.6-related: Improved RDO Considering Deep In-Loop Filter [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 in EE1-1.6. The proposed method is asserted to improve the compression efficiency at encoder by introducing CNN-based filtering into the rate-distortion optimization (RDO) process. Compared with VTM11.0\_nnvc, the proposed method reportedly shows on average {8.85%, 21.61%, and 22.73%}, {11.55%, 23.15%, and 24.62%}, and {9.08%, 17.21%, 18.00%} BD-rate reductions for {Y, Cb, Cr} components, under AI, RA, and LDB configurations, respectively.

Gain compared to EE1-1.1.6 is around 0.5-0.6%, some increase of encoder runtime.

Some simplification of model, but no details disclosed on that (not encoder only method).

Proponents are not intending to participate in EE with this.

[JVET-Z0155](https://jvet-experts.org/doc_end_user/current_document.php?id=11604) EE1-1.7-related: Improved RDO Considering Deep In-Loop Filter and Deblocking [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 EE1-1.7. 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. Compared with VTM11.0\_nnvc, the proposed method reportedly shows on average {8.52%, 22.43%, 23.46%}, { 11.68%, 23.93%, 25.33%}, and {9.43%, 18.26%, 19.64%} BD-rate reductions for {Y, Cb, Cr} components, under AI, RA, and LDB configurations, respectively.

Gain compared to EE1-1.1.7 is around 0.2-0.4%, some increase of encoder runtime. This is an encoder-only method.

Proposal for information, no action proposed.

### Other NN technology related contributions (7)

Contributions in this area were discussed in session 5 at 0845–0920 UTC on Thursday 21 April 2022 (chaired by JRO), in session 9 at 2215–2250 UTC on Friday 22 April 2022, and in session 10 at 0840–XXXX UTC on XXday 2X April 2022 (chaired by JRO).

[JVET-Z0074](https://jvet-experts.org/doc_end_user/current_document.php?id=11509) AHG11: Neural Network Based Motion Compensation Enhancement for Video Coding [C. Ma, R.-L. Liao, Y. Ye (Alibaba)]

This contribution proposes to enhance the motion compensated prediction of coding units with neural network. The proposed method is performed on all inter-coded coding units with a flag in CU level to control the proposed method on or off. Experimental results demonstrate that, compared with VTM-11.0-nnvc-1.0, the proposed method achieves 1.31%, 6.96%, and 6.39% BD-rate reductions for Y, U, and V components under RA configuration, respectively.

Follow-up of JVET-Y0090, improved performance. New elements: CU-level flag, support for non-square blocks, applied to chroma as well.

In total, 120 models for different block sizes and different QPs.

Similar gain in luma as in previous method (but simplified model, e.g. complexity decreased due to switching). Gain in chroma is new.

Investigation if gain is additive with NN based loop filters would be interesting.

Analysis how frequent the mode is used? Not presently available.

More extensive ablation study?

Contribution for information – no particular action suggested.

Further study to improve performance.

[JVET-Z0077](https://jvet-experts.org/doc_end_user/current_document.php?id=11512) AHG11: Extension of DOVC to Regular 2D Videos [Q. Qin, C. Jung (Xidian Univ.), D. Zou, M. Li (OPPO)] [late]

This contribution reports extension of the deep omnidirectional video compression framework (DOVC) to regular 2D videos. Two DOVC frameworks were provided in JVET-X0043 and JVET-Y0051, namely DOVC-RGB and DOVC-YUV, respectively, and DOVC-YUV is conversion of DOVC-RGB from the RGB 4:4:4 domain to the YUV 4:2:0 domain. In this contribution, we apply DOVC-YUV to the regular 2D video dataset after retraining to verify its effectiveness. We generate 5 different models for both luma and chroma components according to 5 lambdas, i.e. 256, 512, 1024, 2048, and 4096. Compared with VTM-16.0 under RA configuration, DOVC-YUV achieves average BD-rate reductions of {15.525% (Y), 18.948% (U), and 14.995% (V)} and average BD-PSNR gains of {-0.2045dB (Y), -0.3541dB (U), and -0.2761dB (V)} on the complete dataset. Furthermore, DOVC-YUV takes advantage of GPU parallel processing and thus the average encoding time of DOVC-YUV is only 0.022 times that of VTM-16.0.

The method uses GOP length 8, i.e. 7 frames between keyframes are synthesized in an end-to-end NN approach. Different models used for different resolutions.

The proponents suggest study of the method in EE.

VTM16 was used as anchor.

It was suggested to also study the benefit of using VVC intra instead of BPG for keyframe coding, or also code some of the keyframes as B pictures, to come closer to the RA condition of I period. This might close the gap compared to VTM.

Investigate in EE.

[JVET-Z0082](https://jvet-experts.org/doc_end_user/current_document.php?id=11517) AHG11: Content-adaptive neural network post-filter [M. Santamaria, R. Yang, F. Cricri, J. Lainema, R. G. Youvalari, H. Zhang, G. Rangu, H. R. Tavakoli, H. Afrabandpey, M. M. Hannuksela (Nokia)]

This contribution presents a content adaptive neural network (NN) based post-filter. The filter is trained offline on general video/image content. The content adaption is achieved by means of overfitting the NN post-filter on the test video. The result of this process is a weight-update which is coded with the MPEG NNR standard and signalled within the video bitstream as an NNR post-filter SEI message. The approach was evaluated on top of VTM 11.0 NNVC 1.0. It is reported that the overall BD-rate in RA configuration is -5,01% (Y), -18,95% (U), -17,33% (V) over classes A1, B and C. It is also reported that content adaptation improves the overall BD-rate gains of class C in RA configuration by -1,23% (Y), -11,65% (U), -11,01% (V). Weight updates in fine tuning signalled with NNR version 2.

Different from previous proposal (JVET-Y0059), no boundary strength input to neural network.

Encoding time (CPU) 1.25x, decoding 45-50x. Encoding time does not include retraining.

Rates include overhead for signalling the network. What is the size of the NNR bitstream? Proponents do not know.

Overfitting is done for whole sequence.

NN based post processing would not be normative.

Network is relatively small, due to finetuning specifically for a given sequence no deep network is needed.

Rate for weight updates is not included in intra refresh random access periods.

Investigate in EE1, also implementation of SEI message proposal, and the rate that is consumed for the filter information. Also SADL implementation planned.

[JVET-Z0088](https://jvet-experts.org/doc_end_user/current_document.php?id=11523) AHG11: A CNN-based Super Resolution Method Combined with Existing RPR Functionality [S. Peng, D. Jiang, J. Lin, C. Fang, X. Zhang (Dahua)]

This contribution presents a super-resolution method that combines CNN and existing RPR functionality. In this method, the CNN model is derived from the previous contribution JVET-Y0087, and the image up-sampled by RPR is added to the output of CNN. Compared with VTM-11.0-NNVC, the experimental results show 10.34% and 6.83% BD-rate gains on average for luma, under AI and RA configurations, respectively.The architecture of the proposed CNN model is shown in Fig.1. The network contains 16 advanced residual blocks(ARB), several convolutional layers, a concatenate layer, a shuffle layer and a shortcut connection from conv1 to conv2. In addition, a global connection from input to output is included.

The parameters of each convolution layer are [*cin*, *k*, *k*, *cout*], where *cin*, *k* and *cout* represent the number of input channels, the size of convolution kernel and the number of output channels respectively.

The *YRpr* represents reconstruction up-sampled by RPR, *YLR-Rec* represents reconstruction, *YLR-Pred* represents prediction, and QP represents quantization parameter, which are used as the input of the model. QP is expanded to a matrix of the same shape as *YLR-Rec* before entering the concatenate layer. The output of the model is the super-resolution reconstruction *YSR*.

In addition, the architecture of the designed ARB is shown in Fig.2, where ReLU, GAP and Softmax are activation function, global average pooling layer and normalization function, respectively. The channel attention extraction module is circled by a dotted line in Fig.2.



*Architecture of the proposed CNN model*



*Architecture of the designed ARB*

As shown in Fig. 1, the network generates a residual added on top of the upsampled RPR output. This is a new aspect.

It was commented that it would be beneficial to use a similar adaptation approach as in JVET-Z0065. Resolutions other than 4K should be included.

From current results, appears comparable to JVET-Z0097 but less complex.

It was pointed out that the current results are no rate-matched for the downsampled version (which was a condition for the EE to make proposals comparable).

Investigate in EE.

[JVET-Z0089](https://jvet-experts.org/doc_end_user/current_document.php?id=11524) AHG11: An Improved Unet-Based In-Loop Filter Method [C. Fang, J. Lin, D. Jiang, X. Zhang, S. Peng (Dahua)]

This contribution presents a convolutional neural network-based in-loop filtering method with QP based models. In this contribution, we modify the training strategy and Unet-Based model from JVET-Y0086 for the inter frame.

Compared with VTM-11.0-NNVC, the proposed method reportedly shows on average {6.03%, 20.34%, 21.13%} BD-rate reductions for {Y, Cb, Cr} under AI configuration, and {2.33%, 14.55%, 13.07%} BD-rate reductions for {Y, Cb, Cr}, under RA configurationThe basic architecture of Unet-Based neural network is shown in Fig.1.

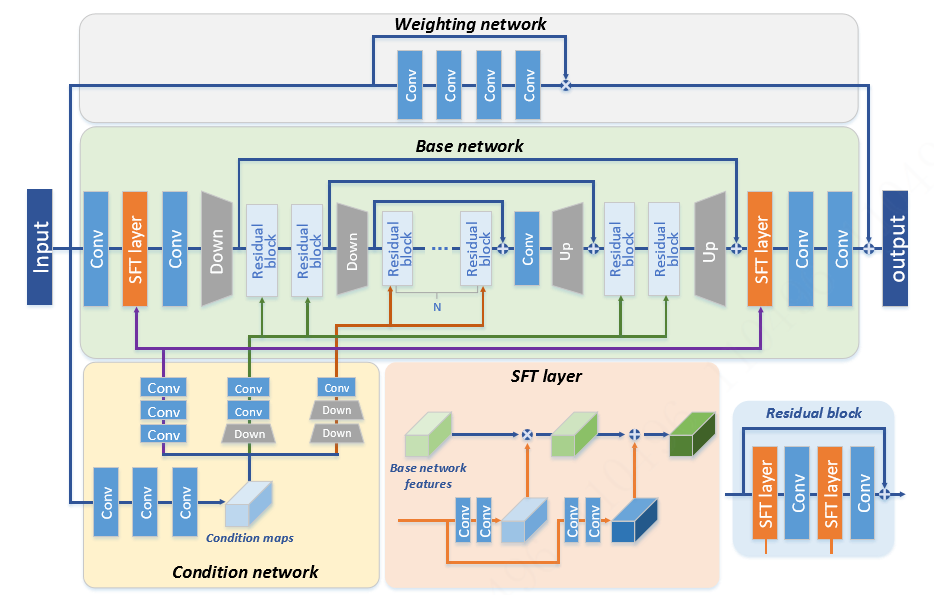
The input size for luma NN is 160x160, including the current CTU and 16 neighboring samples to each side of the current CTU. The chroma samples are up-sampled to 160x160, as an additional channel of the input tensor.

The input size for chroma NN is 80x80, including the current CTU and 8 neighboring samples to each side of the current CTU. The luma samples are down-sampled to 80x80, as an additional channel of the input tensor.

The architecture has 12 residual blocks in total. The kernel of the convolution layer in residual block is 3x3, and the channel number of the convolution layser in residual block is 64.

Other details related to the network architecture are listed as below:

1. Different groups of models are trained for luma and chroma components, the two chroma components share the same network parameters.
2. The CNN in-loop filter tool is applied after SAO and before ALF. An enable flag in SPS is set to control CNN in-loop filter tool, and a frame level enable flag for each component is also set to decide whether utilize NN filtering method to it or not.



*Fig.1*. *Architecture of Unet-Based neural network.*

3. For the inter frame, the channel number of the residual block is added to 96, and the 1x1 convolution layer is added before every 3x3 convolution layer, as shown in Fig.2.



*Fig.2*. *Architecture of residual block for inter model*

Compared to proposals investigated in EE, the tradeoff complexity versus performance is far from attractive. According to proponents, training may need to be improved.

[JVET-Z0101](https://jvet-experts.org/doc_end_user/current_document.php?id=11548) AHG11: Post-process filter based on fusion of CNN and transformer [T. Liu, W. Cui, C. Hui, F. Jiang (Harbin Inst. Tech.), Y. Gao, S. Xie, P. Wu (ZTE)]

This contribution presents a post-process filter based on fusion network, which combines convolutional neural network and transformer by channel-wise attention mechanism. The performance is evaluated using PSNR and MS-SSIM based DB-rate reduction in VVC Test Model 11.0 with Neural Network for Video Coding (NNVC) technology 1.0. It is reported that the average DB-rate reduction of class A1, A2, B, C, D is {-1.66%, -6.31%, -3.63%} on PSNR and {-1.34%, -5.64%, -2.56%} on MS-SSIM under random access configuration. It is also reported that the average DB-rate reduction of class F is {10.60%, 0.19%, 1.80%} on PSNR and {-0.97%, -1.69%, 0.23%} on MS-SSIM under random access configuration.The network architecture is shown in Fig. 1. It includes three parts: head, backbone and reconstruction. The head part consists of one 3×3 convolutional layer, which is used to extract the shallow features of the input frame. The backbone consists of M deep fusion blocks (DFBs) containing 6 residual blocks (RBs). The reconstruction part consists of one 3×3 convolutional layer. In addition, a residual connection is used between input lossy frame and output enhanced frame to learn global residual information.

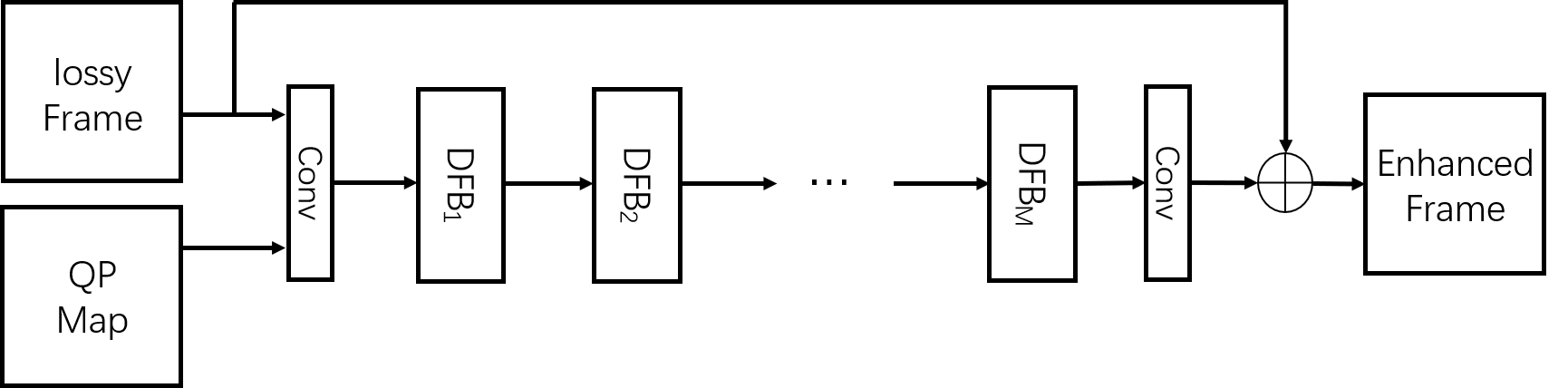


Fig. 1 Architecture of DFNN

One of the inputs to DFNN is 3-channels lossy frame, of which UV components are resized to the same size as Y component, and another input is the QP map of the lossy frame. Thus, the input is a 4-channels tensor, and the output is a 3-channel tensor.

The architecture of proposed DFB is shown in Fig. 2. DFB includes RBs to pre-process input features, GAB as well as LAB comprised of convolutional layers and Swin Transformer blocks [1] respectively to extract restoration features from different perspectives, and linear layers to calculate channel wise fusion weight of GAB and LAB. The output of DFB is the sum of weighted LAB feature and GAB feature.

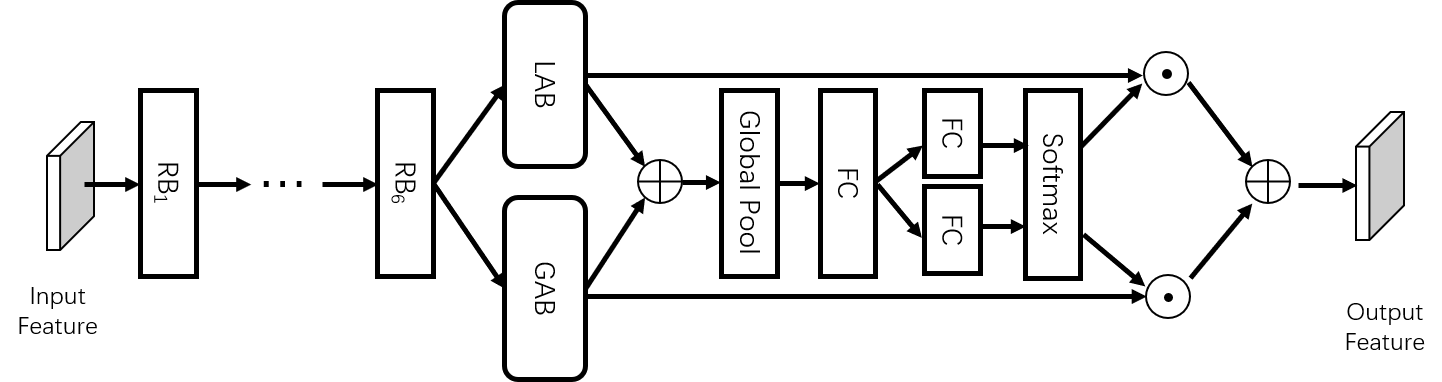


Fig. 2 Architecture of DFB.

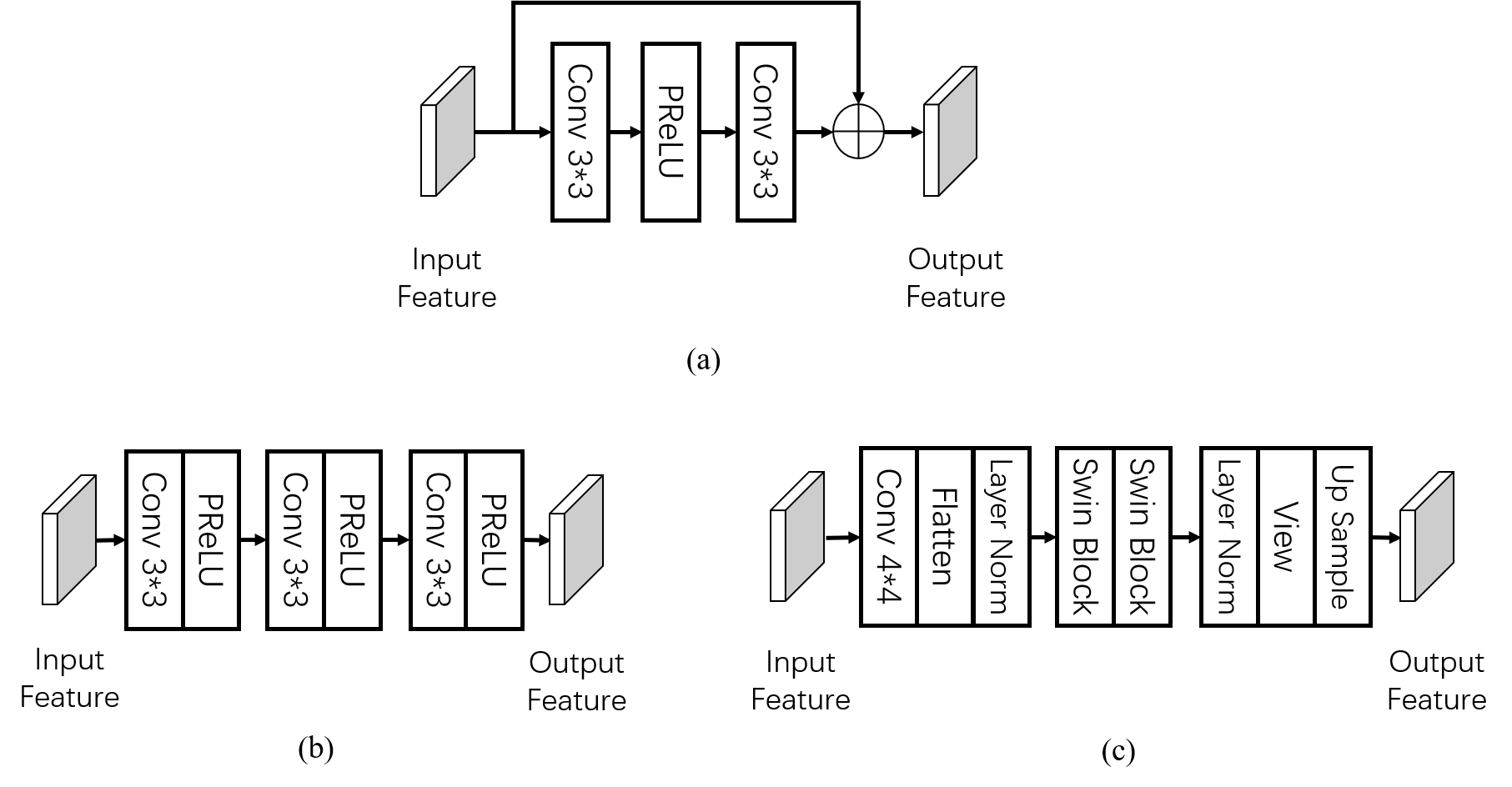


Fig. 3 (a) Architecture of RB. (b) Architecture of LAB. (c) Architecture of GAB.

The architecture of the proposed residual block (RB) is shown in Fig. 3(a). Each RB has two 3×3 convolutional layers and a Parametric Rectified Linear Unit (PReLU) as activation function between these two layers. Residual connection is also applied to RB. LAB consists of three cascaded 3×3 convolutional layer and PReLU activation function. GAB consists of two shift window transformer blocks [1] in the middle. In order to reshape the input feature to be accessible by Swin Block, GAB uses a 4×4 convolutional layer with same stride to embed input feature, then it flattens width and height into one dimension. In addition, a layer normalization function is followed. After the process of Swin Block, reverse operations have to be applied, which contain layer normalization function, viewing operation to recover the dimension and pixel shuffle operation as up sample function.

Decoding time >30x VTM (on GPU vs. CPU)

As loss function, weighted Charbonnier loss was used (similar to L1 norm, with 10:1:1 weight Y:U:V)

Proposal for information.

[JVET-Z0144](https://jvet-experts.org/doc_end_user/current_document.php?id=11593) AHG11: CNN-Based Post-Processing Filter for Video Compression with Multi-Scale Feature Representation [Z. Qi, C. Jung (Xidian Univ.), Y. Liu, M. Li (OPPO)]

This contribution proposes a convolutional neural network (CNN)-based post-processing filter for video compression. The proposed filter combines the discrete wavelet transform (DWT) with CNN to utilize its multi-scale and multi-directional feature representation in enhancement. It consists of two sub-networks: Step-like sub-band network (SLSB) and mixed enhancement network (ME). SLSB includes a Res2Net Group (R2NG) composed of Res2Net modules to represent multiscale features. ME uses dilated convolution and standard convolution to expand the receptive field without blind spots. Experimental results demonstrate that the proposed CNN filter achieves average 2.59%, 3.73%, and 3.34% BD-rate reductions over VTM 11.0-NNVC anchor for Y channel in RA, AI, and LDP configurations (A1, A2, B, C, D and E classes in CTC), respectively.The whole architecture of the proposed CNN post-processing filter is shown in Fig. 1. The network backbone has two sub-networks: SLSB network and ME network. First, the video frame to be processed is divided into four wavelet subbands after being decomposed by DWT. Then, the wavelet subbands are sent to the step-like subband Network. According to the relationship between wavelet subbands, the network performs step-by-step restoration of wavelet subbands from high frequency to low frequency, and from low information to high information. Subsequently, the wavelet subbands are restored to the output frame of the first-stage sub-network through IDWT, and then sent to the ME network. The ME network uses mixed convolution composed of dilated convolution and standard convolution as the basic block for enhancement, which can effectively expand the receptive field of the network and enhance the network's ability of learning mapping. Finally, the output of the ME network is the reconstructed video frame.

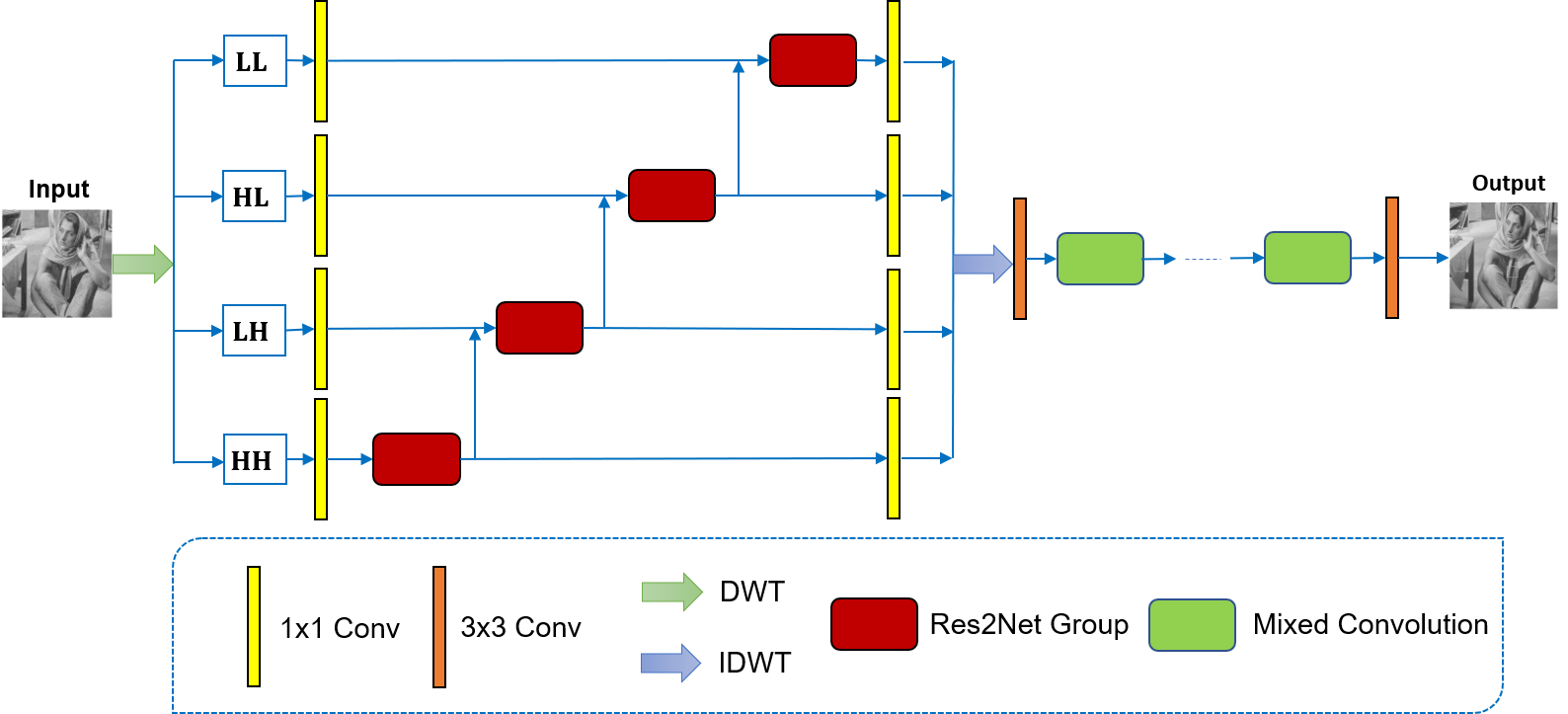


Fig. 1 Whole architecture of the proposed network for Post-Processing of video compression.

After the input video frame is decomposed by DWT, SLSB inputs each wavelet sub-band through 1x1 convolutional layers and step-by-step to the step-like sub-band network. SLSB processes the wavelet sub-bands one-by-one from the highest frequency to the lowest frequency and uses the processed wavelet sub-bands to aid in recovery of the wavelet sub-bands accordingly. Each wavelet sub-band corresponds to a Res2NetGroup in SLSB. Since it can support the multi-scale representation ability of CNN at a finer-grained level, it is used as the basic module. The Res2Net group (R2NG) is further illustrated in Fig. 2.

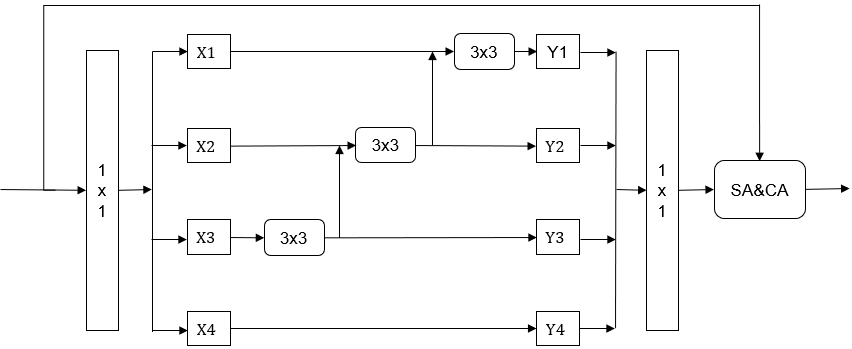


Fig. 2 Network structure of the Res2Net group (R2NG). , represent the input and output of this group, respectively; 1x1 and 3x3 represent convolutional layers with convolution kernels of 1 and 3, respectively; and SA and CA represent the spatial attention and channel attention modules, respectively [2].

Each sub-band corresponds to one Res2Net Group (R2NG), which consists of three Res2Net [3] modules as the backbone, as shown in Fig. 2, with 3x3 convolutional layers added at the beginning and end. In R2NG, residual connections are used. Compared with the traditional bottleneck block, the Res2Net module has stronger multi-scale feature extraction ability with similar computational load.

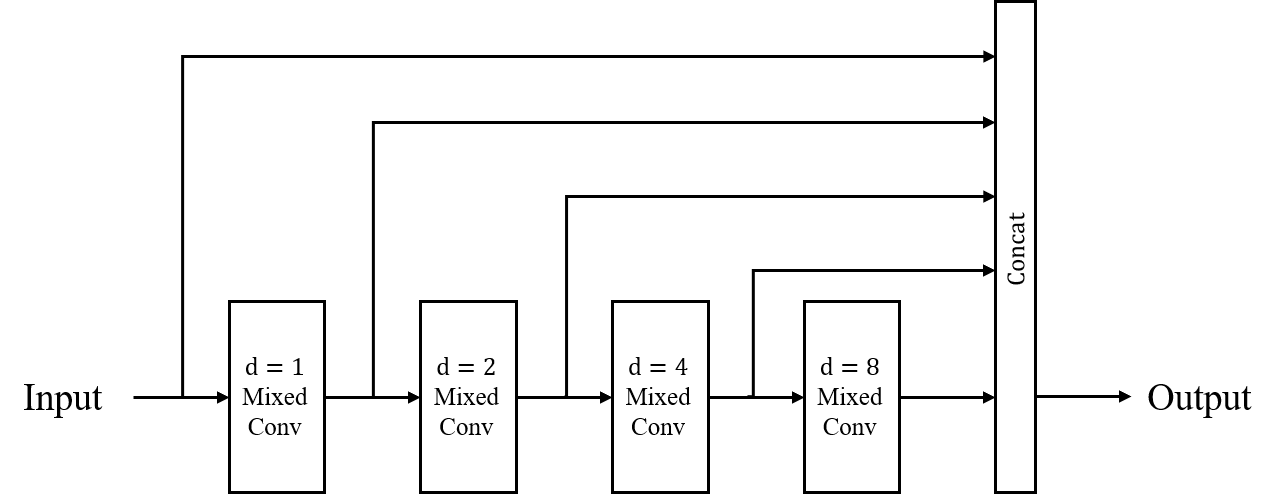


Fig. 3 Structure of the Mixed Enhancement network. d represents the dilated coefficient, and the ME network is formed by weakly densely connected mixed convolutions composed of common convolution and dilated convolutions of d=1, 2, 4 and 8.

Main idea for ME is to expand the size of the receptive field. ME uses mixed convolution [4] composed of dilated convolution and standard convolution as the basic block. Dilated convolution expands the receptive field without increasing the amounts of parameters and reducing the resolution. We densely connect a set of mixed convolutions with dilated coefficients of 1, 2, 4, and 8 into a Mixed Conv group, as shown in Fig. 3. The ME network is built with three Mixed Conv groups and two 3x3 convolutional layers. Since the NN coefficients are hard-coded in the decoder, the coefficients are not needed to be signalled in the bitstream.

The anchor was different from the official EE1 anchor. Chroma results are missing. kMAC/pixel are missing. These aspects should be updated.

After update, decide if the method is worthwhile for investigation in EE, in comparison to benefit of other post processing methods.

## AHG12: Enhanced compression beyond VVC capability (56)

### Summary and BoG reports

[JVET-Z0024](https://jvet-experts.org/doc_end_user/current_document.php?id=11655) EE2: Summary Report on Enhanced Compression beyond VVC capability [V. Seregin, J. Chen, L. Li, K. Naser, J. Ström, M. Winken, X. Xiu, K. Zhang]

Was presented in session 3 at 2100–2300 UTC on Wednesday 20 April 2022 (chaired by JRO), and in session 6 at 2100–23 UTC on Thursday 21 April 2022, and session 9 2100-2210 on Friday 22 April 2022 (chaired by JRO).

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, and entropy coding tests.

The software basis for this EE is ECM-4.0, released at <https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tags/ECM-4.0>. ECM-4.0 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-y-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-y-ee2/simulation-results> if cross-check reports are not submitted as they are optional for EE tests.

1. **List of tests**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Tests** | **Tester** | **Cross-checker** |
| **1 Intra prediction** | | | |
| 1.1 | Slope adjustment for CCLM | Nokia  J. Lainema  [JVET-Z0049](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0049-v1.zip) | OPPO  Z. Xie  [JVET-Z0174](https://jvet-experts.org/doc_end_user/current_document.php?id=11624) |
| 1.2a | DIMD chroma mode | Alibaba  X. Li  [JVET-Z0051](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0051-v1.zip) | InterDigital  K. Naser  [JVET-Z0185](https://jvet-experts.org/doc_end_user/current_document.php?id=11635) |
| 1.2b | Fusion of chroma intra prediction modes | Alibaba  X. Li  [JVET-Z0051](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0051-v1.zip) | InterDigital  K. Naser  [JVET-Z0185](https://jvet-experts.org/doc_end_user/current_document.php?id=11635) |
| 1.2c | Test 1.2a + Test 1.2b | Alibaba  X. Li  [JVET-Z0051](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0051-v1.zip) | InterDigital  K. Naser  [JVET-Z0185](https://jvet-experts.org/doc_end_user/current_document.php?id=11635) |
| 1.2d | Test 1.2a with reduced processing + Test 1.2b | Alibaba  X. Li  [JVET-Z0051](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0051-v1.zip) | InterDigital  K. Naser  [JVET-Z0185](https://jvet-experts.org/doc_end_user/current_document.php?id=11635) |
| 1.3a | Combination of Test 1.1 and Test 1.2c | Nokia  J. Lainema  Alibaba  X. Li  [JVET-Z0050](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0050-v1.zip) | Qualcomm  Y.-J. Chang  [JVET-Z0167](https://jvet-experts.org/doc_end_user/current_document.php?id=11617) |
| 1.3b | Combination of Test 1.1 and Test 1.2d | Nokia  J. Lainema  Alibaba  X. Li  [JVET-Z0050](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0050-v1.zip) | Qualcomm  Y.-J. Chang  [JVET-Z0167](https://jvet-experts.org/doc_end_user/current_document.php?id=11617) |
| **2 Inter prediction** | | | |
| 2.1a | Extended active reference pictures | Qualcomm  H. Huang  [JVET-Z0054](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0054-v1.zip) | Alibaba  R.-L. Liao |
| 2.1b | Block-level reference picture reordering | Qualcomm  H. Huang  [JVET-Z0054](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0054-v1.zip) | Alibaba  R.-L. Liao |
| 2.1c | Test 2.1a + Test 2.1b | Qualcomm  H. Huang  [JVET-Z0054](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0054-v1.zip) | Alibaba  R.-L. Liao |
| 2.2a | Enhanced bi-directional motion compensation | Kwai  Y.-W. Chen  [JVET-Z0136](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0136-v1.zip) | Tencent  G. Li  InterDigital  A. Robert  [JVET-Z0208](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0208-v1.zip) |
| 2.2b | Test 2.2a with BDOF modifications | Qualcomm  Z. Zhang  Kwai  Y.-W. Chen  [JVET-Z0136](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0136-v1.zip) | Tencent  G. Li |
| 2.2c | Test 2.2b with discarded BDOF offsets | Kwai  Y.-W. Chen  Qualcomm  Z. Zhang | withdrawn |
| 2.3 | Template matching based OBMC | vivo  Z. Lv  [JVET-Z0061](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0061-v1.zip) | Qualcomm  Y.-J. Chang  [JVET-Z0166](https://jvet-experts.org/doc_end_user/current_document.php?id=11615) |
| 2.4 | Template matching based reordering for GPM split modes | Qualcomm  C.-C. Chen  [JVET-Z0056](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0056-v1.zip) | KDDI  Y. Kidani |
| 2.5a | ARMC with refined motion | Bytedance  Y. Wang  [JVET-Z0058](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0058-v1.zip) | Wuhan Univ.  G. Ren  [JVET-Z0163](https://jvet-experts.org/doc_end_user/current_document.php?id=11612) |
| 2.5b | Test 2.5a with reduced complexity | Bytedance  Y. Wang  [JVET-Z0058](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0058-v1.zip) | Wuhan Univ.  G. Ren  [JVET-Z0163](https://jvet-experts.org/doc_end_user/current_document.php?id=11612) |
| 2.6a | 12-tap interpolation filter for chroma | Bytedance  X. Xie | withdrawn |
| 2.6b | 8-tap interpolation filter for chroma | Bytedance  X. Xie | withdrawn |
| 2.6c | 6-tap interpolation filter for chroma | Bytedance  X. Xie  [JVET-Z0117](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0117-v1.zip) | Kwai  X. Xiu |
| 2.7a | History based affine inheritance without tables stored in the line buffer | Bytedance  K. Zhang  [JVET-Z0139](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0139-v1.zip) | Xiaomi  P. Andrivon |
| 2.7b | History based affine inheritance with tables stored in the line buffer | Bytedance  K. Zhang  [JVET-Z0139](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0139-v1.zip) | Xiaomi  P. Andrivon |
| 2.7c | Test 2.7a + affine candidates derived from non-adjacent blocks | Bytedance  K. Zhang  [JVET-Z0139](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0139-v1.zip) | Tencent  L.-F. Chen |
| 2.7d | Test 2.7b + affine candidates derived from non-adjacent blocks | Bytedance  K. Zhang | withdrawn |
| 2.7e | Test 2.7 with template matching disabled | Bytedance  K. Zhang | withdrawn |
| 2.8a | Non-adjacent affine model derivation without constrained memory usage | Kwai  W. Chen  [JVET-Z0139](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0139-v1.zip) | Qualcomm  Y. Zhang |
| 2.8b | Non-adjacent affine model derivation with constrained memory usage | Kwai  W. Chen  [JVET-Z0139](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0139-v1.zip) | Qualcomm  Y. Zhang |
| 2.8c | Test 2.8 with template matching disabled | Kwai  W. Chen | withdrawn |
| 2.9a | Test 2.8a + Test 2.7 | Kwai  W. Chen | withdrawn |
| 2.9b | Test 2.8b + Test 2.7 | Kwai  W. Chen  [JVET-Z0139](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0139-v1.zip) | Tencent  L.-F. Chen |
| **3 Screen content coding** | | | |
| 3.1 | Cross-component palette coding | Bytedance  B. Vishwanath  [JVET-Z0055](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0055-v1.zip) | Nokia  J. Lainema |
| 3.2 | IBC with extended reference area | Bytedance  J. Xu  [JVET-Z0153](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0153-v1.zip) | Ofinno  D. Ruiz Coll |
| 3.3 | Enlarged HMVP table for IBC | Bytedance  N. Zhang  [JVET-Z0075](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0075-v2.zip) | Ofinno  D. Ruiz Coll |
| 3.4 | IBC with template matching | InterDigital  A. Robert  [JVET-Z0084](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0084-v1.zip) | Ofinno  D. Ruiz Coll |
| 3.5a | Invalid BVP candidate adjustment based on IBC reference region | Ofinno  D. Ruiz Coll  [JVET-Z0160](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0160-v1.zip) |  |
| 3.5b | Replacement of zero-vector candidates in IBC Merge/AMVP list | Ofinno  D. Ruiz Coll  [JVET-Z0160](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0160-v1.zip) |  |
| 3.5c | Test 3.5a + Test 3.5b | Ofinno  D. Ruiz Coll  [JVET-Z0160](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0160-v1.zip) |  |
| 3.6a | Test 3.2 + Test3.3 | Bytedance  J. Xu, N. Zhang  [JVET-Z0165](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0165-v2.zip) | Ofinno  D. Ruiz Coll |
| 3.6b | Test 3.2 + Test 3.4 | Bytedance  J. Xu  InterDigital  A. Robert  [JVET-Z0095](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0095-v1.zip)  [JVET-Z0165](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0165-v2.zip) | Ofinno  D. Ruiz Coll |
| 3.6c | Test 3.2 + Test 3.5c | Bytedance  J. Xu  Ofinno  D. Ruiz Coll  [JVET-Z0165](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0165-v2.zip) |  |
| 3.6e | Test 3.3 + Test 3.4 | Bytedance  N. Zhang  InterDigital  A. Robert  [JVET-Z0095](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0095-v1.zip)  [JVET-Z0165](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0165-v2.zip) | Ofinno  D. Ruiz Coll |
| 3.6f | Test 3.3 + Test 3.5c | Bytedance  N. Zhang  Ofinno  D. Ruiz Coll  [JVET-Z0165](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0165-v2.zip) | Alibaba  R.-L. Liao |
| 3.6g | Test 3.4 + Test 3.5c | InterDigital  A.Robert  Ofinno  D. Ruiz Coll  [JVET-Z0095](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0095-v1.zip)  [JVET-Z0165](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0165-v2.zip) |  |
| 3.6h | Test 3.2 + Test 3.3 + Test 3.4 | Bytedance  J. Xu, N. Zhang  InterDigital  A. Robert  [JVET-Z0095](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0095-v1.zip)  [JVET-Z0165](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0165-v2.zip) | Ofinno  D. Ruiz Coll |
| 3.6i | Test 3.2 + Test 3.3 + Test 3.5c | Bytedance  N. Zhang  Ofinno  D. Ruiz Coll  [JVET-Z0165](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0165-v2.zip) | InterDigital  A. Robert  [JVET-Z0207](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0207-v1.zip) |
| 3.6j | Test 3.2 + Test 3.4 + Test3.5c | Bytedance  J. Xu, N. Zhang  InterDigital  A.Robert  Ofinno  D. Ruiz Coll  [JVET-Z0165](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0165-v2.zip) |  |
| 3.6k | Test 3.3 + Test 3.4 + Test 3.5c | Bytedance  J. Xu, N. Zhang  InterDigital  A.Robert  Ofinno  D. Ruiz Coll  [JVET-Z0165](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0165-v2.zip) |  |
| 3.6l | Test 3.2 + Test 3.3 + Test 3.4 + Test 3.5c | Bytedance  J. Xu, N. Zhang  InterDigital  A.Robert  Ofinno  D. Ruiz Coll  [JVET-Z0165](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0165-v2.zip) | Alibaba  R.-L. Liao |
| **4 Entropy coding** | | | |
| 4.1a | Probability estimation with adaptive weights | Kwai  X. Xiu  [JVET-Z0134](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0134-v1.zip) | Qualcomm  V. Seregin  Alibaba  J. Chen |
| 4.1b | Test 4.1a + inherited context initialization | Kwai  X. Xiu  [JVET-Z0134](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0134-v1.zip) | Qualcomm  V. Seregin  Alibaba  J. Chen |
| 4.2a | Temporal CABAC initialization | Qualcomm  V. Seregin  [JVET-Z0133](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0133-v1.zip) | Kwai  X. Xiu |
| 4.2b | Adaptive window size adjustment | Qualcomm  V. Seregin  [JVET-Z0133](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0133-v1.zip) | Kwai  X. Xiu |
| 4.2c | Test 4.2a + Test 4.2b | Qualcomm  V. Seregin  [JVET-Z0133](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0133-v1.zip) | Kwai  X. Xiu  OPPO  Kazushi Sato  [JVET-Z0173](https://jvet-experts.org/doc_end_user/current_document.php?id=11623) |
| 4.3a | Test 4.1a + Test 4.2b | Qualcomm  V. Seregin  Kwai  X. Xiu  [JVET-Z0135](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0135-v1.zip) | Xiaomi  P. Andrivon  LGE  J. Zhao  Alibaba  J. Chen |
| 4.3a | Test 4.1a + Test 4.2b + Test 4.2a | Qualcomm  V. Seregin  Kwai  X. Xiu  [JVET-Z0135](https://jvet-experts.org/doc_end_user/documents/26_Teleconference/wg11/JVET-Z0135-v1.zip) | Xiaomi  P. Andrivon  LGE  J. Zhao  Alibaba  J. Chen |

***Intra prediction***

**Test 1.1: Slope adjustment for CCLM**

CCLM uses a model with 2 parameters to map luma values to chroma values. The slope parameter “a” and the bias parameter “b” define the mapping as follows:

chromaVal = a \* lumaVal + b

It is proposed signal an adjustment “u” to the slope parameter to update the model to the following form:

chromaVal = a’ \* lumaVal + b’,

where

a’ = a + u

b’ = b – u \* yr

As a result, the mapping function is rotated around the point with luminance value yr chosen to be the average of the reference luma samples as shown in Figure 1.

Slope adjustment integer parameter *u* ranges between -4 and 4, inclusive, and signaled in the bitstream, the unit of the parameter is 1/8th of a chroma sample value per one luma sample value (for 10-bit content).



**Figure 1. Current (left) and proposed CCLM (right)**

**Test 1.2: DIMD chroma mode and fusion of intra modes**

In Test 1.2a, DIMD chroma mode is tested, chroma prediction of the current block is based on the collocated reconstructed luma samples where horizontal and vertical gradients are calculated for each collocated reconstructed luma sample of the current chroma block to build a HoG as shown in Figure 2.



**Figure 2. Collocated reconstructed luma samples for DIMD chroma mode**

When the intra prediction mode derived from the DIMD chroma mode is the same as the intra prediction mode derived from the DM mode, the intra prediction mode with the second largest histogram amplitude value is used as the DIMD chroma mode. A CU level flag is signaled to indicate whether the proposed DIMD chroma mode is applied.

In Test 1.2b, DM mode and the four default modes can be fused with the MMLM\_LT mode as follows:

where is the predictor obtained by applying the non-LM mode, is the predictor obtained by applying the MMLM\_LT mode and is the final predictor of the current chroma block. The two weights, and are determined by the intra prediction mode of adjacent chroma blocks as when the above and left adjacent blocks are both coded with LM modes, {}={1, 3}; when the above and left adjacent blocks are both coded with non-LM modes, {}={3, 1}; otherwise, {}={2, 2} , and is set equal to 2.

For non-LM mode, a flag is signaled to indicate whether the fusion is applied, the fusion is only applied in I slices.

In Test 1.2c the DIMD chroma mode and the fusion of chroma intra prediction modes are combined where DIMD chroma mode is applied, and for I slices, the DM mode, the four default modes and the DIMD chroma mode can be fused with the MMLM\_LT mode, while for non-I slices, only the DIMD chroma mode can be fused with the MMLM\_LT mode using equal weights.

In Test 1.2d, the same combination as in Test 1.2c is tested while DIMD chroma mode is derived based on the neighboring reconstructed Y, Cb and Cr samples in the second neighboring row and column as shown in Figure 3, instead of using collocated reconstructed luma samples.



**Figure 3. Reconstructed neighbor sample used in DIMD chroma mode**

**Test 1.3: Combination of intra tests**

Test 1.3a is a combination of Test 1.1 and Test 1.2c.

Test 1.3b is a combination of Test 1.1 and Test 1.2d.



Several experts expressed opinion that the proposed methods are straightforward and in particular the combination tests 1.3 indicate small (but in terms of intra coding attractive) gain. 1.3b is even more simplified (not using collocated luma) without reducing the performance significantly.

Decision: Adopt JVET-Z0050 (Test 1.3b)

***Inter prediction***

**Test 2.1: Block-level reference picture reordering**

In Test 2.1a, the number of active reference pictures in random access configurations is increased to the number of available reference pictures at encoder, and encoder fast algorithm is applied:

1. The reference picture with the index larger than 1 in the affine motion estimation process is skipped if it is not selected in the regular AMVP mode.
2. The reference picture with the index larger than 1 is skipped if the ME cost for the reference index 1 is larger than 1.25 times the cost of the reference index 0.

In Test 2.1b, block level reference pictures reordering based on template matching is tested. For the uni-prediction AMVP mode, the reference pictures in List 0 and List 1 are interweaved to generate a joint list. For each hypothesis of the reference picture in the joint list template matching is performed to calculate the cost. The joint list is reordered based on ascending order of the template matching cost. The index of the selected reference picture in the reordered joint list is signaled in the bitstream. For the bi-prediction AMVP mode, a list of pairs of reference pictures from List 0 and List 1 is generated and similarly reordered based on the template matching cost. The index of the selected pair is signaled in the bitstream.

In Test 2.1c, the block level reference picture reordering method is tested with the increased number of active reference pictures (Test 2.1a + Test 2.1b).

**Test 2.2: Enhanced bi-directional motion compensation**

In Test 2.2a, one uni-directional MC sample is regarded as out-of-boundary (OOB) when one of its reference samples is located outside the reference picture beyond half sample. For each prediction sample in a bi-directional block, when one of its uni-directional prediction samples is OOB and the other one is non-OOB, the corresponding bi-directional sample is set equal to the non-OOB sample instead of weighted average of the two uni-directional prediction samples.

BDOF gradients are calculated using the original uni-prediction before OOB processing, and BDOF offset is added to unmodified by OOB handling samples, other BDOF offsets are discarded.

In Test 2.2b, the OOB checking is conducted prior to the BDOF process, where the OOB sample in one uni-directional prediction block is replaced by its collocated prediction sample from the other direction if the collocated sample is non-OOB. In this test, BDOF process is not modified.

**Test 2.3: Template matching based OBMC**

Template matching based OBMC scheme is tested. In the method, sample blending operation is processed for each 4x4 top and left blocks at CU boundaries, and three template matching costs Cost1, Cost2, Cost3 are computed as SAD between the reconstructed samples of a template shown in Figure 4 and its corresponding reference samples derived by MC process according to the following three types of motion information described for block A as an example:

Cost1 is calculated according to A motion information.

Cost2 is calculated using the above neighbor A motion information.

Cost3 is calculated as a weighted prediction using Aandthe above neighbor A motion information with the weighting factors equal to 3/4 and 1/4, respectively.



**Figure 4. OBMC template**

The final predictor is done based on the Cost1, Cost2, and Cost3 comparison.

The original MC result using current block’s motion information is denoted as s1, and the MC result using neighboring block’s motion information is denoted as s2. The final prediction result is denoted as s.

* If Cost1is the smallest, then s(i,j)= s1(i,j).
* If (Cost2 + (Cost2 >> 2) + (Cost2 >> 3)) <= Cost1, then blending mode 1 is used.

For luma blocks, the number of blending sample rows is 4.

s(i,0) = (26×s1(i,0) + 6×s2(i,0) + 16)≫5

s(i,1) = (7×s1(i,1) + s2(i,1) + 4)≫3

s(i,2) = (15×s1(i,2) + s2(i,2) + 8)≫4

s(i,3) = (31×s1(i,3) + s2(i,3) + 16)≫5

For chroma blocks, the number of blending sample rows is 1.

s(i,0) = (26×s1(i,0) + 6×s2(i,0) + 16)≫5

* If Cost1 <= Cost2, then blending mode 2 is used.

For luma blocks, the number of blending sample rows is 2.

s(i,0) = (15×s1(i,0) + s2(i,0) + 8)≫4

s(i,1) = (31×s1(i,1) + s2(i,1) + 16)≫5

For chroma blocks, the number of blending sample rows/columns is 1.

s(i,0) = (15×s1(i,0) + s2(i,0) + 8)≫4

* Otherwise, blending mode 3 is used.

For luma blocks, the number of blending sample rows is 4.

s(i,1) = (7×s1(i,1) + s2(i,1) + 4)≫3

s(i,2) = (15×s1(i,2) + s2(i,2) + 8)≫4

s(i,3) = (31×s1(i,3) + s2(i,3) + 16)≫5

For chroma blocks, the number of blending sample rows is 1.

s(i,0) = (7×s1(i,0) + s2(i,0) + 4)≫3

**Test 2.4: Template matching based reordering for GPM split modes**

The reordering method for GPM split modes is a two-step process performed after the respective reference templates of the two GPM partitions in a coding unit are generated, as follows:

* GPM partition edge is extended into the reference templates of the two GPM partitions, resulting in 64 reference templates, and computing the TM cost for each reference template
* reordering GPM split modes based on their TM cost values in ascending order and marking the best 32 as available split modes.

The edge on the template is extended from the CU, as Figure 5 illustrates, but GPM blending is not used in the templates across the edge.

Diagram

Description automatically generated

**Figure 5. GPM templates and partition edge**

After ascending reordering using TM cost, an index is signaled using Golomb-Rice code to indicate the use of GPM split mode.

**Test 2.5: ARMC with refined motion**

In the method, each merge candidate in the merge candidate list is refined using TM/multi-pass DMVR and the refined motion is used in ARMC to reorder the merge candidate list as shown on Figure 6. ARMC with refined motion is applied to template matching merge (TM merge) mode, adaptive DMVR mode, and template matching for advanced motion vector prediction (TM-AMVP) mode. 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.

形状

低可信度描述已自动生成

**Figure 6. AMC with refined motion**

In Test 2.5b, when template matching is used to derive the refined motion, the template size is set equal to 1 and only the first 8 merge candidates are reordered with the refined motion in TM merge mode.

**Test 2.6: Longer interpolation filters for chroma**

6-tap chroma interpolation filters are tested in addition to 4-tap filters used for motion compensation. A flag is signaled in SPS to indicate whether 6-tap or 4-tap chroma interpolation filter is used. The 4-tap chroma interpolation filter is applied for sequences in class A1/A2 while the 6-tap interpolation filter is applied for sequences in other classes.

**Test 2.7: History based affine inheritance**

In the test, affine model can be inherited from a previous affine-coded blocks which may not be neighboring to the current block. The tested method consists of the following aspects:

1. Two history tables for affine inheritance to derive affine merge candidates
2. Pair-wise affine candidates derived from history- and non-history-based affine candidates
3. Increased sub-block merge list size from 5 to 15
4. Regular merge candidates derived from the first affine history table

A history-parameter table is established, each entry stores a set of affine parameters: *a*, *b*, *c* and *d*, each of which is represented by a 16-bit signed integer. Entries are categorized by reference list and reference index, and 5 reference indices are supported for each reference list in the history table.

For each category, at most 7 entries can be stored, resulting in 70 entries totally in the history table. At the beginning of each CTU row, the number of entries for each category is initialized as 0.

A history-affine-parameter-based candidate is derived from a neighbouring 4×4 block denoted as A0, A1, A2, B0, B1, B2 or B3 in Figure 8 and a set of affine parameters stored in a corresponding entry in the history table. The MV of a neighbouring 4×4 block served as the base MV. The MV of the current block at position (*x*, *y*) is calculated as:

,

where (*mvhbase*, *mvvbase*) represents the MV of the neighbouring 4×4 block, (*xbase*, *ybase*) represents the center position of the neighbouring 4×4 block. (*x*, *y*) can be the top-left, top-right and bottom-left corner of the current block to obtain the corner-position MVs (CPMVs) for the current block, or it can be the center of the current block to obtain a regular MV for the current block.

A second history-parameter table with base MV information is appended. There are 9 entries in the second table, wherein an entry comprises a base MV, a reference index, 4 affine parameters for each reference list, and a base position. An additional merge candidate can be generated from the second table with the base MV information and the corresponding affine models stored as an entry. The difference between the first and the second tables is illustrated in Figure **7**.



**Figure 7. Affine candidate derived from the history tables**

Pair-wise affine merge candidates are generated by two affine merge candidates which are history-derived or not history-derived. A pair-wise affine merge candidates is generated by averaging the CPMVs of existing affine merge candidates in the list.

Such candidates are added into the sub-block-based merge candidate list and affine AMVP candidate. The size of sub-block-based merge candidate list is increased from 5 to 15.

Besides, such candidates are used to derive MVs located at the center of the current block, as regular merge candidates.

In Test 2.7b, affine history tables stored in the above and above right CTUs can be used by the blocks of the current CTU.

The requested test with disabled template matching was not provided.

**Test 2.8: Non-adjacent affine model derivation**

Non-adjacent spatial neighbor positions used in the regular merge mode are added for affine merge and affine AMVP modes. The tested method consists of the following aspects:

1. Non-adjacent constructed and non-adjacent inherited affine candidates
2. Affine candidates derived from neighboring and non-adjacent blocks
3. Increased sub-block merge list size from 5 to 15
4. ARMC subgroup size of affine merge is increased from 3 to 15

The motion information of the non-adjacent spatial neighbors is utilized to generate additional inherited and constructed affine merge candidates.

For inherited candidates, the same derivation process of the inherited affine merge candidates in the VVC is kept unchanged except that the CPMVs are inherited from non-adjacent spatial neighbors. The non-adjacent spatial neighbors are checked based on their distances to the current block, i.e., from near to far. At a specific distance, only the first available neighbor (that is coded with the affine mode) from each side (e.g., the left and above) of the current block is included for inherited candidate derivation. As indicated by the red dash arrows in Figure 8 (a), the checking orders of the neighbors on the left and above sides are bottom-to-up and right-to-left, respectively.



**Figure 8. Spatial neighbors for deriving affine merge candidates: (a) for deriving inherited affine merge candidates (b) for deriving constructed affine merge candidates**

For constructed candidates, as shown in the Figure 8 (b), the positions of one left and above non-adjacent spatial neighbors are firstly determined independently. After that, the location of the top-left neighbor can be determined accordingly which can enclose a rectangular virtual block together with the left and above non-adjacent neighbors. Then, as shown in the Figure 7, the motion information of the three non-adjacent neighbors is used to form the CPMVs at the top-left (A), top-right (B) and bottom-left (C) of the virtual block, which is finally projected to the current CU to generate the corresponding constructed candidates.



**Figure 9. From non-adjacent neighbors to constructed affine merge candidates**

One second type of constructed affine candidates that are derived from non-adjacent neighbors are introduced for affine merge mode. Those new constructed affine candidates are derived based on the same affine candidate construction scheme as in Test 2.7a. However, instead of using history-based look-up table, the non-translational affine parameters are inherited from the non-adjacent spatial neighbors. Specifically, the second type of affine constructed candidates are generated from the combination of 1) the translational affine parameters of adjacent neighboring 4x4 blocks; and 2) the non-translational affine parameters {a, b, c, d} inherited from the non-adjacent spatial neighbors as defined in Figure 8 (a).

Due to the inclusion of the additional non-adjacent candidates, the size of the affine merge candidate list is increased from 5 to 15. The subgroup size of the affine merge mode for the adaptive reordering method is increased from 3 to 15.

In Test 2.8b, below changes are applied to constrain the memory usage:

* The area from where the non-adjacent neighbors come is restricted to be within the current CTU (i.e., no additional storage requirements for line buffer).
* The storage granularity for affine motion information, including CPMVs and reference indexes, is reduced from 8x8 to 16x16 (i.e., only the affine motion from the top-left 8x8 block is saved). Additionally, the saved CPMVs are projected to each 16x16 block before storage, such that the position and size information are not needed.
* Only the top-left and top-right CPMVs are stored (i.e., always using 4-parameter affine model).

The requested test with disabled template matching was not provided.

**Test 2.9: Combined history based and non-adjacent affine candidates**

Test 2.7c: Test 2.7a + the first type of constructed non-adjacent affine candidates from Test 2.8a.

The test contains the following aspects from Test 2.7 and Test 2.8:

1. Two history tables for affine inheritance to derive affine merge candidates
2. Pair-wise affine candidates derived from history- and non-history-based affine candidates
3. Non-adjacent constructed affine candidates
4. Increased sub-block merge list size from 5 to 15
5. Regular merge candidates derived from the first affine history table

Test 2.9b: Test 2.7a + the first type of constructed non-adjacent affine candidates from Test 2.8a + the inherited non-adjacent affine candidates with memory constraints from Test 2.8b.

The test contains the following aspects from Test 2.7 and Test 2.8:

1. Two history tables for affine inheritance to derive affine merge candidates
2. Pair-wise affine candidates derived from history- and non-history-based affine candidates
3. Non-adjacent constructed and non-adjacent inherited affine candidates with memory constraint
4. Increased sub-block merge list size from 5 to 15
5. ARMC reordering size is increased from 15 to 45. The best 15 candidates are inserted into the affine merge candidate list, i.e., the size of the affine merge candidate list is kept unchanged.
6. Regular merge candidates derived from the first affine history table

Memory requirements for 4K resolution sequences.

|  |  |
| --- | --- |
| Test | Memory |
| Test 2.7a | 0.8 KB |
| Test 2.7b | 12.5 KB |
| Test 2.8a | 1, 868 KB |
| Test 2.8b | 4.2 KB |
| Test 2.7c | 0.8 KB |
| Test 2.9b | 5 KB |



2.1a is an encoder-only change using in RA mode the maximum possible number of ref pictures that VVC allows. Similar gain is possible with VVC (additional results in JVET-Z0054). Combination 2.1c shows that the gain of 2.1b (block-level reordering based on relative simple template matching without search) is retained.

It is mentioned that JVET-Z0072 proposes another encoder-only change of reference picture usage (for RA and LB) with comparably higher gain, both for ECM and VVC.

2.1b gives 0.1% gain with relatively small impact on complexity, which would also translate into gain over VVC.

Decision: Adopt JVET-Z0054 (Test 2.1b)

2.2 avoids wrong prediction at picture boundary by checking usage of samples beyond picture boundary, and using uni prediction in such cases. Variant b applies this before BDOF, which would require additional memory for copying samples (BDOF itself is unchanged in 2.2b, but modified in 2.2a).

The decoder runtime is increased by 4%, According to the proponents, this is due to the fact that sample-wise checking is applied over the whole picture.

It is also asked why 2.2b improves in LB as well, as BDOF is not used there. It was later explained that both 2.2a and 2.2b also have some modification of affine bi-prediction (avoiding out-of-boundary prediction) which is causing a small gain.

From the discussion, it turns out that the complexity impact of the proposed method is not yet sufficiently understood, and more clarification is needed which of the two methods (2.2a/2.2b) would be the better choice. Proponents should clarify this offline with the experts who raised concerns, and also identify the possibility of reducing decoder run time.

Follow-up discussion in session 11 2145 UTC: After offline consideration, it was confirmed that the method 2.2b has extra memory copy for prediction samples. On the other hand, it also has slightly better performance than 2.2a (in RA, 0.27% vs. 0.22%).

The proponents of 2.2a reported verbally that they have done some code optimization which would reduce the decoding time to 101%.

Due to the more straightforward design and lower complexity 2.2a was assessed to be the preferable method and supported by non-proponent parties.

Decision: Adopt JVET-Z0136 (test EE2 – 2.2a)

2.3 uses template matching to derive the OBMC blending method. Decoder run time is slightly decreased, which is asserted by proponents to be due to the fact that OBMC may not be used in some cases. It is however mentioned that the worst case complexity might likely be increased.

In terms of numbers reported, the benefit of compression (though relatively small) versus encoder/decoder run time is considered attractive by various experts.

Decision: Adopt JVET-Z0061 (test EE2 – 2.3)

2.4 uses template matching for re-ordering the GPM modes (check for all 64 modes, template 1 line/column necessary at the decoder). Relative to the last meeting’s proposal, additional gain was achieved by encoder optimization. The complexity is certainly increased, but this is not noticeable in terms of decoder run time, Results also confirmed by cross checkers.

Decision: Adopt JVET-Z0056 (test EE2 – 2.4)

2.5a and 2.5b (the latter with reduced complexity) do not provide attractive tradeoff performance vs. decoder complexity. Proponents report that they have an EE related proposal (JVET-Z0145) which might be better suitable. It was asked why the method is disabled depending on block size, and depending on QP. No clear rationale was given.

2.6 (chroma interpolation with 6 tap): It was asked why there is switching between 4-tap (for A classes) and 6-tap (for lower resolutions). Opinion was raised that the switching (using a flag) may not be necessary. Additional results also show that without switching the chroma gain is increased to 1.25%. Filters larger than 6-tap are not deemed to be beneficial, as the additional gain is small, complexity further increases, and they are accessing a relatively larger area than luma filters (in 4:2:0).

Filters are based on DCT-IF design.

Though the gain is not large, several experts expressed the opinion that it is a more consistent design aligning the length of the filters for chroma such that they use the same area as the 12-tap luma filters.

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).

Tests 2.7…2.9 are targeting improved performance of affine parameter coding. 2.7a and 2.7c are most attractive in terms of small memory footprint for storing parameters of adjacent blocks. 2.7c has better performance both for RA and LB. It is however noted that 2.7c might require more cycles for list derivation (which is not reflected in encoder/decoder run time). At this stage of exploration, better compression has higher priority than detailed implementation considerations.

Decision: Adopt JVET-Z0139 (version EE2-2.7c)

***Screen content coding***

**Test 3.1: Cross-component palette coding**

In cross-component palette coding, a cross-component look-up table is derived based on reconstructed neighboring samples by storing a luma sample quantized to 8 bits and a corresponding chroma sample value. The look-up table is derived from the top, left, top-left and bottom-right blocks in the current CTU and their extensions by 8 chroma reference lines in the neighboring CTUs.

For each chroma sample, the co-located luma sample value, denoted as Yc is fetched as a key to retrieve a mapping chroma sample value from the table. If Yc cannot be found, Yc ± 1, Yc ± 2 and Yc ± 3 are checked in order as keys to find the mapping chroma sample value. If none of the candidate keys are found, the mapping chroma sample value is set to be the average of the chroma values in the table.

In addition to LM down-sampling filter in ECM, five candidate down-sampling filters corresponding to down-sampling locations Y1, Y2…, Y5 as shown in Figure 10 are introduced to generate the best co-located luma sample value for the YCbCr 4:2:0 color format. The encoder tries all candidate down-sampling filters and signals an index per block of the best filter to the decoder.

Shape, polygon

Description automatically generated

**Figure 10. Luma down-sampling positions**

**Test 3.2: IBC with extended reference area**

In this test, the reference area for IBC is extended to two CTU rows above. Figure 11 illustrates the reference area for coding CTU (m, n).



**Figure 11. Reference area (green) for IBC when CTU (m, n) is coded**

**Test 3.3: Enlarged HMVP table for IBC**

In the test, the HMVP table size for IBC is increased to 25. After up to 20 IBC merge candidates are derived with full pruning, they are reordered together. After reordering, the first 6 candidates with the lowest template matching costs are selected as the final candidates in the IBC merge list.

**Test 3.4: IBC with template matching**

In the test, the IBC-TM merge list has been modified compared to the one used by regular IBC merge mode such that the candidates are selected according to a pruning method with a motion distance between the candidates as in the regular TM merge mode. The zero motion candidates have been replaced by (-W, 0), (0, -H), (-W, -H) MVs.

In the IBC-TM merge mode, the selected candidates are refined with the template matching method, TM-merge flag is signaled to indicate the mode.

In the IBC-TM AMVP mode, up to 3 candidates are selected from the IBC-TM merge list. Each of those 3 selected candidates are refined using the template matching method and sorted according to their resulting TM cost.

TM refinement is performed at integer position, and in IBC-TM AMVP mode, it is performed either at integer or 4-pel precision depending on the AMVR value. The refinement is done within the existed IBC reference area.

**Test 3.5: BVP candidate adjustment based on IBC reference region**

In Test 3.5a, clipping to the nearest IBC buffer boundaries is applied to one or both components BVP candidates pointing outside the IBC reference region before being used in the IBC Merge/AMVP list derivation.

In Test 3.5b, zero MV candidates in IBC merge and AMVP modes are replaced by 3 candidates are located on the nearest corners of the reference region, and 3 additional candidates are determined in the middle of the three sub-regions (A, B, and C), whose coordinates are determined by the width, and height of the current block and the ΔX and ΔY parameters, as is depicted in Figure 12.



**Figure 12. Zero MV candidates replacement**

Test 3.5c is a combination of Test 3.5a and 3.5b.

**Test 3.6: Combination of IBC tests**

The following combinations were tested.

Test 3.6a: extended IBC area (Test 3.2) + enlarged IBC HMVP (Test 3.3)

Test 3.6b: extended IBC area (Test 3.2) + IBC TM (Test 3.4)

Test 3.6c: extended IBC area (Test 3.2) + BV clipping & zero BV replacement (Test 3.5c)

Test 3.6e: enlarged IBC HMVP (Test 3.3) + IBC TM (Test 3.4)

Test 3.6f: enlarged IBC HMVP (Test 3.3) + BV clipping & zero BV replacement (Test 3.5c)

Test 3.6g: BV clipping & zero BV replacement (Test 3.5c) + IBC TM (Test 3.4)

Test 3.6h: extended IBC area (Test 3.2) + enlarged IBC HMVP (Test 3.3) + IBC TM (Test 3.4)

Test 3.6i: extended IBC area (Test 3.2) + IBC TM (Test 3.4) + BV clipping & zero BV replacement (Test 3.5c)

Test 3.6j: extended IBC area (Test 3.2) + IBC TM (Test 3.4) + BV clipping & zero BV replacement (Test 3.5c)

Test 3.6k: enlarged IBC HMVP (Test 3.3) + IBC TM (Test 3.4) + BV clipping & zero BV replacement (Test 3.5c)

Test 3.6l: extended IBC area (Test 3.2) + enlarged IBC HMVP (Test 3.3) + IBC TM (Test 3.4) + BV clipping & zero BV replacement (Test 3.5c)



Test 3.1: “Cross-component palette coding” is rather a new prediction mode using a palette construction from the neighborhood. This is used to improve the prediction of chroma, where a residual is still optionally encoded. The anchor compared against was not using the original palette mode of VVC. However, the proponents report that most of the gain is retained when palette was enabled in additional results.

Gain is relatively low even for the case of screen content. Not worthwhile to introduce a new coding tool which is working specifically for a very special type of content.

It is recommended that in future EE on screen content the original palette mode should be enabled (not CTC at this time).

IBC proposals 3.2 … 3.6: Various combinations available. Most benefit from enlarged reference area (3.2). Further, template matching (3.4) which is also used in other parts of ECM gives additional gain. Replacement of zero displacement candidates is straightforward and also gives some additional gain, where the solution 3.5b is just modifying the merge/AMVP lists. Enlarging the HMVP is similar to what is done in MV coding for inter, simple to implement and also gives gain. Altogether this would be similar result to combination 3.6l, except that 3.5a is not included, as it does not give additional gain to 3.5b (according to 3.5c). Conclusion from session 9 on Friday 22:

Decision: Adopt JVET-Z0153 (3.2), JVET-Z0075 (3.3), JVET-Z0084 (3.4), JVET-Z0160 (3.5b)

***Entropy coding***

**Test 4.1: Improved probability estimation for CABAC**

In test 4.1a, instead of using semi average of the two states p0 and p1, a weighted average is performed

p = (ω0 ∙ p0 + ω1 ∙ p1 )≫s,

where ω0 and ω1 are the weights selected from a pre-defined set {10,12,16,20,22}; *s* is the bitwise right-shift value, which is equal to 5 when (ω0 + ω1 ) ≤ 32 and 6 otherwise.

Three different sets of weights are pre-determined for each context model at I-, B- and P-slice types. It takes 6 bits to store the indices of two weights (3bits for each) for each context. There are 508 contexts in ECM-4.0 and three different sets of weights needed for I-, B- and P-slice types. Therefore, the total storage of maintaining the weights are 6bits × 3 × 508 = 1,143 bytes.

In test 4.1b, context initialization stored at previously coded picture after coding the last CTU can be used to initialize an inter slice having the same slice type and QP. For each inter slice, one control flag is signaled to select the context initialization scheme used for the slice. When the flag is equal to zero, it indicates that the context models of the slice are initialized using one of the existing context initialization tables (as indicated by the flag *sh\_cabac\_init\_flag*). Otherwise, the context models of the slice are initialized by inheriting the ones from the previously coded picture.

**Test 4.2: CABAC initialization from previous inter slice and windows adjustment**

In test 4.2a, context initialization stored at previously coded picture after coding the last CTU can be used to initialize an inter slice having the same slice type, QP, and temporal ID. The buffer size for storing previous initializations is set equal to 5 for each slice type, when the buffer is full the entry with the smallest QP and temporal ID is removed first before storing the initialization.

In Test 4.2b, short and long window sizes used in CABAC update are adjusted by two delta parameters stored in a look-up table per context and retrieved by a previous coded bin used as an index. The actual window sizes used to code the current bin after adjustment is (shift0+delta0) and (shift1+delta1), where shift0 and shift1 are existed predefined windows sizes stored in the context initialization tables. The window adjustment values take 4 bits each, so both values are stored in 1 byte and the total size of the LUT for 508 contexts is 1,016 bytes.

**Test 4.3: Combination of entropy coding tests**

Test 4.3a is a combination of Test 4.1a (adaptive weights) and Test 4.2b (windows adjustment)

Test 4.3b is a combination of Test 4.1a (adaptive weights), Test 4.2b (windows adjustment), and Test 4.2a (temporal initialization)



Note: “inherited context initialization”=”temporal initialization”

According to test 4.3b, all three elements (probability estimation with adaptive weights, temporal initialization, and adaptive window size adjustment) contribute to the overall gain. Most gain comes by temporal initialization, but comparison of 4.1b and 4.3b shows that also the adaptive window size is beneficial. Comparing 4.2c and 4.3b shows the additional gain of 4.1a.

An opinion is raised that the adaptive window size requires a second table which may slow down CABAC. Concern is also raised on the temporal initialization, as this introduces parsing dependency between frames.

It is reported that a maximum of 5 buffers are foreseen for the temporal initialization. It is asked if this would be sufficient for non-CTC condition. It is argued that it would allow to avoid retraining for initialization, which is more specifically tuning to CTC. It is also argued that it has been practically already implemented elsewhere.

A cross-checker points out that he found a deviation relative to the previously proposed method of temporal initialization. This was however introduced for supporting temporal scalability (inheriting from same tid) as requested in the previous meeting.

Many experts expressed support for adopting the combination 4.3b

Decision: Adopt JVET-Z0135, Test 4.3b

[JVET-Z0210](https://jvet-experts.org/doc_end_user/current_document.php?id=11661) BoG report on EE2 related proposals [Y. Ye]

(update from version 5)

This is a report of the BoG on EE2 related proposals. The BoG held meetings at the following times during the 26th JVET meeting:

• April 20th 23:20 – April 21st 01:27

• April 21st 23:20 – April 22nd 01:37

• April 22nd 23:20 – April 23rd 01:49

• April 25th 23:20 – April 26th 01:35

• April 26th 23:20 – April 27th 01:35

The BoG recommends to:

• Decision: Include the following proposal in the next release of ECM: JVET-Z0131, JVET-Z0127 (option 2)

• Investigate the following proposals in the next round of EE2 (specifics and/or variations to be investigated in the EE can be found in the detailed notes): JVET-Z0125, JVET-Z0142, JVET-Z0145, JVET-Z0152, JVET-Z0157, JVET-Z0059, JVET-Z0137, JVET-Z0130, JVET-Z0062, JVET-Z0069, JVET-Z0124, JVET-Z0140, JVET-Z0146, JVET-Z0147, JVET-Z0149

• Discuss JVET-Z0157 (IBC for camera captured content) within the context of CTC in the JVET plenary (see notes in section 4.3)

• Decision (SW): Fix the RPR functionality in the next ECM release with some elements from JVET-Z0062, JVET-Z0067, and JVET-Z0079

• Discuss whether to study luma-only RPR as proposed in JVET-Z0067 within the context of a “444 profile” in JVET main track. This was discussed in session 18. It was clarified that such an approach is not comparable with a traditional 4:4:4 profile which are rather targeting high-end (professional) application. Moreover, switching between 4:4:4 and 4:2:0 chroma formats is currently not possible at sub-sequence level. It could however be beneficial in the context of RPR to only subsample luma, to avoid losses which are often observed in subsampled chroma. This would be a new type of RPR tool which would require definition of a new profile with impact on other coding tools at block level and overall architecture, e.g. motion comp and intra prediction switching between 4:2:0 and 4:4:4 chroma sampling. Further study on design implications is recommended. No study in EE2 at this moment.

Add the individual document discussion from the newest version of the BoG report.

The recommendations above were approved in the JVET plenary (session 15). V. Seregin and other EE coordinators are asked to take action drafting the new EE2 description.

Presented in session 11 at 2100 UTC on Monday 25 April, and in session15 on Wednesday 27 April

### EE2 contributions: Enhanced compression beyond VVC capability (20)

Contributions in this area were discussed in the context of the summary report JVET-Z0024 (see section 5.3.1).

[JVET-Z0049](https://jvet-experts.org/doc_end_user/current_document.php?id=11483) EE2-1.1: Slope adjustment for CCLM [J. Lainema, A. Aminlou, P. Astola, R. G. Youvalari (Nokia)]

[JVET-Z0174](https://jvet-experts.org/doc_end_user/current_document.php?id=11624) Cross-check of JVET-Z0049 (EE2-1.1: Slope adjustment for CCLM) [Z. Xie, L. Xu (OPPO)] [late]

[JVET-Z0050](https://jvet-experts.org/doc_end_user/current_document.php?id=11484) EE2-1.3: Combined intra prediction tests [J. Lainema, A. Aminlou, P. Astola, R. G. Youvalari (Nokia), X. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba)]

[JVET-Z0167](https://jvet-experts.org/doc_end_user/current_document.php?id=11617) Crosscheck of JVET-Z0050 (EE2-1.3: Combined intra prediction tests) [Y.-J. Chang (Qualcomm)] [late]

[JVET-Z0051](https://jvet-experts.org/doc_end_user/current_document.php?id=11485) EE2-1.2: On chroma intra prediction [X. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba)]

[JVET-Z0185](https://jvet-experts.org/doc_end_user/current_document.php?id=11635) Crosscheck of JVET-Z0051 (EE2-1.2: On chroma intra prediction) [K. Naser (InterDigital)] [late]

[JVET-Z0054](https://jvet-experts.org/doc_end_user/current_document.php?id=11488) EE2-2.1: On the extended number of active reference pictures and block level reference picture reordering [H. Huang, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-Z0055](https://jvet-experts.org/doc_end_user/current_document.php?id=11489) EE2-3.1: Cross-component palette coding [B. Vishwanath, K. Zhang, L. Zhang (Bytedance)]

[JVET-Z0182](https://jvet-experts.org/doc_end_user/current_document.php?id=11632) Crosscheck of JVET-Z0055 (EE2-3.1: Cross-component palette coding) [J. Lainema (Nokia) [late]

[JVET-Z0056](https://jvet-experts.org/doc_end_user/current_document.php?id=11491) EE2-2.4: Template matching based reordering for GPM split modes [C.-C. Chen, H. Huang, Y. Zhang, Z. Zhang, Y.-J. Chang, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-Z0222](https://jvet-experts.org/doc_end_user/current_document.php?id=11673) Crosscheck of JVET-Z0056 (EE2-2.4: Template matching based reordering for GPM split modes) [Y. Kidani, K. Kawamura (KDDI)] [late]

[JVET-Z0058](https://jvet-experts.org/doc_end_user/current_document.php?id=11493) EE2-2.5: Adaptive reordering of merge candidates with refined motion [Y. Wang, K. Zhang, N. Zhang, Z. Deng, L. Zhang (Bytedance)]

[JVET-Z0163](https://jvet-experts.org/doc_end_user/current_document.php?id=11612) Crosscheck of JVET-Z0058 (EE2-2.5: Adaptive reordering of merge candidates with refined motion) [G. Ren, Z. Chen (Wuhan Univ.)] [late]

[JVET-Z0061](https://jvet-experts.org/doc_end_user/current_document.php?id=11496) EE2-2.3: Template matching based OBMC [Z. Lv, C. Zhou, J. Zhang (vivo)]

[JVET-Z0166](https://jvet-experts.org/doc_end_user/current_document.php?id=11615) Crosscheck of JVET-Z0061 (EE2-2.3: Template matching based OBMC) [Y.-J. Chang (Qualcomm)] [late]

[JVET-Z0075](https://jvet-experts.org/doc_end_user/current_document.php?id=11510) EE2-3.3: Enlarged HMVP table for IBC [N. Zhang, K. Zhang, L. Zhang (Bytedance)]

[JVET-Z0196](https://jvet-experts.org/doc_end_user/current_document.php?id=11646) Crosscheck of JVET-Z0075 (EE2-3.3: Enlarged HMVP table for IBC) [D. Ruiz Coll (Ofinno)] [late]

[JVET-Z0084](https://jvet-experts.org/doc_end_user/current_document.php?id=11519) EE2-3.4: IBC with Template Matching [A. Robert, K. Naser, T. Poirier, Y. Chen, F. Galpin (InterDigital)]

[JVET-Z0197](https://jvet-experts.org/doc_end_user/current_document.php?id=11647) Crosscheck of JVET-Z0084 (EE2-3.4: IBC with Template Matching) [D. Ruiz Coll (Ofinno)] [late]

[JVET-Z0095](https://jvet-experts.org/doc_end_user/current_document.php?id=11530) EE2-3.6: Combined tests involving EE2-3.4 [A. Robert, K. Naser, T. Poirier, Y. Chen, F. Galpin (InterDigital)]

[JVET-Z0198](https://jvet-experts.org/doc_end_user/current_document.php?id=11648) Crosscheck of JVET-Z0095 (EE2-3.6: Combined tests involving EE2-3.4) and JVET-Z0165 (EE2-3.6: Combination tests of EE2-3.2+EE2-3.3+EE2-3.4+EE2-3.5): tests EE2-3.6a, EE2-3.6b, EE2-3.6e, and EE2-3.6h [D. Ruiz Coll (Ofinno)] [late]

[JVET-Z0117](https://jvet-experts.org/doc_end_user/current_document.php?id=11564) EE2-2.6: Long-tap interpolation filtering on chroma components [X. Xie, K. Zhang, L. Zhang, J. Li, M. Wang, S. Wang (Bytedance)]

[JVET-Z0133](https://jvet-experts.org/doc_end_user/current_document.php?id=11582) EE2-4.2: CABAC initialization from previous inter slice and windows adjustment [V. Seregin, H. Golestani, N. Hu, P. Garus, M. Karczewicz (Qualcomm)]

[JVET-Z0173](https://jvet-experts.org/doc_end_user/current_document.php?id=11623) Cross-check of test 4.2c of JVET-Z0133 (EE2-4.2: CABAC initialization from previous inter slice and windows adjustment) [K. Sato (OPPO)] [late]

[JVET-Z0134](https://jvet-experts.org/doc_end_user/current_document.php?id=11583) EE2-Test4.1: Improved probability estimation for CABAC [X. Xiu, W. Chen, C.-W. Kuo, H.-J. Jhu, N. Yan, X. Wang (Kwai)]

[JVET-Z0135](https://jvet-experts.org/doc_end_user/current_document.php?id=11584) EE2-Test4.3: Combined tests of EE2-4.1 and EE2-4.2 [V. Seregin, H. Golestani, N. Hu, P. Garus, Marta Karczewicz (Qualcomm), X. Xiu, W. Chen, C.-W. Kuo, H.-J. Jhu, N. Yan, X. Wang (Kwai)]

[JVET-Z0181](https://jvet-experts.org/doc_end_user/current_document.php?id=11631) Crosscheck of JVET-Z0135 (EE2-Test4.3: Combined tests of EE2-4.1 and EE2-4.2) [J. Zhao (LGE)] [late]

[JVET-Z0188](https://jvet-experts.org/doc_end_user/current_document.php?id=11638) Cross-check of JVET-Z0135 (EE2-Test4.3 - tests 4.3a/4.3b) [P. Andrivon, F. Le Léannec, M. Radosavljević (Xiaomi)] [late]

[JVET-Z0136](https://jvet-experts.org/doc_end_user/current_document.php?id=11585) EE2-Test2.2: Enhanced bi-directional motion compensation [Y.-W. Chen, X. Xiu, H.-J. Jhu, N. Yan, X. Wang (Kwai), H. Huang, Y-J. Chang, C.-C. Chen, M. Karczewicz, V. Seregin, Y. Zhang, Z. Zhang, (Qualcomm)]

[JVET-Z0201](https://jvet-experts.org/doc_end_user/current_document.php?id=11651) Crosscheck of JVET-Z0136 (EE2-Test2.2: Enhanced bi-directional motion compensation) [G. Li (Tencent)] [late]

[JVET-Z0208](https://jvet-experts.org/doc_end_user/current_document.php?id=11659) Cross-check of JVET-Z0136 (EE2-Test2.2a: Enhanced bi-directional motion compensation) [A. Robert (InterDigital)] [late]

[JVET-Z0139](https://jvet-experts.org/doc_end_user/current_document.php?id=11588) EE2-2.7, 2.8, 2.9: History-parameter-based affine model inheritance and non-adjacent spatial neighbors for affine merge mode [W. Chen, X. Xiu, H.-J. Jhu, C.-W. Kuo, X. Wang (Kwai), K. Zhang, L. Zhang, Z. Deng, N. Zhang, Y. Wang (Bytedance)]

[JVET-Z0189](https://jvet-experts.org/doc_end_user/current_document.php?id=11639) Cross-check of JVET-Z0139 (EE2-2.7 - tests 2.7a/2.7b) [P. Andrivon, F. Le Léannec, M. Radosavljević (Xiaomi)] [late]

[JVET-Z0211](https://jvet-experts.org/doc_end_user/current_document.php?id=11662) Crosscheck of JVET-Z0139 (EE2-2.7, 2.8, 2.9: History-parameter-based affine model inheritance and non-adjacent spatial neighbors for affine merge mode): Tests EE2-2.7c and EE2-2.9b [L.-F. Chen (Tencent)] [late]

[JVET-Z0153](https://jvet-experts.org/doc_end_user/current_document.php?id=11602) EE2-3.2: IBC Reference Area Extension [J. Xu, N. Zhang (ByteDance)]

[JVET-Z0195](https://jvet-experts.org/doc_end_user/current_document.php?id=11645) Crosscheck of JVET-Z0153 (EE2-3.2: IBC Reference Area Extension) [D. Ruiz Coll (Ofinno)] [late]

[JVET-Z0160](https://jvet-experts.org/doc_end_user/current_document.php?id=11609) EE2-3.5: BVP candidate adjustment based on IBC reference region [D. Ruiz Coll, A. Filippov, V. Rufitskiy, T. M. Bae (Ofinno)]

[JVET-Z0221](https://jvet-experts.org/doc_end_user/current_document.php?id=11672) Crosscheck of JVET-Z0160 (EE2-3.5: BVP candidate adjustment based on IBC reference region) [K. Cao (Qualcomm)] [late]

[JVET-Z0165](https://jvet-experts.org/doc_end_user/current_document.php?id=11614) EE2-3.6: Combination tests of EE2-3.2+EE2-3.3+EE2-3.4+EE2-3.5 [N. Zhang, J. Xu, K. Zhang, L. Zhang (Bytedance), A. Robert, K. Naser, T. Poirier, Y. Chen, F. Galpin (InterDigital), D. Ruiz Coll, A. Filippov, V. Rufitskiy (Ofinno)] [late]

[JVET-Z0207](https://jvet-experts.org/doc_end_user/current_document.php?id=11658) Cross-check of JVET-Z0165 (Combination tests of EE2-3.2+EE2-3.3+EE2-3.4+EE2-3.5): test EE2-3.6i [A. Robert (InterDigital)] [late]

Note that a co-author of a contribution should not be a cross-checker.

[JVET-Z0198](https://jvet-experts.org/doc_end_user/current_document.php?id=11648) Crosscheck of JVET-Z0095 (EE2-3.6: Combined tests involving EE2-3.4) and JVET-Z0165 (EE2-3.6: Combination tests of EE2-3.2+EE2-3.3+EE2-3.4+EE2-3.5): tests EE2-3.6a, EE2-3.6b, EE2-3.6e, and EE2-3.6h [D. Ruiz Coll (Ofinno)] [late]

Note that a co-author of a contribution should not be a cross-checker.

### EE2 related contributions (12)

Contributions in this area were discussed in session X at XXXX–XXXX UTC on XXday 2X April 2022 (chaired by JRO).

[JVET-Z0048](https://jvet-experts.org/doc_end_user/current_document.php?id=11482) EE2-related: Modification of LFNST for MIP coded block [J.-Y. Huo, W.-H. Qiao, X. Hao, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.), M. Li, Y. Liu (OPPO)]

[JVET-Z0169](https://jvet-experts.org/doc_end_user/current_document.php?id=11619) Crosscheck of JVET-Z0048 (EE2-related: Modification of LFNST for MIP coded block) [X. Li (Alibaba)] [late]

[JVET-Z0066](https://jvet-experts.org/doc_end_user/current_document.php?id=11501) EE2-related: Template matching using extended MVP candidate list [K. Kim, D. Kim, J.-H. Son, J.-S. Kwak (WILUS)]

[JVET-Z0179](https://jvet-experts.org/doc_end_user/current_document.php?id=11629) Crosscheck of JVET-Z0066 (EE2-related: Template matching using extended MVP candidate list) [C. Zhou (vivo)] [late]

[JVET-Z0068](https://jvet-experts.org/doc_end_user/current_document.php?id=11503) EE2-2.6 related: Longer chroma interpolation filters [K. Andersson, R. Yu (Ericsson)]

[JVET-Z0180](https://jvet-experts.org/doc_end_user/current_document.php?id=11630) Crosscheck of JVET-Z0068 (EE2-2.6 related: Longer chroma interpolation filters) [J. Gan (OPPO)] [late]

[JVET-Z0123](https://jvet-experts.org/doc_end_user/current_document.php?id=11571) EE2-related: On chroma interpolation filters for motion compensation [J. Gan, Y. Yu, H. Yu (OPPO)]

[JVET-Z0187](https://jvet-experts.org/doc_end_user/current_document.php?id=11637) Cross-check of JVET-Z0123 (EE2-related: On chroma interpolation filters for motion compensation) [K. Andersson (Ericsson)] [late]

[JVET-Z0125](https://jvet-experts.org/doc_end_user/current_document.php?id=11573) EE2-related: Regression based affine candidate derivation [Y. Zhang, H. Huang, V. Seregin, M. Coban, M. Karczewicz (Qualcomm)]

[JVET-Z0214](https://jvet-experts.org/doc_end_user/current_document.php?id=11665) Crosscheck of JVET-Z0125 (EE2-related: Regression based affine candidate derivation) [W. Chen (Kwai)] [late]

[JVET-Z0216](https://jvet-experts.org/doc_end_user/current_document.php?id=11667) Crosscheck of JVET-Z0125 (EE2-related: Regression based affine candidate derivation) [R. G. Youvalari, J. Lainema (Nokia)] [late]

[JVET-Z0130](https://jvet-experts.org/doc_end_user/current_document.php?id=11578) EE2-related: Motion compensation boundary padding [Z. Zhang, H. Huang, C.-C. Chen, Y.-J. Chang, V. Seregin, M. Coban, M. Karczewicz (Qualcomm), F. Le Léannec, P. Andrivon, M. Radosavljević, E. Thomas (Xiaomi)]

[JVET-Z0162](https://jvet-experts.org/doc_end_user/current_document.php?id=11611) Crosscheck of JVET-Z0130 (EE2-related: Motion compensation boundary padding) [Y. J. Piao, M. S. Park (Samsung)] [late]

[JVET-Z0190](https://jvet-experts.org/doc_end_user/current_document.php?id=11640) Cross-check of JVET-Z0130 (EE2-related: Motion compensation boundary padding [F. Galpin (InterDigital)] [late]

[JVET-Z0131](https://jvet-experts.org/doc_end_user/current_document.php?id=11579) EE2-related: Block Vector Difference Binarization [K. Cao, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-Z0200](https://jvet-experts.org/doc_end_user/current_document.php?id=11650) Crosscheck of JVET-Z0131 (EE2-related: Block Vector Difference Binarization) [D. Ruiz Coll (Ofinno)] [late]

[JVET-Z0142](https://jvet-experts.org/doc_end_user/current_document.php?id=11591) EE2-related: On MMVD and Affine MMVD Extension [M. Salehifar, Y. He, K. Zhang, N. Zhang, L. Zhang (Bytedance)]

[JVET-Z0183](https://jvet-experts.org/doc_end_user/current_document.php?id=11633) Crosscheck of JVET-Z0142 (EE2-related: On MMVD and Affine MMVD Extension) [J. Chen (Alibaba)] [late]

[JVET-Z0145](https://jvet-experts.org/doc_end_user/current_document.php?id=11594) EE2-2.5 related: More test results for ARMC with refined motion [Y. Wang, K. Zhang, N. Zhang, Z. Deng, L. Zhang (Bytedance)]

[JVET-Z0158](https://jvet-experts.org/doc_end_user/current_document.php?id=11607) Cross-check of JVET-Z0145 EE2-2.5 related: More test results for ARMC with refined motion [Z. Lv (vivo)] [late]

[JVET-Z0225](https://jvet-experts.org/doc_end_user/current_document.php?id=11676) Crosscheck of JVET-Z0145 (EE2-2.5 related: More test results for ARMC with refined motion) [J. Chen (Alibaba)] [late]

[JVET-Z0226](https://jvet-experts.org/doc_end_user/current_document.php?id=11677) Crosscheck of JVET-Z0145 (EE2-2.5 related: More test results for ARMC with refined motion) [X. Xiu (Kwai)] [late]

[JVET-Z0152](https://jvet-experts.org/doc_end_user/current_document.php?id=11601) EE2-related: IBC Merge Mode with Block Vector Differences N. Zhang, J. Xu, K. Zhang, M. Salehifar, L. Zhang (Bytedance)]

[JVET-Z0157](https://jvet-experts.org/doc_end_user/current_document.php?id=11606) EE2-3.2-related: IBC Adaptation for Camera-Captured Contents [J. Xu (Bytedance)]

[JVET-Z0199](https://jvet-experts.org/doc_end_user/current_document.php?id=11649) Crosscheck of JVET-Z0157 (EE2-3.2-related: IBC Adaptation for Camera-Captured Contents) [D. Ruiz Coll (Ofinno)] [late]

[JVET-Z0184](https://jvet-experts.org/doc_end_user/current_document.php?id=11634) EE2-3.5-related: Additional results of zero-vector candidates replacement in the IBC Merge/AMVP list [D. Ruiz Coll, A. Filippov, V. Rufitskiy (Ofinno)] [late]

[JVET-Z0233](https://jvet-experts.org/doc_end_user/current_document.php?id=11684) Crosscheck of JVET-Z0184 (EE2-3.5-related: Additional results of zero-vector candidates replacement in the IBC Merge/AMVP list) [J. Xu (Bytedance)] [late]

### ECM modifications beyond EE2 (24)

Contributions in this area were discussed in session X at XXXX–XXXX UTC on XXday 2X April 2022 (chaired by JRO).

[JVET-Z0059](https://jvet-experts.org/doc_end_user/current_document.php?id=11494) Non-EE2: Adaptive width for GPM blending area [Y. Kidani, H. Kato, K. Unno, K. Kawamura (KDDI)]

[JVET-Z0191](https://jvet-experts.org/doc_end_user/current_document.php?id=11641) Crosscheck of JVET-Z0059 (Non-EE2: Adaptive width for GPM blending area) [C.-C. Chen (Qualcomm)] [late]

[JVET-Z0204](https://jvet-experts.org/doc_end_user/current_document.php?id=11654) Crosscheck of JVET-Z0059 (Non-EE2: Adaptive width for GPM blending area) [Z. Deng (ByteDance)] [late]

[JVET-Z0243](https://jvet-experts.org/doc_end_user/current_document.php?id=11694) Crosscheck of JVET-Z0059 (Non-EE2: Adaptive width for GPM blending area) [X. Xiu (Kwai)] [late]

[JVET-Z0062](https://jvet-experts.org/doc_end_user/current_document.php?id=11497) AHG12: RPR luma filter modifications and RPR enabling fixes in ECM [R. Yu, K. Andersson (Ericsson)]

[JVET-Z0228](https://jvet-experts.org/doc_end_user/current_document.php?id=11679) Crosscheck of RPR enabling fix in JVET-Z0062 (AHG12: RPR luma filter modifications and RPR enabling fixes in ECM), JVET-Z0067 (Non-EE2: RPR bug fixes and new results of RPR with luma-only re-scaling) and JVET-Z0079 (Non-EE2: Support RPR on ECM-4.0 and handling method for template matching based coding tools [Z. Zhang (Qualcomm)] [late]

[JVET-Z0230](https://jvet-experts.org/doc_end_user/current_document.php?id=11681) Crosscheck of JVET-Z0062 (AHG12: RPR luma filter modifications and RPR enabling fixes in ECM) [Z. Zhang (Qualcomm)] [late]

[JVET-Z0063](https://jvet-experts.org/doc_end_user/current_document.php?id=11498) AHG12: Fix for parsing of MHP information in ECM [R. Yu (Ericsson)]

In ECM-4.0, the tool multi-hypothesis prediction (MHP) allows signalling of multiple pieces of additional motion information for an inter block. The MHP is noted to be applicable for a current inter block when it is coded using normal merge, affine merge, MMVD or AMVP with BCW enabled. Furthermore, when the current block is coded using affine AMVP with BCW, the MHP is only applicable when the block is coded with affine AMVR off (e.g., with its motion vector difference resolution at quarter-pel). In other words, the presence of MHP motion information conditionally relies on the affine AMVR information. However, it is noted that the parsing of affine AMVR information is done after the MHP motion information is parsed. It is asserted that such a design has a parsing issue since the parsing of MHP motion information depends on undefined affine AMVR information. The contribution proposes a fix to resolve such a parsing issue. The proposed fix has been tested under ECM-4.0 RA and LDB common test configuration. The overall BD-rate impact for luma is reported to be 0.00% and -0.01%, respectively.

Was presented in session 11 (chaired by JRO)

Bitstreams are not decodable if using BcwFast=0.

JVET-Z0083 deals with the same issue. See further notes there.

[JVET-Z0236](https://jvet-experts.org/doc_end_user/current_document.php?id=11687) Crosscheck of JVET-Z0063 (AHG12: Fix for parsing of MHP information in ECM) [Y. Zhang (Qualcomm)] [late]

[JVET-Z0064](https://jvet-experts.org/doc_end_user/current_document.php?id=11499) AHG12: Convolutional cross-component model (CCCM) for intra prediction [P. Astola, J. Lainema, R. G. Youvalari, A. Aminlou, K. Panusopone (Nokia)]

This contribution proposes to apply a convolutional cross-component model (CCCM) based chroma prediction to improve compression efficiency of ECM. The proposed 7-tap convolutional filter maps luma values into chroma values when a CCCM prediction mode is activated by a PU level flag. The filter input consists of 5 spatial luma samples, a nonlinear term and a bias term. Filter coefficients are derived for each chroma block separately using regression based MSE minimization on reference samples in the PU’s neighborhood. The impact on coding efficiency and runtimes over ECM-4.0 is reportedly {for Y, U, V, EncT, DecT}:

AI { -1.43%, -3.51%, -3.50%, 102%, 103%}, RA { -0.79%, -2.43%, -2.77%, 101%, 100%}

The gains are asserted to be larger on higher resolution sequences and on TGM content.

Was presented in session 11 (chaired by JRO)

Better for higher resolutions, and in particularly high gain for class A2, mainly from park running sequence.

Probably more complex than other cross-component models (e.g., relatively large reference area), but interesting intra coding gain also for some other sequences.

Investigate in EE.

JVET-Z0140 is a similar proposal.

[JVET-Z0227](https://jvet-experts.org/doc_end_user/current_document.php?id=11678) Cross-check of JVET-Z0064 (Convolutional cross-component model for intra prediction) [V. Seregin (Qualcomm)] [late]

[JVET-Z0067](https://jvet-experts.org/doc_end_user/current_document.php?id=11502) Non-EE2: RPR bug fixes and new results of RPR with luma-only re-scaling [P. Bordes, F. Galpin, H. Guermoud, E. François (InterDigital)]

[JVET-Z0228](https://jvet-experts.org/doc_end_user/current_document.php?id=11679) Crosscheck of RPR enabling fix in JVET-Z0062 (AHG12: RPR luma filter modifications and RPR enabling fixes in ECM), JVET-Z0067 (Non-EE2: RPR bug fixes and new results of RPR with luma-only re-scaling) and JVET-Z0079 (Non-EE2: Support RPR on ECM-4.0 and handling method for template matching based coding tools [Z. Zhang (Qualcomm)] [late]

[JVET-Z0232](https://jvet-experts.org/doc_end_user/current_document.php?id=11683) Crosscheck of JVET-Z0067 (Non-EE2: RPR bug fixes and new results of RPR with luma-only re-scaling) [Z. Zhang (Qualcomm)] [late]

[JVET-Z0069](https://jvet-experts.org/doc_end_user/current_document.php?id=11504) AHG12: Longer chroma filters for RPR in ECM [K. Andersson, R. Yu (Ericsson)]

[JVET-Z0231](https://jvet-experts.org/doc_end_user/current_document.php?id=11682) Crosscheck of JVET-Z0069 (AHG12: Longer chroma filters for RPR in ECM) [Z. Zhang (Qualcomm)] [late]

[JVET-Z0078](https://jvet-experts.org/doc_end_user/current_document.php?id=11513) Non-EE2: Support MMVD and Affine MMVD without template matching process [H. Jang, J. Nam, N. Park, J. Lim, S. Kim (LGE)]

[JVET-Z0215](https://jvet-experts.org/doc_end_user/current_document.php?id=11666) Cross-check of JVET-Z0078 (Non-EE2: Support MMVD and Affine MMVD without template matching process) [H. Huang (Qualcomm] [late]

[JVET-Z0079](https://jvet-experts.org/doc_end_user/current_document.php?id=11514) Non-EE2: Support RPR on ECM-4.0 and handling method for template matching based coding tools [H. Jang, J. Nam, N. Park, J. Lim, S. Kim (LGE)]

[JVET-Z0228](https://jvet-experts.org/doc_end_user/current_document.php?id=11679) Crosscheck of RPR enabling fix in JVET-Z0062 (AHG12: RPR luma filter modifications and RPR enabling fixes in ECM), JVET-Z0067 (Non-EE2: RPR bug fixes and new results of RPR with luma-only re-scaling) and JVET-Z0079 (Non-EE2: Support RPR on ECM-4.0 and handling method for template matching based coding tools [Z. Zhang (Qualcomm)] [late]

[JVET-Z0083](https://jvet-experts.org/doc_end_user/current_document.php?id=11518) Non-EE2: Fix on MHP parsing condition [N. Park, J. Nam, J. Lim, S. Kim (LGE)]

This contribution is to address a potential issue in ECM4.0. Specifically, AMVR index is referenced before it is parsed. Therefore, in order to resolve the problem, two solutions are proposed.

- Solution #1: remove dependency between AMVR and MHP

- Solution #2: relocate MHP parsing behind AMVR syntax and remove dependency between MHP and MVSD

For Solution #1, simulation results reportedly show no changes in coding performance and run-time compared to ECM4.0 as anchor in CTC [2] because combination of MHP and AMVR for affine inter mode is not permitted from fast encoder option. But in case of disabling fast encoder option (BcwFast = 0) in cfg file, ECM 4.0 software works by the proposed modification, otherwise parsing error is observed.

For Solution #2, simulation results reportedly show -0.01%, -0.03% BD-rate changes compared to ECM4.0 as anchor in RA, LB configurations.

Was presented in session 11 (chaired by JRO)

According to crosscheckers, solution 1 from JVET-Z0083 is simpler than solution 2. A crosschecker who has checked both proposals also assesses solution 1 from Z0083 simpler than Z0063.

Several experts expressed support for the simplest solution.

Decision(BF): Adopt JVET-Z0083 solution 1.

[JVET-Z0217](https://jvet-experts.org/doc_end_user/current_document.php?id=11668) Crosscheck JVET-Z0083 (Non-EE2: Fix on MHP parsing condition) [X. Xiu (Kwai)] [late]

[JVET-Z0238](https://jvet-experts.org/doc_end_user/current_document.php?id=11689) Crosscheck of JVET-Z0083 (Non-EE2: Fix on MHP parsing condition) [Y. Zhang (Qualcomm)] [late]

[JVET-Z0085](https://jvet-experts.org/doc_end_user/current_document.php?id=11520) Non-EE2: AmvpMerge without decoder side mv refinement process [H. Jang, J. Nam, N. Park, J. Lim, S. Kim (LGE)]

In the ECM-4.0, amvpMerge is disabled when DMVD is disabled by the SPS flag since the bilateral matching based Merge candidate reordering and mv refinement are to be applied as default. However, even though DMVD is disabled, amvpMerge still can generate amvpMerge candidate list without bilateral matching based merge list reordering and mv refinement. In this proposal, it is suggested to enable amvpMerge without merge candidate reordering process and mv refinement process when the DMVD is disabled in SPS.

The simulation result of the proposed method shows (-0.14%, -0.20%, -0.16%) over ECM-4.0 with DMVD off for RA. The amvpMerge is automatically disabled at low delay case in the slice level therefore there is no changes for LDB configuration.

Was presented in session 13 (chaired by JRO)

Does not affect CTC where DMVD is on. The proposal is supported, as disabling AMVP merge when DMVD off is unnecessary, and the suggested change is straightforward.

Decision: Adopt JVET-Z0085.

[JVET-Z0229](https://jvet-experts.org/doc_end_user/current_document.php?id=11680) Crosscheck of JVET-Z0085 (Non-EE2: AMVP merge without decoder side mv refinement process) [Z. Zhang (Qualcomm)] [late]

[JVET-Z0100](https://jvet-experts.org/doc_end_user/current_document.php?id=11547) Non-EE2: Weighted Chroma Prediction (WCP) [J.-Y. Huo, H.-Q. Du, Z.-Y. Zhang, Y.-Z. Ma, F.-Z. Yang (Xidian Univ.), M. Li, Y. Liu (OPPO)]

This contribution proposes a weighted chroma prediction (WCP) method in which predicted chroma samples of the current block are derived based on the weights and neighboring chroma samples. It is proposed that the weights are designed according to the difference between neighboring luma samples and the reconstructed luma samples of a coding block. The test results on top of ECM-4.0 are as follows:

AI(Y/U/V): -0.09%/-0.70%/-0.67%, EncT 98%, DecT 101%

Was presented in session 13 (chaired by JRO)

The method determines a weight based on luma difference. The weight is determined from a trained LUT (based on softmax).

Encoding time is unreliable. Most likely encoding time increases, as another mode is checked by RDO.

Gain is not large, but also almost 0.1% on luma which indicates that the improved chroma prediction reduces the bit rate.

It was asked if the method would still give gain in combination with JVET-Z0064.

Investigate in EE.

[JVET-Z0102](https://jvet-experts.org/doc_end_user/current_document.php?id=11549) Non-EE2: Correction of the ARMC re-ordering step regarding the zero candidates [G. Laroche, P. Onno, R. Bellessort (Canon)]

This contribution presents a correction of the Adaptive Merge Candidates list Reordering process (ARMC). It is reported that when excluding the zero candidates from the reordering process of the ARMC, this exclusion provides a coding efficiency benefit while saving several cost computations performed by template matching.

Compared to ECM-4.0, the average BDR gains and runtimes reported in this contribution are as follows:

-0.04%/-0.04%/-0.04% - 98%, 97%for the RA configuration,

-0.11%/-0.14%/-0.16% - 100% , 97%for the Low Delay B configuration.

Was presented in session 13 (chaired by JRO)

Obvious design inconsistency. Confirmed and supported by cross-checker to be a simple and effective change, improving performance and reducing encoding and decoding time.

Decision: Adopt JVET-Z0102

[JVET-Z0248](https://jvet-experts.org/doc_end_user/current_document.php?id=11699) Cross-check of JVET-Z0102 (Non-EE2: Correction of the ARMC re-ordering step regarding the zero candidates) [L. Zhao (Bytedance)] [late]

[JVET-Z0103](https://jvet-experts.org/doc_end_user/current_document.php?id=11550) Non-EE2: On ARMC improvements [G. Laroche, P. Onno (Canon)]

This contribution presents some modifications of the Adaptive Merge Candidates list Reordering process (ARMC). The proposed method, applied after the first reordering process, reorders the list according to the distortion computed during the first ordering to target a more diverse list of Merge candidates for the final list. In addition, the merge derivation is adapted to this process to obtain additional coding gains. It is also reported that this particular merge derivation adaptation does not provide any gain when directly applied on top of ECM4.0.

Compared to ECM-4.0, the average BDR gains and runtimes reported in this contribution and combined with JVET-Z0102 are as follows:

-0.27%/-0.24%/-0.18% - 101%, 102% for the RA configuration,

-0.47%/-0.58%/-0.70% - 102%, 102% for the Low Delay B configuration.

Was presented in session 13 (chaired by JRO)

In previous EE, no successful improvement of ARMC. This one looks more interesting, and shall be investigated in EE along with JVET-Z0145.

[JVET-Z0105](https://jvet-experts.org/doc_end_user/current_document.php?id=11552) AHG12: Virtual boundary processing for the new In-Loop filter tools in ECM [A. M. Kotra, N. Hu, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution enables the virtual boundary (VB) processing for the new In-loop filters in ECM-4.0, which are Cross-component sample adaptive offset (CCSAO), Bilateral filter (BIF) and the improved ALF and CCALF.

The proposed method is evaluated on two software bases. The first software base being JVET-Y1063 [2], which proposed a new GDR implementation which internally uses the VB processing. However, JVET-Y0163 software disabled the new In-loop filter tools of ECM-3.0 as the virtual boundary processing is missing. This contribution implements the VB processing and reports the BD-Rate numbers on top of JVET-Y0163 GDR implementation. For the LDP configuration, the BD-Rate impact is as follows -2.67%(Y), -13.39%(U), -13.32%(V).

Secondly, on top of ECM-4.0 common test condition, VB is tested by explicitly adding one horizontal virtual boundary and one vertical virtual boundary and the results in RA are as follows:

0.25%(Y), 0.16%(U), 0.17% (V).

The software for both the bases is included along with the contribution package.

Furthermore, to validate the proposed VB processing implementation, JVET-Z0141 version 3 provided results by combining the proposed VB processing with JVET-Z0141 GDR implementation. It is reported that no leaks were observed in the generated GDR bitstreams.

Was presented in session 13 (chaired by JRO)

This virtual boundary processing would only be invoked when needed, e.g. at the boundary between clean an dirty area in GDR, or at non-adjacent cube face boundaries in 360 video.

This is different from virtual boundary at CTU boundaries for ALF in VVC. In that case, subjective artifacts might occur, whereas in the applications above VB processing reduces artifacts.

The design appears sufficiently mature, according to cross-checks, and combination with JVET-Z0141 was tested as well.

Decision: Adopt JVET-Z0105 (not CTC)

[JVET-Z0220](https://jvet-experts.org/doc_end_user/current_document.php?id=11671) Crosscheck of JVET-Z0105 (AHG12: Virtual boundary processing for the new In-Loop filter tools in ECM) [H.-J. Jhu, C.-W. Kuo (Kwai)] [late]

[JVET-Z0239](https://jvet-experts.org/doc_end_user/current_document.php?id=11690) Cross check of JVET-Z0105 (AHG12: Virtual boundary processing for the new In-Loop filter tools in ECM) [S. Hong (Nokia)] [late]

[JVET-Z0124](https://jvet-experts.org/doc_end_user/current_document.php?id=11572) Non-EE2: Spatial GPM [F. Wang, Y. Yu, H. Yu, D. Wang (OPPO)] [late]

Initial version rejected as “placeholder”.

[JVET-Z0194](https://jvet-experts.org/doc_end_user/current_document.php?id=11644) Crosscheck of JVET-Z0124 (Non-EE2: Spatial GPM) [H.-J. Jhu (Kwai)] [late]

[JVET-Z0126](https://jvet-experts.org/doc_end_user/current_document.php?id=11574) Non-EE2: Improvement on Local Illumination Compensation [Z. Xie, Y. Yu, H. Yu, D. Wang (OPPO)]

[JVET-Z0192](https://jvet-experts.org/doc_end_user/current_document.php?id=11642) Crosscheck of JVET-Z0126 (Non-EE2: Improvement on Local Illumination Compensation) [C.-C. Chen (Qualcomm)] [late]

[JVET-Z0127](https://jvet-experts.org/doc_end_user/current_document.php?id=11575) Non-EE2: On the maximum number of MHP merge candidates [H. Huang, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-Z0218](https://jvet-experts.org/doc_end_user/current_document.php?id=11669) Cross-check of JVET-Z0127 (Non-EE2: On the maximum number of MHP merge candidates) [H. Jang (LGE)] [late]

[JVET-Z0137](https://jvet-experts.org/doc_end_user/current_document.php?id=11586) Non-EE2: Adaptive Blending for GPM [H. Gao, X. Xiu, W. Chen, H.-J. Jhu, C.-W. Kuo, N. Yan, X. Wang (Kwai)]

[JVET-Z0223](https://jvet-experts.org/doc_end_user/current_document.php?id=11674) Crosscheck of JVET-Z0137 (Non-EE2: Adaptive Blending for GPM) [Y. Kidani, K. Kawamura (KDDI)] [late]

[JVET-Z0140](https://jvet-experts.org/doc_end_user/current_document.php?id=11589) AHG12: Enhanced CCLM [C.-W. Kuo, X. Xiu, N. Yan, H.-J. Jhu, W. Chen, H. Gao, X. Wang (Kwai)]

[JVET-Z0240](https://jvet-experts.org/doc_end_user/current_document.php?id=11691) Crosscheck of JVET-Z0140 (AHG12: Enhanced CCLM) [J. Lainema (Nokia)] [late]

[JVET-Z0143](https://jvet-experts.org/doc_end_user/current_document.php?id=11592) Non-EE2: Chroma intra modes derived from collocated luma blocks and neighboring chroma blocks [Y.-J. Chang, K. Cao, B. Ray, V. Seregin, M. Karczewicz (Qualcomm)]

Was again discussed in session 17 (chaired by JRO). According to the results shown in v2 of the document, the improvement compared to EE2 1.3b is 0.01% in luma, and 0.2-0.3% in chroma. This is not substantial enough to override a method investigated in the EE.

Follow-up discussion in session 18. The proponents point out that their main motivation is avoiding DIMD in chroma intra prediction (whereas 1.3b has still a reduced-complexity DIMD).

It was confirmed by multiple experts that the decision about adoption of EE2-1.3b should not be reverted.

Some experts thought that this proposal might have aspects of complexity reduction which need to be better understood.

Study in EE.

[JVET-Z0206](https://jvet-experts.org/doc_end_user/current_document.php?id=11657) Crosscheck of JVET-Z0143 (Non-EE2: Chroma intra modes derived from collocated luma blocks and neighboring chroma blocks) [P. Astola (Nokia)] [late]

[JVET-Z0146](https://jvet-experts.org/doc_end_user/current_document.php?id=11595) AHG12: Using samples before deblocking filter for adaptive loop filter [N. Hu, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-Z0212](https://jvet-experts.org/doc_end_user/current_document.php?id=11663) Cross-check of JVET-Z0146 (AHG12: Using samples before deblocking filter for adaptive loop filter) [K. Andersson (Ericsson)] [late]

[JVET-Z0147](https://jvet-experts.org/doc_end_user/current_document.php?id=11596) Non-EE2: Improvement of Inter-MTS in ECM [B. Ray, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-Z0235](https://jvet-experts.org/doc_end_user/current_document.php?id=11686) Crosscheck of JVET-Z0147 (Non-EE2: Improvement of Inter-MTS in ECM) [T. Hashimoto (Sharp)] [late]

[JVET-Z0149](https://jvet-experts.org/doc_end_user/current_document.php?id=11598) Non-EE2: Adaptive Filter Shape Switch for ALF [W. Yin, K. Zhang, L. Zhang (Bytedance)]

[JVET-Z0176](https://jvet-experts.org/doc_end_user/current_document.php?id=11626) Cross-check of JVET-Z0149 (Non-EE2: Adaptive Filter Shape Switch for ALF) [Z. Xie (OPPO)] [late]

[JVET-Z0159](https://jvet-experts.org/doc_end_user/current_document.php?id=11608) Non-EE2: Reconstruction-Reordered IBC for screen content coding [Z.Deng, K.Zhang, L.Zhang (Bytedance)]

This contribution proposes a Reconstruction-Reordered IBC (RR-IBC) mode for screen content video coding. When it is applied to an IBC coded block, the reconstruction block is flipped. Two symmetric flip methods, horizontal flip and vertical flip, are supported for RR-IBC coded blocks. Syntax flags are signalled for IBC AMVP, indicating whether the reconstruction block is flipped and how to flip it. For IBC merge, the flip type is inherited from neighbouring blocks without signalling. Moreover, a motion shift is added to a block vector candidate obtained from a neighbouring block according to a symmetric rule, in case that the neighbouring block is coded with RR-IBC. Simulation results based on ECM-4.0 are reported as below:

AI: Class F -1.33%, -1.20%, -1.36%, 103%, 99%; Class TGM -2.66%, -2.54%, -2.50%, 102%, 99%

RA: Class F -0.93%, -1.27%, -1.33%, 101%, 100%; Class TGM -1.10%, -0.91%, -0.89%, 101%, 100% Presented in session 17 (chaired by JRO)

It is commented that this could potentially be more beneficial with the enlarged reference area adopted for ECM5.

Investigate in EE.

[JVET-Z0219](https://jvet-experts.org/doc_end_user/current_document.php?id=11670) Non-EE2: SPS flag to control TM-based merge/amvp and multi-pass DMVR separately [H. Jang, J. Nam, N. Park, J. Lim, S. Kim (LGE)] [late]

This contribution suggests to add an SPS flag to control template matching based merge/amvp and multi-pass DMVR separately. Current ECM-4.0 has only one SPS flag named as DMVD, which controls TM based merge/amvp and multi-pass DMVR. Since there is only one flag, TM based merge/amvp and multi-pass DMVR should always be disabled at the same time when DMVD is set to 0. Therefore, to support functionality to disable/enable TM operation separately, this contribution proposes to add one SPS flag and related low-level change.

- Aspect 1 : adding SPS flag to control separately tm-based merge/amvp and multi-pass DMVR.

- Aspect 2 : low-level change to avoid disabling BM merge when TM merge is disabled.

Presented in session 17 (chaired by JRO).

Main intent is disabling all TM related elements altogether. In a real standard, this might better be an issue of a profile definition, however it appears useful in the context of the exploration

The proposal comes with 2 different methods, both requiring low-level changes additional to the SPS flag.

Results not complete yet for both methods. According to a cross-checker, the design is consistent.

Further study recommended (not relevant for EE at this moment, but could be interesting in future).

[JVET-Z0237](https://jvet-experts.org/doc_end_user/current_document.php?id=11688) Crosscheck of JVET-Z0219 (Non-EE2: SPS flag to control TM-based merge/amvp and multi-pass DMVR separately) [W. Lim, S.-C. Lim (ETRI)] [late]

[JVET-Z0246](https://jvet-experts.org/doc_end_user/current_document.php?id=11697) Crosscheck of JVET-Z0219 method 1 (Non-EE2: SPS flag to control TM-based merge/amvp and multi-pass DMVR separately) [J. Gan (OPPO)] [late]

# High-level syntax (HLS) and related proposals (11)

## AHG9: SEI message studies and proposals (7)

Contributions in this area were discussed in sessions 7 and 8 at 0615–0705 UTC and 0725-0925 on Friday 22 April 2022, and in session 15 at 0640-0805 UTC on Wednesday 27 April 2022 (chaired by JRO).

[JVET-Z0046](https://jvet-experts.org/doc_end_user/current_document.php?id=11480) VTM Software Implementation for GREEN-MPEG SEI Messaging [C. Herglotz, M. Kränzler, A. Kaup (FAU)]

This document proposes a reference implementation for signaling the Green MPEG complexity metrics for decoder-power reduction and quality metrics for quality recovery after low-power encoding, which are both defined for VVC SEI messages. The implementation complies with the issued DIS ISO/IEC 23001-11 3rd edition Energy-Efficient Media Consumption (Green Metadata). The implementation is available for the VTM-16.0 encoder and decoder implementation. We propose to include the code for a future version of VTM.

Decoder implementation is straightforward. For the encoder, it is proposed to include the bitstream feature analyser (previously presented in JVET-W0071) in VTM as well. VTM software coordinators shall investigate the code of the BSFA encoder and the practicality of integrating it into VTM until the next meeting.

Decision(SW): Adopt JVET-Z0046 (decoder part only)

[JVET-Z0052](https://jvet-experts.org/doc_end_user/current_document.php?id=11486) AHG9: NNR post-filter SEI message [M. M. Hannuksela, M. Santamaria, F. Cricri, E. B. Aksu, H. R. Tavakoli (Nokia), T. Chujoh, Y. Yasugi, T. Ikai (Sharp)]

This contribution proposes an NNR post-filter SEI message into a working draft of a VSEI amendment. The NNR post-filter SEI message may carry an ISO/IEC 15938-17 bitstream that specifies a neural-network-based post-filter. The NNR post-filter SEI message proposed in this contribution combines aspects of JVET-Y0073, JVET-Y0074, JVET-Y0075, and JVET-Y0115.

The complexity indication is not fully specified -> see JVET-Z0151

It was commented that using different mode idc for NNR v1 and v2 may not be necessary and not desirable because this information can be extracted from the NNR stream, and would then also allow using future NNR extensions.

It was commented that using 21 bits for the id is not consistent with other SEI designs. Multiple filters would be possible in the same bit stream, therefore a long id value was chosen.

It was commented that usage of persistence/cancellation mechanisms in case that the previous id was unknown should be made consistent with other SEI designs.

It was asked under which circumstances the decoder can select the patch size for processing. This is the case for convolutional networks, not for fully connected architectures.

It was commented that using more than 2 bits for mode\_idc would be preferable for possible extensibility.

Is software available? Partially yes, related to JVET-Z0082. This at least verifies that the concept in combination with NNR works, including incremental signalling of NNR v2.

The URI mechanism seems to require some more clarification about its usage and usefulness. The intended purpose could basically be achieved by sending user data via a proprietary SEI message. Referring to a specific URI might be more appropriate (if such a thing would exist). This could be done (if needed) in a future extension.

[JVET-Z0120](https://jvet-experts.org/doc_end_user/current_document.php?id=11568) AHG9: Shutter interval information SEI message for VSEI [S. McCarthy, F. Pu, W. Husak, P. Yin, T. Lu, A. Arora, T. Shao (Dolby), J. R. Arumugam, S. Agrawal, K. Patankar (Ittiam)]

This contribution proposes that the shutter interval information (SII) SEI message already specified in HEVC be adopted in the next version of VSEI. A functionally equivalent SII SEI message is also already specified in AVC. The SII SEI message indicates the shutter interval for the associated video source pictures prior to encoding, e.g., for camera-captured content, the shutter interval is the amount of time that an image sensor is exposed to produce each source picture. Shutter interval information can be useful for several use cases, for example: enabling backwards compatibility of high frame rate (HFR) bit streams with receivers that support only standard frame rate (SFR) bit streams such as specified in ETSI TS 101 154 and ATSC A/341; prediction of subjective quality of motion; frame rate conversion; and detection of non-camera-captured content such as screen captures and computer-generated content. Additional descriptions, methods, and showcase VTM software is provided to illustrate the use of the SII SEI message with backwards compatible HFR bit streams.

Identical to HEVC version, with name of one syntax element aligned to VVC.

The processing described in the proposal is included in the software package as a showcase of SFR/HFR conversion (non-normative). This reflects ATSC A/341 specification.

It was commented by other aspects that having a software implementation of such functionality is desirable in general.

Decision: Adopt JVET-Z0120 into a first working draft of VSEI amendment. For timeline, better wait if other SEI messages would be progressing. It also needs alignment with VVC, but not necessarily in the ongoing VVC amendment.

Decision(SW): Include the software part into VTM.

[JVET-Z0202](https://jvet-experts.org/doc_end_user/current_document.php?id=11652) Crosscheck of JVET-Z0120 (AHG9: Shutter interval information SEI message for VSEI) [S. Deshpande (Sharp)] [late]

[JVET-Z0121](https://jvet-experts.org/doc_end_user/current_document.php?id=11569) AHG9: Neural-network post filtering SEI message [S. McCarthy, A. Arora, T. Shao, P. Yin, T. Lu, F. Pu, W. Husak (Dolby)]

This contribution proposes the following SEI messages be adopted into a working draft of a VSEI amendment:

1. A neural-network post-processing filter characteristics SEI message that signals information that can be used to determine the capabilities required to perform the indicated post-processing filtering; and

2. A picture-adaptive neural-network post-processing filter SEI message that signals the specific model and parameters that may be used for the current picture.

Main difference compared to JVET-Z0052/0151 is the definition of two different SEI messages, to be used for whole CVS or for a single picture.

It was questioned if two SEI messages are necessary. The proposal JVET-Z0051 could be used for all cases (whole sequence, certain picture types, or individual pictures).

The SEI message also includes signalling of QP. It is commented that QP is VVC specific, not defined in VSEI, and may also be highly dependent on an encoder’s usage, in particular when variable QP is used. Also usage of “picture type” inter/intra is inconsistent – should be slice type.

See further notes under JVET-Z0151.

[JVET-Z0129](https://jvet-experts.org/doc_end_user/current_document.php?id=11577) AHG9: SEI for transparency information as dedicated layer for transparent screens [E. Thomas, P. Andrivon, F. Le Leannec, M. Radosavljević, M.-L. Champel (Xiaomi)]

It is asserted that a conventional video content displayed on a transparent display may result in a degraded experience for the viewer where important parts of the image may be lost such as logos, computer windows or characters, etc. Additionally, it is further asserted that the transparency effect perceived by the viewer of those transparent displays is controlled by the sample values of the displayed images and depends on the panel technology and the ambient lightning conditions.

Based on those observations, it is proposed to define metadata payload to signal the content creator intent with respect to how the decoded pictures should look like on a transparent display, i.e., which area should be transparent and to what degree.

To this end, following up on JVET-Y0104, a SEI signaling is presented which enables the use of a dedicated layer for the transport of the necessary information about the intended transparency to be achieved after post-decoding image operations.Presented in session 15.

In the previous proposal, the transparency information was directly coded in the SEI message. In the current proposal, it is encoded in layer 1 of a stream (which also carries the SEI message).

Transparency information was generated for some test sequences from classes C and E (based on depth analysis) – background is made transparent.

Similar as an alpha plane. However, the purpose is different: No blending with an alternative background is performed, the background is made transparent. Alpha channels are already supported in VVC, and it is assumed that the background is supplied externally at the decoder (or via systems level) – this could also be “no background” in case of transparent displays. In alpha channels, only foreground and its transparency are sent, no signalling what is done with that. This would fully fulfill the purpose of a transparent display application.

What is the relation with the existing ACI SEI? It is argued that they could both be used together, ACI for blending and transparency for the display. This would either require a separate SEI message, or a two-layer alpha map which is currently not supported.

It appears that the transparency-only application could be done with ACI.

How relevant would the combination be? Would some of this compositing be done at systems level?

Further study on these aspects recommended.

[JVET-Z0151](https://jvet-experts.org/doc_end_user/current_document.php?id=11600) AHG9: On NNR post-filter SEI message [Y. Yasugi, T. Chujoh, T. Ikai (Sharp)]

This contribution proposes syntax and semantics refinements about tensor input/output data type, patch size, and complexity information based on the NNR post-filter SEI message proposed in JVET-Z0052.

Aspects: Color format indication (different for input and output), free patch size support (also non-square rectangular, and making syntax elements dependent on constant\_patch\_size\_flag), support for integer tensors, complexity information.

JVET-Z0052 already allows different colour formats for input and output.

It was commented that making the syntax elements for patch size dependent on the flag would give up some flexibility that a decoder could choose larger patches.

It was commented that the signalling of block-wise processing also in JVET-Z0052 may in general require more considerations, may be dependent on network architecture if it works or not. See discussion under JVET-Z0052 above – for certain architecures such as fully connected constant patch size would need to be used.

The specification of integer input/output (in addition to floating point of JVET-Z0052) is asserted useful.

The additional complexity information is asserted useful, as otherwise a decoder would first have to analyse the NNR stream before knowing if it can operate the network.

In summary, the following elements are asserted useful in addition to JVET-Z0052: Non-square rectangular partitions, integer input/output, complexity information.

As an overall conclusion on JVET-Z0052, JVET-Z0121 and JVET-Z0151, which have many commonalities: Proponents should discuss offline which elements of JVET-0121 would be useful to be included in JVETR-Z0052 (additional to the elements listed above from JVET-0151, some of which are also in JVET0121), and if possible provide a merged proposal on a NN post processing SEI message. In general, this topic and its specification seems to have arrived at sufficient maturity to be included in a VSEI draft. See JVET-Z0244.

[JVET-Z0244](https://jvet-experts.org/doc_end_user/current_document.php?id=11695) AHG9: NN post-filter SEI [M. M. Hannuksela, M. Santamaria, F. Cricri, E. B. Aksu, H. R. Tavakoli (Nokia), T. Chujoh, Y. Yasugi, T. Ikai (Sharp), S. McCarthy, A. Arora, T. Shao, P. Yin, T. Lu, F. Pu, W. Husak (Dolby)] [late]

This contribution proposes neural-network (NN) post-filter supplemental enhancement information into a working draft of a VSEI amendment. This contribution merges aspects of JVET-Z0121 and JVET-Z0151 into JVET-Z0052.

The base text without highlighting is the JVET-Z0052 text and the items of JVET-Z0151 that were agreed in the first JVET review, i.e., the following:

1. non-square rectangular partitions

2. integer input/output

3. complexity information

The offline work among the proponents resulted into the following proposals categorized with highlighted specification text as follows:

- Changes in response to the first JVET review:

1. nnpfc\_id was aligned with fp\_arrangement\_id as follows:

a. ue(v) and value ranges are used like in fp\_arrangement\_id.

b. The explicit phrasing “determined by the application” was excluded since in the NN post-filtering case, nnpfc\_id is determined by the encoder.

2. Changes that were made related to nnpfc\_mode\_idc:

a. The nnpfc\_mode\_idc value to indicate a URI was excluded in response to the comments in the first JVET review: "The URI mechanism seems to require some more clarification about its usage and usefulness. The intended purpose could basically be achieved by sending user data via a proprietary SEI message. Referring to a specific URI might be more appropriate (if such a thing would exist). This could be done (if needed) in a future extension."

b. The nnpfc\_mode\_idc values for standalone NNR bitstream and NNR update were united into one value in response to the comments in the first JVET review: "It was commented that using different mode idc for NNR v1 and v2 may not be necessary and not desirable because this information can be extracted from the NNR stream, and would then also allow using future NNR extensions."

c. Instead of u(2), ue(v) was taken into use in response to the following comment in the JVET review: "It was commented that using more than 2 bits for mode\_idc would be preferable for possible extensibility."

- The cyan highlighting in the document indicates a proposal of two SEI messages, one for filter characteristics (with a scope until the end of CLVS) and another for picture-wise filter activation. It is asserted that this design enables to indicate the characteristics for several filters on CLVS basis. The picture-wise filter activation message enables selection of the filter on picture basis, without repeating the filter characteristics. The idea of CLVS-level filter characteristics SEI message and picture-level activation SEI message originates from JVET-Z0121.

If the proposal of two SEI messages is not acceptable, it is proposed to fall back to the Z0052 design on the SEI message with indicated persistence.

- Items originating from JVET-Z0121:

1. Addition of a mode indicator specifying that the post-processing filter is determined by external means not specified in this Specification. This value can be used, for example, when there are systems means to indicate a post-processing filter.

2. Addition of "unknown or unspecified" purpose value.

3. Addition of a figure to illustrate interleaving of luma data channels of nnpfc\_inp\_order\_idc equal to 3 (informative).

4. Addition of padding type syntax element and zero and reflection padding (in addition to the replication padding that was already present in JVET-Z0052).

5. The syntax elements nnpfc\_parameter\_type\_flag and nnpfc\_log2\_parameter\_bit\_length\_minus3 for the neural network complexity structure were added to merge aspects from both JVET-Z0151 and JVET-Z0121:

nnpfc\_parameter\_type\_flag equal to 0 indicates that the neural network uses only integer parameters. nnpfc\_parameter\_type\_flag equal to 1 indicates that the neural network may use floating point or integer parameters.

nnpfc\_log2\_parameter\_bit\_length\_minus3 equal to 0, 1, 2, and 3 indicates that the neural network does not use parameters of bit length greater than 8, 16, 32, and 64, respectively.

One argument for having two SEI message is that the parameters relating to the overall architecture are only sent once per sequence, and then only parameters of different models used for specific pictures would need to be sent in the other SEI.

Could there be multiple filters in one picture? Potentially, one for luma and one for chroma.

NNR could use external mechanisms – is this supported? It is left vague, as the NNR bitstream is considered as a black box.

Is the incremental NNR functionality operating? Has been implemented in software for JVET-Z0082, which will also be used in the next EE.

nnpf\_mod\_idc equal zero is expressing that filter is provided by external means. What benefit would that have? One example is that a filter is externally specified by NNR, and updated via the SEI.

Generally, the text is assessed to be mature.

This is asserted to be a first step for supporting NN based technology in a standard, but not specifying a specific technology, but rather supporting a mechanism to operate a filter, which could also include (and would most likely include) low-complexity designs. Further, as it is post processing, it is not mandatory for decoders to operate it.

Decision: Adopt JVET-Z0244 (WD of VSEI amendment)

It was later decided to issue a request for VSEI amendment.

## Film grain synthesis (3)

Contributions in this area were discussed in session 15 at 0815–0925 UTC on Wednesday 27 April 2022 (chaired by JRO).

[JVET-Z0047](https://jvet-experts.org/doc_end_user/current_document.php?id=11481) AHG13: Improvements of film grain analysis [P. de Lagrange, E. François, Z. Ameur (InterDigital), M. Radosavljević (Xiaomi)]

This contribution proposes improvements for film grain analysis part. It introduces improved film grain scaling factor estimation and introduces two new options in the VTM code. Those new options are related with the ability to use externally calculated masks of flat regions and externally calculated filtered frames. Current version only supports the use of MCTF and Canny edge detector in conjunction with morphological operations that are coded within VTM. Film grain synthesis part at the decoder side is not affected by this contribution.How many cutoff frequency pairs? One, as the previous software version. It was suggested that further improvement would be possible supporting more (RDD5 supports up to 10).

The results are confirmed by crosscheckers. It is also confirmed by them that the subjective quality is improved by better estimation of scaling factor.

Decision(SW): Adopt JVET-Z0047.

[JVET-Z0205](https://jvet-experts.org/doc_end_user/current_document.php?id=11656) Crosscheck of JVET-Z0047 (AHG13: Improvements of film grain analysis) [F. Pu (Dolby)] [late]

[JVET-Z0115](https://jvet-experts.org/doc_end_user/current_document.php?id=11562) AHG13: Proposed film grain synthesis reference model [P. de Lagrange, F. Urban, E. François (InterDigital)]

Was discussed in BoG (A. Tourapis) at 2320-0120+1 on Wednesday 27 April 2022

This document describes a film grain synthesis model that supports both modes of FGS SEI (frequency-filtering and autoregressive), as well as other parameter formats, with attention paid to hardware complexity. It is proposed to study this model in ad-hoc group on film grain technologies, as a candidate for the reference model of a potential “soft conformance” definition.

Q: Comment that artificial noise is important for archiving purposes.

Q: Up to now we have been studying what FG models are being used. It seems that now we are going on a different direction and asking to maybe standardize a particular model.

A. This is only proposing a reference model of a conformance requirement is imposed.

Software supports YUV 4:2:0 data at 8/10 bits as inputs and the film grain information is provided in a text file (single model). The film grain is then added on all frames of the YUV sequence. Multiple models are not supported currently natively (could be done in a separate application). The SW also only supports film grain addition in the native content domain (no color transformations are supported).

A comment was made that the industry needs clear specifications for film grain.

It was pointed out that this is just an example and it was not clear why this example was better than other implementations (especially for HW).

There were several arguments on how this will be addressed. For example, we can define new profile (or profiles?) for VVC or maybe there could be work on a standalone specification (or specifications) for film grain modeling. These could have different advantages or disadvantages (e.g., former may result in a market fragmentation or could create issues with existing deployed or upcoming devices that do not support such profiles, while the later may not be as emphatic but may be easier to deploy and may be able to allow multiple incarnations of different film grain technologies).

How would this impact the technical report? Originally it was thought that the technical report may help the market better understand how to use and encourage it to use film grain technologies. Would doing more help better in the adoption or would just the technical report and any supporting SW be sufficient? In general, it is felt that the value of the technical report is significant and completing it would likely benefit the industry and the adoption of FGS technologies. Having some reference SW implementations, on both the encoder and decoder side, could also help this further.

It was also added that this could have implications not only to VVC but also previous standards such as HEVC and creating new profiles there might be disruptive. Careful thought is needed of how to address this issue and how to proceed.

Some elements of this contribution could be introduced into the technical report.[JVET-Z0132](https://jvet-experts.org/doc_end_user/current_document.php?id=11581) AHG13: On film grain synthesis [Y. He, M. Coban, M. Karczewicz (Qualcomm), P. de Lagrange, E. François (InterDigital), M. Radosavljević (Xiaomi), S. McCarthy, W. Husak (Dolby)]

This contribution proposes study of film grain synthesis and relevant conformance requirements in AHG13 with the following items:

1. Conformance requirements

2. Auto-regressive model reference software development

3. Reference film grain synthesis process

It is suggested to define that a conformant device shall either produce bit-exact replication of the reference synthesis, or fulfills a “soft” conformance requirement that no “perceptually significant difference” occurs relative to the reference synthesis. It shall not be specified what “perceptually significant” means. This basically means that a conformant post-processing device shall at least have some mechanism to generate film grain synthesis for all pictures.

It is generally agreed that the film grain synthesis should allow some degree of freedom to implementers, so defining only bit-exact conformance would be inappropriate.

It is also pointed out that post processing may not only include FGS, but also other elements such as sharpening, etc.

It is pointed out that the quality of devices is also regulated by market competition, and this has been the usual way why post processing was never regarded to need normative conformance requirements.

Currently, no proposal is on the table how the “soft” conformance would be specified, and in which standard. Would a corresponding device declare itself to be conformant? Further, it would be inappropriate to require conformance for an SEI message of version 1 in a later version. A possible option could be a definition of a profile.

It was suggested to further discuss these issues in the AHG, and add a mandate on this.

A BoG (A. Tourapis) was created to further discuss these issues, and also review JVET-Z0115 which relates to synthesis (including AR model based).

The BoG reported in session 16 Thursday 0500-0540 UTC

During this report, also the status of the WD for the TR on FGS was discussed. It was suggested not creating a new version at this meeting, and prepare a new version proposal as an input to the next meeting in the context of the AHG activity. At the current meeting, there was no input document that could be integrated at the current stage.

It was confirmed that the TR can refer to external references if they are stable and accurately describing the method, and this should be done in a neutral way, or try to interpret it. The report should not try developing a “better” description of a technology than available elsewhere.

## Non-SEI HLS aspects (1)

Contributions in this area were discussed in session 16 at 0540–0600 UTC on Thursday 28 April 2022 (chaired by JRO).

[JVET-Z0164](https://jvet-experts.org/doc_end_user/current_document.php?id=11613) YCgCo-R: Request for new code points in CICP [D. Buitenhuis (Vimeo), A. M. Tourapis (Apple)] [late]

In this contribution it is requested that additional code points are introduced in the CICP specification to better handle the YCgCo-R representation in applications that require both the luma and chroma information to be signalled using the same bit depth. Such representations were originaly presented in contributions JVET-T0111 and JVET-U0093, where coding gains for both lossless and lossy applications were documented.

It was agreed that this is beneficial for a reversible YCgCo-R representation with equal bit depth for luma and chroma. It was commented that the text is mature now. Beyond CICP, this should also be added to AVC and HEVC.

For CICP, draft 1. In WG 5, issue a request for new edition and CD.

For HEVC, issue a draft, and consider the possibility of later integrating it into the new edition. It could be in the next edition of H.265 (October), and potentially in FDIS if there would be a related comment in the DIS ballot (FDIS planned for January 2023).

# Plenary meetings, joint meetings, BoG reports, and liaison communications

## JVET plenaries

No intermediate plenaries were held, as document review and decisions were made in single-track mode at this meeting.

(The remainder of this section is kept as a template for future use.)

Some of the discussions and actions at plenary sessions are noted in this section.

XXday X January XXXX–XXXX:

* Joint meetings:
  + …
* BoG reports
  + …
* Planning of outputs
  + …
* Scheduling, remaining doc reviews

Fri. 28 Apr. 0500-0900 UTC (sessions 25 and 26):

* Review of remaining docs from section XX
* See notes under sections 8, 9, 10, and 11.

## Information sharing meetings

Information sharing sessions with other WGs and AGs of the MPEG community were held on Monday 25 Apr. 0500–0730, Wednesday 27 Apr. 0500–0600, and Friday 29 Apr. 2100–2300. The status of the work in the MPEG WGs and AGs was reviewed at these information sharing sessions.

## Joint meeting with … XXXX–XXXX on XXday X January

No joint meetings were held during the current meeting. This section is kept in the report as a template for future use.

The following topics were discussed in this joint session. See also the notes recorded on these topics in other sections of this document.

## BoGs (0)

No break-out groups (BoGs) were established at the current meeting. This section is kept in the report as a template for future use.

The following break-out groups were established at this meeting to conduct discussion and develop recommendations on specific topics.

## Liaison communications

The JVET received the following liaison related communication at its current meeting.

A response WG 5 N XXX was drafted and reviewed during plenary session on Friday 29 April at XXXX UTC. The response described the current status of … .

# Project planning

## Software timeline

ECM5 software (including all adoptions) was planned to be available 3 weeks after the meeting.

VTM17 software was planned to be available on 2022-05-27. (Note that updates 16.1/16.2 are planned to be released shortly after the meeting)

HM16.26 software was planned to be available on 2022-05-13. A new version of HTM is targeted before the next meeting.

## Core experiment and exploration experiment planning (update)

An EE on neural network-based video coding was established, as recorded in output document JVET-Y2023.

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-Y2024.

Initial versions of these documents were presented and approved in the session 25 on Friday 21 January.

## Drafting of specification text, encoder algorithm descriptions, and software

The following agreement has been established: the editorial team has the discretion to not integrate recorded adoptions for which the available text is grossly inadequate (and cannot be fixed with a reasonable degree of effort), if such a situation hypothetically arises. In such an event, the text would record the intent expressed by the committee without including a full integration of the available inadequate text.

## Plans for improved efficiency and contribution consideration

The group considered it important to have the full design of proposals documented to enable proper study.

Adoptions need to be based on properly drafted working draft text (on normative elements) and HM/VTM encoder algorithm descriptions – relative to the existing drafts. Proposal contributions should also provide a software implementation (or at least such software should be made available for study and testing by other participants at the meeting, and software must be made available to cross-checkers in EEs).

Suggestions for future meetings included the following generally-supported principles:

* No review of normative contributions without draft specification text
* VTM algorithm description text is strongly encouraged for non-normative contributions
* Early upload deadline to enable substantial study prior to the meeting
* Using a clock timer to ensure efficient proposal presentations (5 min) and discussions

As general guidance, it was suggested to avoid usage of company names in document titles, software modules etc., and not to describe a technology by using a company name.

## General issues for experiments

It was emphasized that those rules which had been set up or refined during the 12th JVET meeting should be observed. In particular, for some CEs of some previous meetings, results were available late, and some changes in the experimental setup had not been sufficiently discussed on the JVET reflector.

Group coordinated experiments have been planned as follows:

* “Core experiments” (CEs) are the coordinated experiments on coding tools which are deemed to be interesting but require more investigation and could potentially become part of a draft standard by the next meeting or in the near future.
* “Exploration experiments” (EEs) are also coordinated experiments. These are conducted on technology which is not foreseen to become part of a draft standard in near future. 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”, 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 that the description matches the software, and recommendation of value of the tool 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.

Availability of spec text is important to have a detailed understanding of the technology and also to judge what its impact on the complexity of the spec will be. There must also be sufficient time to study it in detail. CE contributions without sufficiently mature draft spec 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 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 reviewed and 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) | N |
| **Test material and visual assessment (AHG4)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Consider plans for additional verification testing of VVC capability, particularly target establishing a test plan for VVC scalability features by the next meeting. * Maintain the video sequence test material database for testing the VVC and HEVC standards and potential future extensions, as well as exploration activities. * Study coding performance and characteristics of available and proposed video test material. * Identify and recommend appropriate test material for testing the VVC standard and potential future extensions, as well as exploration activities. * Identify and characterize missing types of video material, solicit contributions, collect, and make available a variety of video sequence test material. * Maintain and update the directory structure for the test sequence repository as necessary. * Collect information about test sequences that have been made available by other organizations. * Prepare and conduct remote expert viewing for purposes of subjective quality evaluation. * Coordinate with AG 5 in studying and developing further methods of subjective quality evaluation, e.g. based on crowd sourcing. * Prepare availability of viewing equipment and facilities arrangements for future meetings. | V. Baroncini, T. Suzuki, M. Wien (co-chairs), S. Liu, 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. | V. Seregin (chair), J. Chen, F. Le Léannec, K. Zhang (vice-chairs) | N |
| **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. | M. Karczewicz, Y. Ye, L. Zhang (co-chairs), B. Bross, X. Li, K. Naser, H. Yang (vice-chairs) | N |
| **Film grain technologies (AHG13)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the benefits and characteristics of film grain technologies, including autoregressive and frequency-filtering technologies. * 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 signaling, 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 N046 “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 (update)

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-Z1000](https://jvet-experts.org/doc_end_user/current_document.php?id=11704) Meeting Report of the 26th JVET Meeting [J.-R. Ohm] [WG 5 N 124] (2022-05-27)

Initial versions of the meeting notes (d0 … dX) 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 WD N 132] [G. J. Sullivan, A. Tourapis] (2022-05-27)

This includes identifiers for YCoCg-R color space.

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-Z1004](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, C. Rosewarne, G. J. Sullivan, Y. Syed, Y.-K. Wang] (2022-06-30, near next meeting)

[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.

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.

Upon 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.

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]

[JVET](https://jvet-experts.org/doc_end_user/current_document.php?id=11708)-Z1008 Additional color type identifiers for AVC and HEVC (Draft 1) [G. J. Sullivan, A. Tourapis] (2022-06-03)

This includes identifiers for YCoCg-R color space.

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

Remains valid – not updated [JCTVC-V1014](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=10316) Screen Content Coding Test Model 7 Encoder Description (SCM 7) [R. Joshi, J. Xu, R. Cohen, S. Liu, Y. Ye] [WG 11 N 16049]

Remains valid for HM – not updated: [JCTVC-Z1015](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=10689) Common Test Conditions for Screen Content Coding [H. Yu, R. Cohen, K. Rapaka, J. Xu]

No output: JVET-X1016 through JVET-X1019

Remains valid for HM – not updated: [JVET-Y1100](https://jvet-experts.org/doc_end_user/current_document.php?id=11466) Common Test Conditions for HM Video Coding Experiments [K. Sühring, K. Sharman]

contains only range extensions CTC.

Remains valid – not updated: [JVET-T2001](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10540) Versatile Video Coding Draft 10 [B. Bross, J. Chen, S. Liu, Y.-K. Wang]

[JVET-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)

New elements: Encoder optimization from JVET-Y0077, JVET-Y0085, JVET-Y0105, JVET-Y0126, JVET-Y0152, JVET-Y0155

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-Z2005](https://jvet-experts.org/doc_end_user/current_document.php?id=11710) New level and systems-related supplemental enhancement information for VVC (Draft 2) [B. Bross, E. François, A. Tourapis, Y.-K. Wang] [WG 5 CDAM 1 N 129] (2022-05-13)

Number changed from JVET-Y2019.

[JVET-Z2006](https://jvet-experts.org/doc_end_user/current_document.php?id=11711) Additional SEI messages for VSEI (Draft 1) [S. McCarthy, T. Chujoh, M. M. Hannuksela, G. J. Sullivan, Y.-K. Wang] [WG 5 WD N 126] (2022-06-17)

JVET-Z0120, JVET-Z0244 are included.

The corresponding variables in VVC could be included in a later version of document 2019, or as a separate amendment of VVC.

A request for a new edition (WG 5 N 125) was reviewed Friday 29 April at 0850 UTC. Target dates are CDAM in July 2022, DAM in October 2022, FDAM in July 2023

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]

Was 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]

[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)

Merge of HM/VTM CTC of HDR as per JVET-Z0175. No need to review.

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-Z2016](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] (2022-05-13)

This includes the cross-check of training procedures as suggested in BoG JVET-Z0234. No need to review

Remains valid – not updated: [JVET-Y2017](https://jvet-experts.org/doc_end_user/current_document.php?id=11473) Common Test Conditions and evaluation procedures for enhanced compression tool testing [M. Karczewicz and Y. Ye]

Remains valid – not updated: [JVET-U2018](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]

No output: JVET-X2019

Remains valid – not updated: [JVET-Y2020](https://jvet-experts.org/doc_end_user/current_document.php?id=11475) Film grain synthesis technology for video applications (Draft 1) [D. Grois, Y. He, W. Husak, M. Radosavljević, A. Tourapis, W. Wan] [WG 5 [N 120](https://dms.mpeg.expert/doc_end_user/current_document.php?id=82207&id_meeting=189)]

No outputs: JVET-X2021, JVET-X2022

This number is retained for future purposes of planning possible additional VVC verification testing reports and other purposes.

[JVET-Z2023](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 133] (2022-05-13)

An initial draft of this document was reviewed and approved at 0735 UTC on Friday 29 April. It was suggested that it would be interest to study the impact of training (e.g. in term of BPG or VVC be used for coding key pictures) also for the new end-to-end method.

[JVET-Z2024](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 134] (2022-05-27)

An initial draft of this document was reviewed and approved on Friday 29 April at 0750 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 it mandatory, as it might even happen that proposals are withdrawn from EE if no clearance of disclosure is reached within a company.

[JVET-Z2025](https://jvet-experts.org/doc_end_user/current_document.php?id=11714) Algorithm description of Enhanced Compression Model 5 (ECM 5) [M. Coban, F. Le Léannec, K. Naser, J. Ström] [WG 5 N 135] (2022-06-30)

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)

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]

# Future meeting plans, expressions of thanks, and closing of the meeting

Future meeting plans were established according to the following guidelines:

* 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 case of online meetings, no sessions should be held on weekend days. This may imply an earlier starting date as well.

Some specific future meeting plans (to be confirmed) were established as follows:

* Wed. 13 – Fri. 15 and Mon. 18 – Fri. 22 July 2022, 27th meeting under ISO/IEC JTC 1/‌SC 29 auspices, to be held as teleconference meeting.
* Fri. 21 – Fri. 28 October 2022, 28th meeting under ITU-T SG16 auspices in Antalya, TR.
* 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.

The agreed document deadline for the 27th JVET meeting was planned to be Wednesday 6 July 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 26th JVET meeting was closed at approximately 0020 hours UTC on Saturday 30 April 2022.

# Annex A to JVET report: List of documents

# Annex B to JVET report: List of meeting participants

The participants of the twenty-sixth meeting of the JVET, according to the participation records from the Zoom teleconferencing tool used for the meting sessions (approximately XXX people in total, not including those who attended only the joint sessions with other groups), were as follows:

1. First Last (company – country)

# Annex C to JVET report: Recommendations of the 7th 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 123**](https://dms.mpeg.expert/doc_end_user/current_document.php?id=81988&id_meeting=189)

**(link to be updated)**

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)
2. [] L. Wang, X. Xu, S. Liu, F. Galpin. EE1-1.2: neural network based in-loop filter with a single model. JVET-Z0091, 26th Meeting, by teleconference, 20–29 April 2022. [↑](#endnote-ref-3)