|  |  |
| --- | --- |
| **Joint Video Experts Team (JVET)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29**  25th Meeting, by teleconference, 12–21 January 2022 | Document: JVET-Y\_Notes\_d8 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Meeting Report of the 25th Meeting of the Joint Video Experts Team (JVET), by teleconference, 12–21 January 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-fifth meeting during 12–21 January 2022 as an online-only meeting. It had previously been planned to be in Geneva, CH, 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 sixth meeting as WG 5. The JVET meeting was held under the chairmanship of Dr Jens-Rainer Ohm (RWTH Aachen/Germany). For rapid access to particular topics in this report, a subject categorization is found (with hyperlinks) in section 2.14 of this document. It is further noted that work items which had originally been conducted by the Joint Collaborative Team on Video Coding (JCT-VC) were continued in JVET as a single joint team, and explorations towards possible future need of standardization in the area of video coding are also conducted by JVET, as negotiated by the parent bodies.

The JVET meeting began at approximately 1300 hours UTC on Wednesday 12 January 2022. Meeting sessions were held on all days except the weekend days of Saturday and Sunday 15 and 16 January 2022, until the meeting was closed at approximately XXXX hours UTC on XXday 2X January 2022. Approximately XXX people attended the JVET meeting, and approximately XXX input documents (not counting crosschecks), 13 AHG reports, 2 EE summary reports, X BoG reports, and X report on experts viewing conducted during the meeting were discussed. The meeting took place in a coordinated fashion with a teleconference meeting of SG16 – one of the two parent bodies of the JVET, under whose auspices this JVET meeting was held. Various other SC29 Working Groups – where WG 5 is representing the Joint Video Coding Team(s) and their activities from the perspective of the SC 29 parent body – were also meeting in parallel, and some coordination was conducted with those as well. 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 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-fourth JVET meeting in producing the following documents:

* JVET-X1000 Meeting report
* JVET-X1004 Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR
* JVET-X1005 New level for HEVC (Draft 1)
* JVET-X2002 Algorithm description for Versatile Video Coding and Test Model 15 (VTM 15)
* JVET-X2005 VVC operation range extensions (Draft 5)
* JVET-X2006 Additional SEI messages for VSEI (Draft 5)
* JVET-X2008 Conformance testing for versatile video coding (Draft 7)
* JVET-X2016 Common Test Conditions and evaluation procedures for neural network-based video coding technology
* JVET-X2017 Common Test Conditions and evaluation procedures for enhanced compression tool testing
* JVET-X2022 Core Experiment on film grain synthesis
* JVET-X2023 Exploration Experiment on neural network-based video coding (EE1)
* JVET-X2024 Exploration Experiment on enhanced compression beyond VVC capability (EE2)
* JVET-X2025 Algorithm description of Enhanced Compression Model 3 (ECM 3)
* JVET-X2026 Conformance testing for VVC operation range extensions (Draft 2)

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

* JVET-X1004 Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR
* JVET-X1005 New level for HEVC (Draft 1)
* JVET-X2002 Algorithm description for Versatile Video Coding and Test Model 15 (VTM 15)
* JVET-X2005 VVC operation range extensions (Draft 5)
* JVET-X2006 Additional SEI messages for VSEI (Draft 5)
* JVET-X2008 Conformance testing for versatile video coding (Draft 7)
* JVET-X2016 Common Test Conditions and evaluation procedures for neural network-based video coding technology
* JVET-X2017 Common Test Conditions and evaluation procedures for enhanced compression tool testing
* JVET-X2022 Core Experiment on film grain synthesis
* JVET-X2023 Exploration Experiment on neural network-based video coding (EE1)
* JVET-X2024 Exploration Experiment on enhanced compression beyond VVC capability (EE2)
* JVET-X2025 Algorithm description of Enhanced Compression Model 3 (ECM 3)
* JVET-X2026 Conformance testing for VVC operation range extensions (Draft 2)

For the organization and planning of its future work, the JVET established XX “ad hoc groups” (AHGs) to progress the work on particular subject areas. At this meeting, X Exploration Experiments (EE) were defined. The next eight JVET meetings were planned for Wed. 20 – Fri. 22 and Mon. 25– Fri. 29 April 2022 under ISO/IEC JTC 1/‌SC 29, to be conducted as a teleconference meeting; during Wed. 13 – Fri. 22 July 2022 under ISO/IEC JTC 1/‌SC 29 auspices, to be conducted as a teleconference meeting; during October 2022 under ITU-T SG16 auspices, date and location t.b.d.; 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 in Geneva, CH, date t.b.d.; during October 2023 under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.; and during January 2024 under ISO/IEC JTC 1/‌SC 29 auspices, date and location t.b.d.

The document distribution site <https://jvet-experts.org/> was used for distribution of all documents. It was noted that the previous 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-fifth meeting during 12–21 January 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 sixth 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. 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 minor 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_01_Y_Virtual/>.

## Primary goals

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

* JVET-X1000 Meeting report
* JVET-X1004 Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR
* JVET-X1005 New level for HEVC (Draft 1)
* JVET-X2002 Algorithm description for Versatile Video Coding and Test Model 15 (VTM 15)
* JVET-X2005 VVC operation range extensions (Draft 5)
* JVET-X2006 Additional SEI messages for VSEI (Draft 5)
* JVET-X2008 Conformance testing for versatile video coding (Draft 7)
* JVET-X2016 Common Test Conditions and evaluation procedures for neural network-based video coding technology
* JVET-X2017 Common Test Conditions and evaluation procedures for enhanced compression tool testing
* JVET-X2022 Core Experiment on film grain synthesis
* JVET-X2023 Exploration Experiment on neural network-based video coding (EE1)
* JVET-X2024 Exploration Experiment on enhanced compression beyond VVC capability (EE2)
* JVET-X2025 Algorithm description of Enhanced Compression Model 3 (ECM 3)
* JVET-X2026 Conformance testing for VVC operation range extensions (Draft 2)

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/> is still accessible, but was 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, 5 January 2022. Any documents uploaded after 1159 hours Paris/Geneva time on Thursday 6 January 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-Y0163 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-Y0XXX (a proposal on …), uploaded 01-XX.
* …

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

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

* JVET-Y0XXX (a document on …), uploaded 01-XX.
* …

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

The following 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: JVET-Y01XXX, …. Therefore, they were markesd as withdrawn by action of the chair.

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-Y0064, JVET-Y0220, JVET-Y0234, ….

“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 typically be counted as a late contribution). At the current meeting, this situation did not apply.

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.). These 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 were 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-X1000, the Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR JVET-X1004, the New level for HEVC (draft 1) JVET-X1005, the Algorithm description for Versatile Video Coding and Test Model 15 (VTM 15) JVET-X2002, the Operation range extensions for VVC (Draft 5) JVET-X2005, the Additional SEI messages for VSEI (Draft 5) JVET-X2006, the Conformance testing for Versatile Video Coding (draft 7) JVET-X2008, the Common Test Conditions and evaluation procedures for neural network-based video coding technology JVET-X2016, the Common Test Conditions and evaluation procedures for enhanced compression tool testing JVET-X2017, the Description of the EE on Neural Network-based Video Coding JVET-X2023, the Description of the EE on Enhanced Compression beyond VVC capability JVET-X2024, the Algorithm description of Enhanced Compression Model 3 (ECM 3) JVET-W2026, and the Conformance testing for VVC operation range extensions (Draft 2) JVET-X2026, had been completed and were approved. It was noted that the WG 5 output version of JVET-X2008 is still in the phase of editorial finalization before being submitted as FDIS. The software implementations of VTM (version 15.0), ECM (versions 3.0 and 3.1) were also approved.

Only minor editorial issues were found in the meeting report JVET-X1000 – 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
* 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 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, Paris; 8 hours ahead of the time in Los Angeles, etc.). No session should last longer than 2 hrs.

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

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 22nd JVET meeting (which had been the fifth 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.

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 1353 (as of 9 January 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).

It was pointed out that recently an access problem had been detected with the ftp site at University of Hannover that had been used by JCT-VC historically. Mathias Wien and Karsten Sühring will check if all sequences that were hosted there are already available in the new ftp site.

## Terminology

* **ACT**: Adaptive colour transform
* **AFF**: Adaptive frame-field
* **AI**: All-intra
* **AIF**: Adaptive interpolation filtering
* **ALF**: Adaptive loop filter
* **AMP**: Asymmetric motion partitioning – a motion prediction partitioning for which the sub-regions of a region are not equal in size (in HEVC, being N/2x2N and 3N/2x2N or 2NxN/2 and 2Nx3N/2 with 2N equal to 16 or 32 for the luma component)
* **AMVP**: Adaptive motion vector prediction
* **AMT or MTS**: Adaptive multi-core transform, or multiple transform selection
* **AMVR**: (Locally) adaptive motion vector resolution
* **APS**: Adaptation parameter set
* **ARC**: Adaptive resolution conversion (synonymous with DRC, and a form of RPR)
* **ARMC**: Adaptive re-ordering of merge candidates
* **ARSS**: Adaptive reference sample smoothing
* **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 constitutes our draft standard design – 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.
* **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
* **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 12 January at 1300 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
  + Working practices using objective metrics report
    - HSTP-VID-WPOM V1: approved 2020-07-03, published 2020-11
    - ISO/IEC TR 23002-8 (Ed. 1) published 2021-05-20
  + 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 (during the current meeting)
    - ISO/IEC 14496-10:2020 (Ed. 9) FDIS ballot closed 2020-11-27, published 2020-12-15 – corresponding aspects for additional SEI messages are partly in the in-progress DIS (see below)
  + 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), was 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
  + Usage of code points report
    - H.Sup19 V3 approved 2021-04-30, published 2021-06-04
    - ISO/IEC TR 23091-4 (Ed. 3) published 2021-05-23
  + VVC
    - H.266 V1 approved 2020-08-29, published 2020-11-10
    - ISO/IEC 23090-3:2021 (Ed. 1) published 2021-02-16
  + VSEI
    - H.274 V1 approved 2020-08-29, published 2020-11-10
    - ISO/IEC 23002-7:2021 (Ed. 1) published 2021-01-28
  + CICP V2 (includes errata items)
    - ISO/IEC 23091-2 had been forwarded from DIS directly for publication in 2021-04

Post-meeting note: This was published after the meeting on 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
  + 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)
    - ISO/IEC 23002-7:2021 (Ed. 1) VSEI
    - ISO/IEC 23008-2:2020 (Ed. 4) HEVC
    - ISO/IEC 23090-3:2021 (Ed. 1) VVC
    - ISO/IEC 23091-2:2019 (Ed. 1) Video CICP
    - CICP 2nd ed.?
    - HEVC Amd.1?
* Draft standards progression status
  + VVC conformance – FDIS from previous meeting – ballot to be started soon, to be submitted for ITU-T consent from current meeting
  + VCC conformance for operation range extensions – CDAM from previous meeting, ballot closed XX-XX, progression to DAM, no action yet in ITU-T
    - [m58554](https://dms.mpeg.expert/doc_end_user/current_document.php?id=81284&id_meeting=189) Summary of voting on ISO/IEC 23090-15 CDAM1
  + VVC reference SW – under DIS 23090-16 ballot closed 2021-09-30, FDIS and ITU-T consent submission from current meeting – shall also include reference software for version 2 of VVC and VSEI
    - [m57767](https://dms.mpeg.expert/doc_end_user/documents/136_OnLine/wg11/m57767-v1-m57767.zip) Summary of voting on ISO/IEC DIS 23090-16
    - [n20909](https://dms.mpeg.expert/doc_end_user/current_document.php?id=81004&id_meeting=188) Draft disposition of comments received on ISO/IEC DIS 23090-16 (WG 5 N87) – to be converted into DoCR of WG 5
  + AVC additional SEI messages – integrated into DIS of new edition at 23rd meeting, ballot closed 2021-12-27, FDIS from current meeting – already in most recent edition of H.264
    - [m58533](https://dms.mpeg.expert/doc_end_user/current_document.php?id=81263&id_meeting=189) Summary of voting on ISO/IEC DIS 14496-10
  + VSEI extensions – integrated into DIS of new edition (version 2) at 23rd meeting, ballot closed 2021-12-27, FDIS and ITU-T consent submission from current meeting
    - [m58534](https://dms.mpeg.expert/doc_end_user/current_document.php?id=81264&id_meeting=189) Summary of voting on ISO/IEC DIS 23002-7
  + VVC operation range extensions – integrated into DIS of new edition (version 2) at 23rd meeting, ballot closed 2021-12-27, FDIS and ITU-T consent submission from current meeting
    - [m58535](https://dms.mpeg.expert/doc_end_user/current_document.php?id=81265&id_meeting=189) Summary of voting on ISO/IEC DIS 23090-3
  + New level (from JVET-X1005) to be included in CD targeting new edition on HEVC incorporating Amd.1 and corrigenda items in October, and consent in ITU-T by the same time; note that shutter interval is already in latest ITU-T edition of H.265
  + The request for free availability in ISO/IEC has to be made for each edition, amendment and corrigendum, and these will also need a request form to be filled out and be approved in the Recommendations. A freely available URL on ITU part should be provided for the following parts:
    - ISO/IEC 14496-10:202X – AVC new edition
    - ISO/IEC 23002-7:202X – VSEI new edition
    - ISO/IEC 23090-3:202X – VVC new edition
    - ISO/IEC 23090-15:202X – VVC conformance
    - ISO/IEC 23090-16:202X – VVC reference software
    - ISO/IEC 23008-2:2020/Amd 1 – HEVC done 23rd meeting
    - ISO/IEC DIS 23091-2, 2nd edition – CICP done 23rd meeting
* 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-X1000 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 sections 2.1 and 2.3, some numbering errors occurred: VTM should be 15 instead 14, ECM should be 3 instead 2, JVET-X2007 should be JVET-X2008. In the document site and output document list, everything is correct.
  + …
* There was somewhat less of a problem of late non-cross-check documents and no “placeholders” (see section 2.4.2).
* There were quite a few documents where authors’ given names were not abbreviated, and company affiliation was missing in the authors’ list. Reminder to stick to our conventions.
* The primary goals of the meeting were
  + Finalization V1 of VVC conformance and software
  + New edition of 14496-10
  + Errata
  + Conformance DAM for version2 of VVC
  + Version 2 of VVC (High bit rate / high bit depth)
  + Version 2 of VSEI
  + Exploration Experiments
    - Neural network-based video coding
    - Enhanced compression beyond VVC
* Funding of verification testing activities: recommendation of thanks, recommendation calling for funding wrt upcoming tests. Potentially new version of verification test plan (scalable coding)
* Liaison communication
* Number of documents slightly higher than last meeting (110->135)
* Scheduling was discussed, and it was agreed to avoid conducting “track” sessions in parallel (some BoG parallelism might occur)
* Principles of standards development were discussed.
* Meeting plans need to be discussed, in particular regarding the possibility of a hybrid meeting in April 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.

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

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

* Wed. 12 Jan., 1st day
  + Session 1:
    - 1300–1345 Opening remarks, review of practices, agenda, IPR reminder
    - 1345–1505 Reports of AHGs 1–X
  + Session 2:
    - 1525–1605 Reports of AHGs X–13
    - 1605–1735 Review of EE1 summary report
  + Session 3:
    - 2100–2300 Review of EE2
  + Session 4:
    - 2320–0120+1 Review of EE2 and related
* Thu. 13 Jan., 2nd day
  + Session 5:
    - 1300–1500 Review of V2 VVC/VSEI: 4.2, 4.7, 4.12, potentially relevant contributions from 6.x
  + Session 6:
    - 1520–1720 Review of EE1 and related
  + Session 7:
    - 2100–2300 Review of EE2 and related
  + Session 8:
    - 2320–0120+1 Review of EE2 and related
* Fri. 14 Jan., 3rd day
  + Session 9:
    - 1300–1510 Review of V2 VVC/VSEI (remaining 4.2, 4.12, 6.x), subjective viewing 4.6
  + Session 10:
    - 1530–1720 Review of EE1 and related
  + Session 11:
    - 2100–2300 Review of EE2 and related
  + Session 12:
    - 2320–0120+1 Review of VSEI/HLS (4.2, 6.x)
* Mon. 17 Jan., 4th day
  + 0500–0730 MPEG information sharing session
  + 1215–1345 SG16/WP3 plenary (no JVET meeting in parallel)
  + Session 13:
    - 1400–1500 Review of 4.4 verification test and 4.5 test material
  + Session 14:
    - 1520–1720 Review remaining 4.2 text/errata & 6.x HLS
  + Session 15:
    - 2100–2300 Review 4.3 test conditions, 4.10 low latency
  + Session 16:
    - 2320–0120+1 Review 4.10 low latency
* Tue. 18 Jan., 5th day
  + 1200–1620 ITU-T workshop on “AI and multimedia”
  + Session 17:
    - 2100–2300 Review remaining 5.3.4
  + Session 18:
    - 2320–0120+1 Review remaining 5.3.4; remaining 6.x, tentatively other remaining docs
* Wed. 19 Jan., 6th day
  + 0500–0630 MPEG information sharing session
  + 1400–1530 VCEG
  + Session 19:
    - 1550–1750 Review 6.2 FGS,
  + Session 20:
    - 2100–2300 Revisits, EE & other planning; 4.1 deployment status, 4.7 conformance, 4.8 software
  + Session 21:
    - 2320–0140+1 4.11 encoding optimization
* Thu. 20 Jan., 7th day
  + Session 22:
    - 1520–1720 Remaining docs from 4.5, 4.8, 4.9, 4.11 review (chaired by Frank Bossen)
  + Session 23:
    - 2100–2250 Revisits; output document planning, AHG and EE planning
  + Session 24:
    - 2310–0045+1 Review DoCR documents, remaining 4.5/4.8 docs
* Fri. 21 Jan., 8th day
  + Plenary:
    - 1300–1500 Remaining 4.11 docs, Approval of output docs, AHGs, recommendations
    - 1520–1720 Remaining business, Approval of output docs, AHGs, recommendations
  + 2100–2300 MPEG information sharing session
  + XXXX+1–XXXX+1 WG 5 Closing plenary: Approval of meeting recommendations

## Contribution topic overview

The approximate subject categories and quantity of contributions per category for the meeting were summarized as follows (note that the noted document counts do not include crosschecks, and 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 (4)
  + Test conditions (6)
  + Verification testing (1)
  + Test Material (3)
  + Quality assessment (3)
  + Conformance test development (1)
  + Software development (2)
  + Implementation studies (4)
  + AHG7: Low latency and constrained complexity (7)
  + Encoding algorithm optimization (11) 3 TBP
  + Profile/tier/level specification (5)
  + 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 (2) (section 5.1)
  + AHG11 and EE1: Neural network-based video coding (23) (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 (9) (section 6.1)
  + Film grain synthesis (3) (section 6.2)
  + Non-SEI HLS aspects (0) (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 12 Jan. 2022 in session 1 during 1345–1505 and in session 2 1525–1605 UTC (chaired by JRO).

[JVET-Y0001](https://jvet-experts.org/doc_end_user/current_document.php?id=11358) 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 higher number of input documents (as compared to the previous meeting) submitted to the current meeting. Intense discussion had been carried out on the group email reflector, and all output documents from the preceding meeting had been produced.

Output documents from the preceding meeting had been made initially available at the JVET web site (<https://jvet-experts.org/>) or the ITU-based JVET site (<http://wftp3.itu.int/av-arch/jvet-site/2021_10_X_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-X1000) [Posted 2021-11-13, also submitted as WG 5 N 82]
* Errata report items for VVC, HEVC, AVC, Video CICP, and CP usage TR (JVET-X1004) [Posted 2021-11-17]
* New level for HEVC (Draft 1) (JVET-X1005) [Posted 2021-11-16]
* Algorithm description for Versatile Video Coding and Test Model 15 (VTM 15) (JVET-X2002) [Posted 2021-12-28, also submitted as WG 5 N 92]
* VVC operation range extensions (Draft 5) (JVET-X2005) [Posted 2021-11-12]
* Additional SEI messages for VSEI (Draft 5) (JVET-X2006) [Posted 2021-11-12]
* Conformance testing for Versatile Video Coding (Draft 7) (JVET-X2008) [Posted 2021-11-30, also submitted as WG 5 FDIS N 84]
* Common Test Conditions and evaluation procedures for neural network-based video coding technology (JVET-X2016) [Posted 2021-11-03]
* Common Test Conditions and evaluation procedures for enhanced compression tool testing (JVET-X2017) [Posted 2021-10-29]
* EE on Neural Network-based Video Coding (EE1) (JVET-X2023) [Posted 2021-10-15, last update 2021-11-02, also submitted as WG 5 N 88]
* EE on Enhanced Compression beyond VVC capability (EE2) (JVET-X2024) [Posted 2021-10-14, last update 2021-11-06, also submitted as WG 5 N 89]
* Algorithm description of Enhanced Compression Model 3 (ECM 3) (JVET-X2025) [Posted 2021-12-23, last update 2022-01-07, also submitted as WG 5 N 90]
* Conformance testing for VVC operation range extensions (Draft 2) (JVET-X2026) [Posted 2021-11-14, also submitted as WG 5 CDAM1 N 86]

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

* Recommendations of the 5th WG 5 meeting (WG 5 N 81)
* Disposition of comments received on ISO/IEC DIS 23090-15 Conformance testing for Versatile Video Coding (WG 5 N 83)
* Request for ISO/IEC 23090-15 Amd.1 Conformance testing for VVC operation range extensions (WG 5 N 85)
* Draft disposition of comments received on ISO/IEC DIS 23090-16 VVC reference software (WG 5 N 87)
* List of AHGs established at the 5th WG 5 meeting (WG 5 N 91)

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 25th meeting had been made publicly available on the ITU-hosted ftp site <http://wftp3.itu.int/av-arch/jvet-site/2022_01_Y_Virtual/>.

The AHG recommends its continuation.

[JVET-Y0002](https://jvet-experts.org/doc_end_user/current_document.php?id=11359) JVET AHG report: Draft text and test model algorithm description editing (AHG2) [B. Bross, J. Chen, 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-X2005 VVC operation range extensions (Draft 5)**

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 5 incorporated items:

* Maximum bit rate modification and addition of gci\_all\_rap\_pictures\_constraint\_flag (JVET-X0079v3).
* WPP fix for high bit depth (JVET-X0128)
* GCI flags for high bit depth (JVET-X0076/X0095)
* Addition of intra profile constraints (JVET-X0106)
* Addition of MV wrap-around restriction to CREI SEI (JVET-X0101)
* Some editorial clarifications and cleanups (JVET-X0050)
* Changes on and related the specification of range extensions profiles, including for the value range of vps\_ols\_dpb\_bitdepth\_minus8[ i ] and sps\_bitdepth\_minus8, the general\_profile\_idc values for all the 16-bit profiles, the decoder capability requirements for the range extension profiles, and allowed values in Table A.1 (JVET-X0073/JVET-X0093)
* Addition of a constraint that a multiview view position SEI message shall not be contained in a scalable nesting SEI message (JVET-X0096)
* Text addressing ticket #1513

**JVET-X2006 Additional SEI messages for VSEI (Draft 5)**

This document contains the draft text for changes to the versatile supplemental enhancement information messages for coded video bitstreams (VSEI) standard (Rec. ITU-T H.274 | ISO/IEC 23002-7), to specify additional SEI messages, including the 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 5 incorporated items:

* JVET-X0059 AHG2/AHG9: Comments on the 2nd edition draft text for VSEI

**JVET-X1004 Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR**

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

Incorporated items at the JVET-X meeting:

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

**JVET-X2002 Algorithm description for Versatile Video Coding and Test Model 15 (VTM 15)**

The JVET established the VVC Test Model 15 (VTM15) algorithm description and encoding method at its 24th meeting (6 – 15 October 2021, teleconference). This document serves as a source of general tutorial information on the VVC design and also provides an encoder-side description of VTM15. 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 15 (VTM15) algorithm description and encoding method

* Incorporated JVET-X0128: AHG8: On History-Based Rice Parameter Derivations for Wavefront Parallel Processing
* Incorporated JVET-X0048 CE: Film Grain Synthesis (test CE2.1 and CE2.2)
* Incorporated JVET-X0143 AHG10: VTM Encoder Changes for ALF Usage with Subpicture

**Related input contributions**

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

* JVET-Y0049: AHG2/AHG8: On the range extensions GCI flags
* JVET-Y0050: AHG2/AHG9: On the alpha channel information SEI message
* JVET-Y0056: AHG2: High tier for lower levels
* JVET-Y0057: AHG2: MinCr for still picture profiles
* JVET-Y0063: AHG2: On Main 10 4:4:4 Still Picture profile for VVC v1 and v2
* JVET-Y0072: New Levels for HEVC and VVC [Listed due to title. Document not available as at Jan 11th]
* JVET-Y0190: AHG2/AHG8: Suggestions for the operation range extensions GCI

**Remaining bug tickets**

* [#1517](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1517) Potential Mismatch between VVC Spec and VTM Reference C Decoder Model
* [#1520](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1520) Subclause 8.6.1 - IBC SINGLE\_TREE case processing chroma in monochrome
* [#1521](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1521) Missing a default/inferred value for alf\_cc\_cb/cr\_filters\_signalled\_minus1

Experts are asked to further study these items and identify if action is necessary, and potentially make a correction already in VVC v2 (or software spec).

It was identified that there is no issue on #1521 as the case of the non-inferred value would never occur.

#1517 had already been closed.

#1520 should be possible to be resolved in editing.

**Recommendations**

The AHG recommends to:

* Approve JVET-X1004, JVET-X2002, JVET-X2005, and JVET-X2006 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-Y0003](https://jvet-experts.org/doc_end_user/current_document.php?id=11360) JVET AHG report: Test model software development (AHG3) [F. Bossen, X. Li, K. Sühring, Y. He, K. Sharman, V. Seregin, A. Tourapis]

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

* [VTM 15.0](https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/-/releases/VTM-15.0) (Dec. 2021)
* [HM-16.24](https://vcgit.hhi.fraunhofer.de/jvet/HM/-/releases/HM-16.24) (Oct. 2021)
* [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.

**Software development**

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

The server is located at:

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

The registration and development workflow are documented at:

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

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

**VTM related activities**

The VTM software can be found at

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

The software development continued on the GitLab server. VTM versions 14.1 and 14.2 were tagged on Oct. 18 and Oct. 26, and VTM version 15.0 was tagged on Dec. 3. VTM 15.1 is expected during the 25th JVET meeting.

VTM 14.1 was tagged on Oct. 18, 2021. Changes include:

* JVET-W0133: Constrained RASL encoding for bitstream switching
* JVET-S0154 aspect9 and JVET-S0158 aspect4: bitstream extraction of extracting non-nested SEI from nested SEI
* JVET-S0117 Virtual boundary rewriting for subpicture extraction
* JVET-W0078: Multiview View Position SEI message
* JVET-T0055 item 2: Remove the step 9 of the general sub-bitstream extraction...
* JVET-S0176 item5: constrain the value of sps\_num\_subpics\_minus1 when SLI SEIs show up
* Fix #1508: Use correct chroma format in call to verifyPlane
* Fix #1509: lossless coding of RGB content
* Fixes for multi-layer (spatial scalability) including fix for #1510
* Fix #1514 - Potential memory leak for m\_sdiSEIInFirstAU and m\_maiSEIInFirstAU

VTM 14.2 was tagged on Oct. 26, 2021. Changes include:

* Remove macros from previous cycle

VTM 15.0 was tagged Dec. 3, 2021. Changes include:

* JVET-X0116: Enabled temporal filter for low-delay configurations
* JVET-X0079 and JVET-X0106: Modified bitrates for v2 high tier and constraint on intra profiles
* JVET-X0128 method 2, WPP
* JVET-X0137: Determination of signalled parameters for ETSRC and RLSCP
* JVET\_X0076\_X0095\_V2\_GCI
* JVET-X0101: added WrapAround constraint to Constrained RASL Encoding SEI message
* JVET-X0143: VTM Encoder Changes for ALF usage with Subpicture
* Fix #1515: Overflow in ALF for bitdepth >10
* Fix #1518: wrong RPR flag condition check in xInitSPS()
* Fix for the ARSEI implementation

VTM 15.1 is expected to be tagged during the 25th JVET meeting. Changes include so far:

* Add check of reference frame structure to detect low delay configuration
* JVET-X0048-X0103: implementation of film grain analysis and synthesis
* JVET-X0073: Decrease profile\_idc for 16 bit profiles by 1

***CTC Performance***

The major performance impact for VTM 15.0 under SDR CTC was enabling temporal filtering in low delay configuration. Since the corresponding HM version is not released yet, the following tables still show **VTM 14.0** performance over **HM 16.24** for a fair comparison:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **All Intra Main10** |  |  |
|  |  |  | **Over HM-16.24** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | -29,03% | -32,17% | -34,07% | 1565% | 168% |
| Class A2 | -29,29% | -23,92% | -21,06% | 2522% | 181% |
| Class B | -21,73% | -26,96% | -30,76% | 2798% | 189% |
| Class C | -22,54% | -18,95% | -22,70% | 3955% | 194% |
| Class E | -25,75% | -25,91% | -24,45% | 2264% | 178% |
| **Overall** | -25,06% | -25,37% | -26,85% | 2602% | 183% |
| Class D | -18,46% | -13,31% | -13,41% | 4527% | 166% |
| Class F | -39,33% | -39,73% | -42,22% | 5396% | 182% |
|  |  |  |  |  |  |
|  |  |  | **Random access Main10** |  |  |
|  |  |  | **Over HM-16.24** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | -39,85% | -39,48% | -46,15% | 670% | 165% |
| Class A2 | -43,19% | -40,54% | -39,68% | 747% | 183% |
| Class B | -36,30% | -48,60% | -47,20% | 740% | 169% |
| Class C | -33,16% | -34,83% | -36,95% | 1014% | 193% |
| Class E |  |  |  |  |  |
| **Overall** | -37,55% | -41,49% | -42,75% | 791% | 177% |
| Class D | -31,45% | -31,40% | -31,26% | 1134% | 177% |
| Class F | -45,76% | -49,18% | -50,10% | 581% | 161% |
|  |  |  |  |  |  |
|  |  |  | **Low delay B Main10** |  |  |
|  |  |  | **Over HM-16.24** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -29,24% | -34,80% | -32,41% | 777% | 158% |
| Class C | -25,89% | -17,42% | -17,95% | 947% | 178% |
| Class E | -28,73% | -33,03% | -26,38% | 380% | 147% |
| **Overall** | -28,00% | -28,56% | -26,08% | 694% | 161% |
| Class D | -25,01% | -12,57% | -11,79% | 979% | 183% |
| Class F | -40,20% | -41,56% | -41,87% | 508% | 141% |
|  |  |  |  |  |  |
|  |  |  | **Low delay P Main10** |  |  |
|  |  |  | **Over HM-16.24** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -33,97% | -37,79% | -34,99% | 715% | 168% |
| Class C | -27,68% | -17,28% | -18,05% | 857% | 188% |
| Class E | -32,32% | -36,86% | -30,30% | 374% | 160% |
| **Overall** | -31,46% | -30,72% | -28,17% | 646% | 172% |
| Class D | -26,32% | -11,99% | -10,87% | 901% | 191% |
| Class F | -39,97% | -41,10% | -41,48% | 549% | 150% |

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 15.0** performance compared to **VTM 14.0**:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **All Intra Main10** |  |  |
|  |  |  | **Over VTM-14.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | 0,00% | 0,00% | 0,00% | 100% | 100% |
| Class A2 | 0,00% | 0,00% | 0,00% | 98% | 99% |
| Class B | 0,00% | 0,00% | 0,00% | 99% | 97% |
| Class C | 0,00% | 0,00% | 0,00% | 99% | 99% |
| Class E | 0,00% | 0,00% | 0,00% | 98% | 98% |
| **Overall** | 0,00% | 0,00% | 0,00% | 99% | 98% |
| Class D | 0,00% | 0,00% | 0,00% | 99% | 99% |
| Class F | 0,00% | 0,00% | 0,00% | 98% | 97% |
|  |  |  |  |  |  |
|  |  |  | **Random access Main10** |  |  |
|  |  |  | **Over VTM-14.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | 0,00% | 0,00% | 0,00% | 99% | 101% |
| Class A2 | 0,00% | 0,00% | 0,00% | 99% | 99% |
| Class B | 0,00% | 0,00% | 0,00% | 100% | 99% |
| Class C | 0,00% | 0,00% | 0,00% | 99% | 94% |
| Class E |  |  |  |  |  |
| **Overall** | 0,00% | 0,00% | 0,00% | 99% | 98% |
| Class D | 0,00% | 0,00% | 0,00% | 99% | 97% |
| Class F | 0,00% | 0,00% | 0,00% | 99% | 99% |
|  |  |  |  |  |  |
|  |  |  | **Low delay B Main10** |  |  |
|  |  |  | **Over VTM-14.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -1,14% | -1,09% | -1,38% | 99% | 103% |
| Class C | -0,10% | -0,08% | 0,05% | 98% | 98% |
| Class E | -1,06% | -1,29% | -1,12% | 99% | 96% |
| **Overall** | -0,77% | -0,80% | -0,84% | 99% | 99% |
| Class D | 0,07% | 0,10% | -0,61% | 99% | 96% |
| Class F | 0,00% | 0,00% | 0,00% | 101% | 101% |
|  |  |  |  |  |  |
|  |  |  | **Low delay P Main10** |  |  |
|  |  |  | **Over VTM-14.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -1,68% | -1,46% | -1,76% | 99% | 103% |
| Class C | -0,19% | -0,30% | 0,15% | 99% | 99% |
| Class E | -1,29% | -2,00% | -2,60% | 100% | 97% |
| **Overall** | -1,09% | -1,21% | -1,33% | 99% | 100% |
| Class D | -0,02% | -0,23% | -0,40% | 100% | 98% |
| Class F | 0,00% | 0,00% | 0,00% | 99% | 100% |

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

***Issues in VTM 15.x affecting conformance***

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

* Handling of NoOutputOfPriorPicFlag is disabled due to crash issues (issue [#1415](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1415))
* Missing HLS features (see sections below)

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

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

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

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

**HM related activities**

HM 16.25 is expected to be tagged during or shortly after the 25th JVET meeting. Changes include so far:

* JVET-X0079: Add new level 6.3

The following actions have yet to be included:

* JVET-T0050: Add ability to detect static objects to encoder
* [ARSEI Fix functionality](https://vcgit.hhi.fraunhofer.de/jvet/HM/-/merge_requests/52)
* [JVET-X0116: Enabled temporal filter for low-delay configurations and also fixed a bug for non 4:2:0](https://vcgit.hhi.fraunhofer.de/jvet/HM/-/merge_requests/54)
* [JCTVC-AD0021(JVET-T0056) SEI manifest & SEI prefix indication](https://vcgit.hhi.fraunhofer.de/jvet/HM/-/merge_requests/51)

Merge requests are available but have pending discussions.

As reported in the previous report, 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.

**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-15.0 according to 360-degree video CTC (JVET-U2012) compared to that using VTM-14.0 (VTM-14.0 as anchor).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **PERP: VTM-15.0 over VTM-14.0** | | | | | |
|  | **End-to-end**  **WS-PSNR** | | | **End-to-end**  **S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| Class S2 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| **Overall** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |

The following table compares generalized cubemap (GCMP) coding and padded equi-rectangular projection (PERP) coding using VTM-15.0 (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 | -11.42% | -5.70% | -6.33% | -11.39% | -5.64% | -6.28% |
| Class S2 | -3.67% | 0.66% | 0.84% | -3.67% | 0.76% | 0.90% |
| **Overall** | -8.32% | -3.15% | -3.47% | -8.30% | -3.08% | -3.41% |

The following tables are for PERP and GCMP coding comparison between VTM-15.0 and HM-16.22 (HM as anchor), respectively.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **VTM-15.0 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 | -30.17% | -37.21% | -39.72% | -30.17% | -37.26% | -39.72% |
| Class S2 | -36.22% | -35.94% | -38.27% | -36.20% | -35.98% | -38.32% |
| **Overall** | -32.59% | -36.70% | -39.14% | -32.58% | -36.75% | -39.16% |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **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 | -34.88% | -39.43% | -41.50% | -34.83% | -39.36% | -41.45% |
| Class S2 | -39.20% | -38.33% | -40.41% | -39.21% | -38.34% | -40.45% |
| **Overall** | -36.61% | -38.99% | -41.06% | -36.58% | -38.95% | -41.05% |

**SCM related activities**

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

**SHM related activities**

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

**HTM related activities**

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

**HDRTools related activities**

There had not been any updates of HDRTools.

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

**JM, JSVM, JMVM related activities**

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

**Bug tracking**

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

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

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

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

Bug tracking for HDRTools is located at:

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

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

**Software repositories**

Git repositories that were previously assigned to the JCT-VC group on the GitLab server were re-assigned to the JVET group. The old URLs are still working and will forward the user to the new location, with the display of a warning suggesting to update bookmarks to the new location.

The SVN repository for 360Lib was converted to git and development was moved to the GitLab server. Historical branches can still be accessed in the SVN repository.

**CTC alignment and merging**

Merging of CTC documents related to HEVC and VVC in each area of interest should be considered. JVET-W0152 was submitted to the 23rd JVET providing an overview of the existing CTC documents and suggesting merge options.

JVET-Y0112 contains a proposal for merged VTM and HM CTC for SDR 4:2:0 10-bit video by the AHG3 chairs as a first step for merging CTC documents.

**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 was pointed that in the last meeting also a merge proposal for HDR CTC had been submitted, which however would need further study.

[JVET-Y0004](https://jvet-experts.org/doc_end_user/current_document.php?id=11361) JVET AHG report: Test material and visual assessment (AHG4) [V. Baroncini, T. Suzuki, M. Wien, E. François, S. Liu, A. Norkin, A. Segall, P. Topiwala, S. Wenger, Y. Ye]

***Verification test***

Preliminary verification tests for VVC scalable coding has been prepared by InterDigital and performed by IETR/INSA as reported in JVET-Y0047.

***Remote experts viewing***

Remote experts viewing (REV) sessions have been prepared and conducted in the context of AhG11/EE1 Neural network-based video coding and AhG10 deblocking filter tuning. The results are reported in JVET-Y0212.

***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-Y0041, JVET-Y0042, JVET-Y0071, JVET-Y0123 (see table in next section).

**Related contributions**

|  |  |  |  |
| --- | --- | --- | --- |
| [JVET-Y0047](https://jvet-experts.org/doc_end_user/current_document.php?id=11239) | m58568 | AHG4: first report of spatial scalability verification tests | [P. de Lagrange](mailto:philippe.delagrange@interdigital.com), [F. Urban](mailto:fabrice.urban@interdigital.com), [E. François (Interdigital)](mailto:edouard.francois@interdigital.com), [W. Hamidouche (INSA)](mailto:wassim.hamidouche@insa-rennes.fr) |
| [JVET-Y0041](https://jvet-experts.org/doc_end_user/current_document.php?id=11233) | m58476 | AhG-7 Proposed new class of gaming sequences with depth and optical flow information | G. Martin-Cocher, M. Badawi, T. Poirier, S. Puri, K. Naser (Interdigital) |
| [JVET-Y0042](https://jvet-experts.org/doc_end_user/current_document.php?id=11234) | m58477 | AHG 7 modification of and new classes of sequences | G. Martin-Cocher (Interdigital) |
| [JVET-Y0071](https://jvet-experts.org/doc_end_user/current_document.php?id=11265) | m58619 | New Test Content for Video Conferencing Applications | Z. Sinno, G. Desgouttes, [A. M. Tourapis](mailto:atourapis@apple.com), [D. Singer (Apple)](mailto:singer@apple.com) |
| [JVET-Y0123](https://jvet-experts.org/doc_end_user/current_document.php?id=11317) | m58675 | On Test Sequences | [J. Xu](mailto:xujizheng@bytedance.com), L. Zhang (ByteDance), M. Martin-Cocher (InterDigital) |
| [JVET-Y0212](https://jvet-experts.org/doc_end_user/current_document.php?id=11420) | m59021 | AHG4: REV Result for AHG11/EE1 and AHG10/Deblocking | [M. Wien (RWTH)](mailto:wien@lfb.rwth-aachen.de) |

**Recommendations**

The AHG recommends:

* To consider the results of the remote experts viewing sessions in the discussion of deblocking filtering AhG10 and the NN-based video coding tools in AhG11.
* 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 collect volunteers to conduct the verification test, including volunteers to encode.
* 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.

[JVET-Y0005](https://jvet-experts.org/doc_end_user/current_document.php?id=11362) JVET AHG report: Conformance testing (AHG5) [D. Rusanovskyy, I. Moccagatta, F. Bossen, K. Kawamura, T. Hashimoto, H.-J. Jhu, K. Sühring, Y. Yu]

**Timeline**

The progress on the Conformance testing specification is consistent with the preliminary timeline agreed at the 24th 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

**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
* conformance bitstreams for VVC operation range extensions:
  + 51 bitstream categories have been identified
  + Volunteers have been identified to generate bitstreams in all categories
  + Volunteers have been identified to cross-check bitstreams in 47 categories
  + 73 bitstream descriptions have been provided
  + 102 bitstreams of 47 identified categories have been cross-checked and uploaded
  + 20 bitstreams of 4 identified categories have been generated and are waiting to be cross-checked

**Activities and Discussion**

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

VVC activities:

* Three bitstreams (CROP\_A\_Panasonic, TEMPSCAL\_B\_Panasonic, and TEMPSCAL\_C\_Panasonic) have been re-generated with the sequence BQMall because of potential licensing issue of the sequence RitualDance.

There are not currently any known issues with the other provided VVC conformance bitstream packages. A VTM15 directory of bitstreams was made available. All provided bitstreams can be decoded using VTM15.

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.
* All cross-checked bitstreams can be decoded using VTM14.
* Most of the cross-checked bitstreams cannot be decoded using VTM15 because of this error

ERROR: In function "parseConstraintInfo" in VVCSoftware\_VTM/source/Lib/DecoderLib/VLCReader.cpp:4638: gci\_reserved\_zero\_bit not equal to zero

This problem seems related to the bug reported in JVET-Y0190:

JVET-Y0190 abstract: It is asserted that the current text for general constraints information (GCI) in VVC version 2 contains two bugs. For the first bug, it is asserted that incorrect semantics and handling of gci\_num\_additional\_bits lead to GCI decoding behaviour in VVC version 2 incompatible with VVC version 1 decoders. For the second bug, it is asserted that setting gci\_present\_flag to 1 and gci\_num\_additional\_bits to 0 leads to undefined constraint behaviour for the 6 coding tools which are governed by the 6 additional GCI flags introduced in VVC version 2.

* Some packages have minor issues (i.e., file names do not match package name, incorrect format of MD5, etc.). Plan is to fix them when re-generating the streams following resolution of VTM15.0 error.

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**

Two input contributions:

JVET-Y0127 AHG5: Editors' update on conformance testing for VVC operation range extensions [D. Rusanovskyy, T. Hashimoto, H.-J. Jhu, I. Moccagatta, M. G. Sarwer, Y. Yu]

JVET-Y0190 AHG2/AHG8: Suggestions for the operation range extensions GCI [J. Gan, Y. Yu, H. Yu]

**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:

* Enable VVC operation range extensions streams decoding by VTM15.0
* Solicit and collect missing VVC operation range extensions streams bitstream descriptions
* Solicit and identify volunteers to cross-check 20 bitstreams for the remaining 4 VVC operation range extensions categories
* Review the list of conformance bitstreams for VVC operation range extensions and the conformance specification for VVC operation range extensions Draft 2

It is pointed out that JVET-Y0049 is also relating to GCI flags of range extensions, and JVET-Y0057 (on MinCr) might also have impact on the conformance of v2.

[JVET-Y0006](https://jvet-experts.org/doc_end_user/current_document.php?id=11363) 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-3.0:

JVET-X0083 (test 3.3a): AMVP-merge mode

JVET-X0049 (test 3.4a): Lossless software optimization for BDMVR

JVET-X0071 (EE combination 4.1/4.3a/4.4): Chroma bilateral filter, longer filter for CCALF, and alternative band classifier for ALF

JVET-X0056 test 1: Early termination for DMVR and TM

JVET-X0141: CIIP with TIMD and TM merge

JVET-X0090 solution 3: Combination of CIIP, OBMC and LMCS

JVET-X0124: Cleanup signalling of intra template matching

JVET-X0148 (only the TIMD pipeline modification): PDPC handling for TIMD

JVET-X0149: LUT-based derivation of DIMD and TIMD

JVET-X0144 trade-off 2: Max MTT hierarchy depth set by temporal ID

VET-X0121: Fixes for enabling RPR

JVET-X0122 / merge request #16: Fix the shift number of TIMD absInvAngle

JVET-X0156: Fix for histogram of gradients derivation in DIMD mode

Software improvements and bug fixes:

Fix SSE for CTU 256 (MR 18)

Fix issue #1: compiling error when BASE\_ENCODER=1, BASE\_NORMATIVE=0, and TOOLS=0 (MR 33)

Fix mismatching when TM\_MRG=1 with BSE\_NORMATIVE=0 and TOOLS=0, and fix compilation error when MULTI\_PASS\_DMVR=0 (MR 25)

Fix intra cost overflow issue (solves the mismatch with different compilers) (MR 31 + MR 50)

Fix CABAC initial values for CcSaoControlIdc and MultiHypothesisFlag (MR 26)

Fix intra span issue (MR 45)

Fix ALF encoder bug for HDR PQ optimization (MR 46)

Some other lossless code cleanup

The following fixes were integrated on top of ECM-3.0 and released as ECM-3.1:

Fix for LFNST table size (solves the mismatch with different compilers) (MR 52)

ECM-3.0 was tagged on November 15, 2021 and ECM-3.1was tagged on November 21, 2021.

***CTC Performance***

In this section, ECM-3.0 and ECM-3.1 test results following ECM CTC configuration descried in JVET-X2017 are summarized.

The below tables show ECM-3.0 performance over ECM-2.0 anchor.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | |
|  | **Over ECM-2.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -0.24% | -1.75% | -4.82% | 102% | 103% |
| Class A2 | 0.00% | -3.85% | -5.60% | 102% | 105% |
| Class B | -0.14% | -4.23% | -3.41% | 101% | 105% |
| Class C | -0.35% | -1.58% | -1.64% | 98% | 106% |
| Class E | -0.37% | -2.03% | -3.62% | 98% | 105% |
| **Overall** | -0.22% | -2.80% | -3.65% | 100% | 105% |
| Class D | -0.16% | -1.51% | -1.71% | 96% | 106% |
| Class F | -0.33% | -2.64% | -1.84% | 100% | 105% |
| Class TGM | -0.15% | -1.49% | -1.13% | 101% | 103% |
|  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | |
|  | **Over ECM-2.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -1.65% | -0.71% | -2.12% | 108% | 92% |
| Class A2 | -1.71% | -3.84% | -4.81% | 112% | 86% |
| Class B | -1.23% | -3.09% | -2.32% | 109% | 90% |
| Class C | -1.13% | -1.05% | -0.67% | 113% | 90% |
| Class E |  |  |  |  |  |
| **Overall** | -1.38% | -2.22% | -2.34% | 111% | 90% |
| Class D | -0.88% | -1.65% | -1.48% | 110% | 89% |
| Class F | -0.39% | -1.54% | -1.03% | 110% | 94% |
| Class TGM | -0.28% | -1.19% | -0.97% | 109% | 101% |
|  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | |
|  | **Over ECM-2.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -2.17% | -5.17% | -4.97% | 102% | 102% |
| Class C | -0.80% | -3.35% | -2.86% | 102% | 105% |
| Class E | -1.74% | -4.05% | -5.82% | 116% | 101% |
| **Overall** | -1.60% | -4.28% | -4.48% | 106% | 103% |
| Class D | -0.15% | -3.39% | -2.95% | 99% | 107% |
| Class F | -0.56% | -2.63% | -2.22% | 104% | 106% |
| Class TGM | -0.35% | -1.72% | -1.30% | 107% | 106% |

The next tables show ECM-3.1 performance over ECM-2.0 anchor.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | |
|  | **Over ECM-2.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -0.24% | -1.75% | -4.82% | 102% | 103% |
| Class A2 | 0.00% | -3.85% | -5.60% | 101% | 103% |
| Class B | -0.14% | -4.23% | -3.41% | 100% | 104% |
| Class C | -0.35% | -1.58% | -1.64% | 98% | 104% |
| Class E | -0.37% | -2.03% | -3.62% | 98% | 103% |
| **Overall** | -0.22% | -2.80% | -3.65% | 100% | 103% |
| Class D | -0.16% | -1.51% | -1.71% | 95% | 104% |
| Class F | -0.33% | -2.64% | -1.84% | 100% | 104% |
| Class TGM | -0.15% | -1.49% | -1.13% | 100% | 101% |
|  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | |
|  | **Over ECM-2.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -1.65% | -0.71% | -2.11% | 107% | 93% |
| Class A2 | -1.70% | -3.79% | -4.82% | 113% | 88% |
| Class B | -1.23% | -3.10% | -2.36% | 108% | 90% |
| Class C | -1.13% | -1.05% | -0.67% | 110% | 88% |
| Class E |  |  |  |  |  |
| **Overall** | -1.38% | -2.21% | -2.35% | 109% | 89% |
| Class D | -0.88% | -1.65% | -1.48% | 106% | 87% |
| Class F | -0.39% | -1.54% | -1.03% | 108% | 93% |
| Class TGM | -0.28% | -1.19% | -0.97% | 107% | 99% |
|  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | |
|  | **Over ECM-2.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -2.17% | -5.17% | -4.97% | 100% | 102% |
| Class C | -0.80% | -3.35% | -2.86% | 97% | 103% |
| Class E | -1.74% | -4.05% | -5.82% | 112% | 100% |
| **Overall** | -1.60% | -4.28% | -4.48% | 102% | 102% |
| Class D | -0.15% | -3.39% | -2.95% | 95% | 100% |
| Class F | -0.56% | -2.63% | -2.22% | 101% | 104% |
| Class TGM | -0.35% | -1.72% | -1.30% | 103% | 105% |

Next tables show ECM-3.1 performance over VTM-11.0 + JVET-V0056 (<https://vcgit.hhi.fraunhofer.de/chollmann/VVCSoftware_VTM/-/tree/MCTF_VTM11>) anchor**.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | |
|  | **Over VTM-11.0 + V0056** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -6.98% | -12.44% | -16.63% | 311% | 241% |
| Class A2 | -6.44% | -13.13% | -11.99% | 298% | 232% |
| Class B | -6.05% | -13.75% | -14.21% | 338% | 257% |
| Class C | -7.06% | -10.21% | -10.67% | 322% | 252% |
| Class E | -7.58% | -11.55% | -12.41% | 322% | 294% |
| **Overall** | -6.75% | -12.28% | -13.16% | 320% | 255% |
| Class D | -5.86% | -8.45% | -8.23% | 316% | 265% |
| Class F | -10.80% | -15.69% | -15.66% | 244% | 298% |
| Class TGM | -15.62% | -18.73% | -18.25% | 232% | 293% |
|  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | |
|  | **Over VTM-11.0 + V0056** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -14.95% | -16.42% | -21.85% | 367% | 466% |
| Class A2 | -15.84% | -20.38% | -20.42% | 364% | 511% |
| Class B | -13.57% | -19.94% | -19.32% | 384% | 493% |
| Class C | -15.35% | -17.27% | -16.98% | 387% | 427% |
| Class E |  |  |  |  |  |
| **Overall** | -14.77% | -18.61% | -19.42% | 377% | 473% |
| Class D | -16.10% | -17.65% | -16.99% | 378% | 459% |
| Class F | -13.54% | -18.02% | -17.76% | 345% | 407% |
| Class TGM | -14.65% | -18.96% | -19.20% | 348% | 304% |
|  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | |
|  | **Over VTM-11.0 + V0056** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -12.28% | -24.89% | -24.11% | 302% | 331% |
| Class C | -12.51% | -18.81% | -18.94% | 315% | 294% |
| Class E | -12.11% | -17.24% | -19.15% | 298% | 296% |
| **Overall** | -12.31% | -20.95% | -21.14% | 305% | 310% |
| Class D | -14.09% | -20.17% | -19.36% | 303% | 296% |
| Class F | -12.15% | -18.99% | -18.67% | 295% | 324% |
| Class TGM | -14.02% | -21.79% | -22.13% | 269% | 255% |

**AHG related contributions**

[JVET-Y0102](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0102-v1.zip): On the balance of ECM coding gains between luma and chroma

[JVET-Y0113](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0113-v1.zip): Adjusting luma/chroma BD-rate balance in ECM

**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 suggested to also include results for LP

It is also suggested to include the comparison vs. VTM already when the new ECM is announced via the reflector.

[JVET-Y0007](https://jvet-experts.org/doc_end_user/current_document.php?id=11364) JVET AHG report: Low latency and constrained complexity (AHG7) [T. Poirier, S. Liu, L. Wang, J. Xu]

**Overview of input documents and email discussion related to the AHG**

***Sequences***

JVET-Y0041: AhG-7 Proposed new class of gaming sequences with depth and optical flow information, G. Martin-Cocher, M. Badawi, T. Poirier, S. Puri, K. Naser (Interdigital)

JVET-Y0042: AHG 7 modification of and new classes of sequences, G.Martin-Cocher (Interdigital)

The following contribution also propose new test sequences like gaming and video conferencing in relation with AHG7:

JVET-Y0123: On Test Sequences, J. Xu, L. Zhang (ByteDance), M. Martin-Cocher (InterDigital)

JVET-Y0071: New Test Content for Video Conferencing Applications, Z. Sinno, G. Desgouttes, A. M. Tourapis, D. Singer

***Encoder modifications for new gaming sequences***

JVET-Y0101: AHG-7/AHG-10: Depth motion based fast Multi-Type Tree Splitting, S. Puri, K. Naser, T. Poirier, G. Martin-Cocher (Interdigital)

***New CTCs***

JVET-Y0043: AHG-7 LLCC Scenarios and baseline configurations, G. Martin-Cocher, S. Puri, T. Poirier, K. Naser (Interdigital), J. Xu (Bytedance), D. Nicholson (Ektacom), M. Sychev (Huawei), L. Wang (Nokia), S. Liu, W. Yang (Tencent), J.M. Tiesse (VITEC), M. Karczewicz (Qualcomm)

Discussion on JVET-Y0043 on the reflector included additional suggestions:

* Consider whether different ET/RT trade-off may be needed, particularly to avoid or address over-optimization of tools.
* Limiting CTU size to 128x128 when testing 4K sequences in LLCC/low delay configurations.
* Doing some tools on/off test with VTM LDB but limiting the number of reference pictures to 2.
* If the configurations address a complexity constrained (or real-time) encoder environment, this may be considered for RA configuration.

JVET-Y0060: AhG-7 refined LLCC configurations, G. Martin-Cocher, K. Nasser, T. Poirier, S. Puri (Interdigital)

***Tools supporting new CTCs***

JVET-Y0162: AHG 7: Gradual Decoding Refresh for ECM, S. Hong, L. Wang, K. Panusopone (Nokia), T. Poirier, G. Martin-Cocher (InterDigital)

JVET-Y0163: AHG 7: GDR without encoder constraints for ECM, L. Wang, S. Hong, K. Panusopone, M. M. Hannuksela, J. Lainema (Nokia)

**Recommendation**

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.

It is noted that one of the contributions relating to GDR is implementing the same method from VTM into ECM.

[JVET-Y0008](https://jvet-experts.org/doc_end_user/current_document.php?id=11365) JVET AHG report: High bit depth, high bit rate, and high frame rate coding (AHG8) [A. Browne, T. Ikai, D. Rusanovskyy, M. Sarwer, X. Xiu, Y. Yu]

**Activities**

The AHG used the main JVET reflector, jvet@lists.rwth-aachen.de, with [AHG8] in message headers, but no correspondence marked as AHG8 was sent between the 24th and 25th meetings.

The major area of work for the AHG in this meeting cycle were as follows:

* The development of conformance testing in coordination with AHG5. A revision of JVET-X2026 “Conformance testing for VVC operation range extensions (Draft 2)” was uploaded on 14th November 2021 and a notification was sent to the reflector.
* Updates to the draft text for the VVC version 2 extensions in coordination with AHG2. A revision of JVET-X2005 “VVC operation range extensions (Draft 5)” was uploaded on 12th November 2021.

There are 3 related contributions for the 25th meeting.

**Contributions**

[JVET-Y0049](https://jvet-experts.org/doc_end_user/current_document.php?id=11241) “AHG2/AHG8: On the range extensions GCI flags”, Y.-K. Wang

[JVET-Y0127](https://jvet-experts.org/doc_end_user/current_document.php?id=11321) “AHG5: Editors' update on conformance testing for VVC operation range extensions”, D. Rusanovskyy, T. Hashimoto, H.-J. Jhu, I. Moccagatta, M. G. Sarwer, Y. Yu

[JVET-Y0190](https://jvet-experts.org/doc_end_user/current_document.php?id=11397) “AHG2/AHG8: Suggestions for the operation range extensions GCI”, J. Gan, Y. Yu, H. Yu

**Recommendations**

The AHG recommends the following:

* To proceed with finalization of VVC version 2 at this meeting;
* To review all related contributions;
* To continue high bit depth, high bit rate, and high frame rate studies.

It is pointed out that potentially a combination of scalability with extended bit depth could be tested with the VTM software.

[JVET-Y0009](https://jvet-experts.org/doc_end_user/current_document.php?id=11366) JVET AHG report: SEI message studies (AHG9) [J. Boyce, S. McCarthy, C. Fogg, P. de Lagrange, J. Samuelsson, G. J. Sullivan, A. Tourapis, Y.-K. Wang, S. Wenger]

**Related contributions**

A total of 11 contributions, not including cross-checks, are identified relating to AHG9, of which:

* 2 contributions relate to the mandate to study the SEI messages in VSEI, VVC, HEVC, and AVC;
* 1 contribution relates to the mandate to study the draft text JVET-X2006 ;
* 2 contributions relate to the mandate to study signalling of essential resampling phase indication;
* 0 contributions relate to the mandate to collect software and showcase information of SEI messages;
* 5 contributions relate to the mandate to identify potential needs for additional SEI messages;
* 0 contributions relate 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-Y0053](https://jvet-experts.org/doc_end_user/current_document.php?id=11247) AHG9/AHG13: Film grain blending process for film grain characteristics SEI message [Y. He, M. Coban, M. Karczewicz (Qualcomm)]

[JVET-Y0107](https://jvet-experts.org/doc_end_user/current_document.php?id=11301) AHG9: Text improvement for the film grain SEI [E. Thomas (Xiaomi)]

***Study the draft text JVET-X2006 and suggest improvements***

[JVET-Y0050](https://jvet-experts.org/doc_end_user/current_document.php?id=11242) AHG2/AHG9: On the alpha channel information SEI message [Y.-K. Wang (Bytedance)]

***Study signalling of essential resampling phase indication***

[JVET-Y0103](https://jvet-experts.org/doc_end_user/current_document.php?id=11297) AHG9: Down-sample phase indication (SEI message) [P. Bordes, P. de Lagrange, E. Francois (Interdigital)]

[JVET-Y0156](https://jvet-experts.org/doc_end_user/current_document.php?id=11350) AHG9: SEI message with sample phase indication for consistent rendering [F. Bossen, A. Segall (Sharp)]

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

**Contributions related to proposed NNR post-filter SEI message**

[JVET-Y0073](https://jvet-experts.org/doc_end_user/current_document.php?id=11267) AHG9: Colour component description for post-filter purpose SEI message [ T. Chujoh, Y. Yasugi, K. Takada, T. Ikai (Sharp)]

[JVET-Y0074](https://jvet-experts.org/doc_end_user/current_document.php?id=11268) AHG9: Data conversion description for NNR post-filter SEI message [Y. Yasugi, T. Chujoh, K. Takada, T. Ikai (Sharp)]

[JVET-Y0075](https://jvet-experts.org/doc_end_user/current_document.php?id=11269) AHG9: Complexity description for NNR post-filter SEI message [K. Takada, Y. Yasugi, T. Chujoh, T. Ikai (Sharp)]

[JVET-Y0115](https://jvet-experts.org/doc_end_user/current_document.php?id=11309) AHG9: On post-filter SEI [M. M. Hannuksela, M. Santamaria, F. Cricri, E. B. Aksu, H. R. Tavakoli (Nokia)]

**Other contributions related to needs for additional SEI messages**

[JVET-Y0104](https://jvet-experts.org/doc_end_user/current_document.php?id=11298) AHG9: Transparency information SEI for transparent screens [E. Thomas, P. Andrivon, F. Le Leannec, M.-L. Champel (Xiaomi)]

The aspect of [JVET-Y0044](https://jvet-experts.org/doc_end_user/current_document.php?id=11236) (listed below) related to Video Decoding Interface SEI message relates to the mandate to identify needs for additional SEI messages.

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

[JVET-Y0044](https://jvet-experts.org/doc_end_user/current_document.php?id=11236) AHG9: Signalling of Green metadata and Video Decoding Interface SEI messages in VVC specification [E. Francois (InterDigital), Y. He (Qualcomm), C. Herglotz (FAU), Y. Lim (Samsung)]

The aspect of [JVET-Y0044](https://jvet-experts.org/doc_end_user/current_document.php?id=11236) related to Green metadata SEI message relates to the mandate SEI messages defined in HEVC and AVC for potential use in the VVC context.

**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 3 emails related to coordination and 8 emails related to resampling phase indication sent to the JVET reflector during the AHG period.

**Recommendations**

The AHG recommends to:

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

[JVET-Y0010](https://jvet-experts.org/doc_end_user/current_document.php?id=11367) JVET AHG report: Encoding algorithm optimization (AHG10) [P. de Lagrange, R. Sjöberg, A. Tourapis]

**Related contributions**

A total of 14 contributions, not including cross-checks, are identified relating to AHG10, of which:

* 7 relate to the mandate to study tool adaptation and configuration:
  + JVET-Y0083 AHG10: Report of Teleconference on Viewing Session Preparation for Deblocking
  + JVET-Y0085 AHG10: Report of Deblocking filter setting for VTM
  + JVET-Y0177 AHG 10: Enhanced deblocking settings for VTM CTC
  + JVET-Y0102 On the balance of ECM coding gains between luma and chroma
  + JVET-Y0113 Adjusting luma/chroma BD-rate balance in ECM
  + JVET-Y0126 AHG10: VTM encoder configurations for tests targeting improved coding performance
  + JVET-Y0152 AHG10: Fast skip of TT split partitioning on top of ECM reference software
* 1 relate to the mandate to study the impact of non-normative techniques of preprocessing
  + JVET-Y0155 AHG10: Fixes and clean up for temporal prefilter
* 2 relate to the mandate to study optimized encoding for reference picture resampling and scalability modes in VTM:
  + JVET-Y0048 AHG10: study of layer bitrate allocation for spatial scalability in VTM
  + JVET-Y0061 EE1-2.1: Super Resolution with existing VVC functionality
* 2 relate to the mandate to investigate other methods of improving objective and/or subjective quality, including adaptive coding structures and multi-pass encoding:
  + JVET-Y0101 AHG-7/AHG-10: Depth motion based fast Multi-Type Tree Splitting
  + JVET-Y0077 AHG10: Block importance mapping
* 2 relate to the mandate to study rate control and rate-distortion optimization
  + JVET-Y0105 AHG10: An improved VVC rate control scheme
  + JVET-Y0118 AHG10: On Temporal-Layer-Based ChromaQP Coding

***Tool adaptation and configuration***

**JVET-Y0083 AHG10: Report of Teleconference on Viewing Session Preparation for Deblocking**

This document reports the teleconference on viewing session preparation for deblocking filter settings from JVET-X0063, on 2021-11-17 16:00 – 17:15 UTC. Testing sequences and testing conditions were discussed. A list of test sequences with recommended QPs is provided, for RA and LBD configuration. Length of 5s is suggested, and LDB is preferred. A traditional comparison scale is suggested, but that could be refined if very little difference is seen. Test session was suggested to take place just before or during the first days of the meeting.

**JVET-Y0085 AHG10: Report of Deblocking filter setting for VTM**

This proposal reports objective metrics and internal viewing results for temporal-ID dependent deblocking filter settings (Beta and Tc offsets), as proposed in JVET-X0063 and similar to what is done in ECM. Gain in RA is around 0.3% using PSNRY and expert viewing concludes to no visual degradation. Other learning-based objective metrics confirm gains in most cases, and no severe degradation.

**JVET-Y0177 AHG 10: Enhanced deblocking settings for VTM CTC**

This contribution proposes to change the beta offset setting of DBF to -2 in the CTCs. This parameter controls the amount of spatial activity that is allowed until deblocking is turned off. It is said that it improves objective performance while maintaining subjective quality. The impact on PSNRy BD-rate is -0.25% for RA and -0.43% for LDB.

**JVET-Y0102 On the balance of ECM coding gains between luma and chroma**

Thanks to the introduction of new coding tools, in particular cross-component tools, ECM coding gain over VVC is larger in chroma than in luma.

This contribution proposes to modify ECM encoder configurations in AI, RA and LB to balance the chroma and luma coding gains, through modified chroma QP mapping tables. Two encoder configurations are tested, leading to the following results (for Y/Cb/Cr) over VTM-14.0.

* ECM-3.1: AI: -6.8% / -12.3% / -13.2%; RA: -14.6% / -18.6% / -19.4%; LB : -12.3% / -21% / -21%
* config. 1: AI: -8.1% /-2.2% /-3.4%; RA: -15.4% / -11.5% / -12.8%; LB: -13% / -6.4% / -5.9%
* config. 2: AI: -7.29%/-7.93%/-8.71%; RA: (not yet available) LB : -12.9% / -9.8% / -9.5%

It is proposed to adopt configuration 1 for AI and RA, and configuration 2 for LB

**JVET-Y0113 Adjusting luma/chroma BD-rate balance in ECM**

This contribution reports experiments based on ECM-3.1 aiming at obtaining a better repartition of BD-rate variations between luma and chroma, when using VTM11.0+V0056 as reference. As a result, the contribution proposes adjustments of the configuration parameter “LMCSOffset” for the four configurations of the ECM CTCs for SDR content.

For AI configuration, a LMCSOffset value of -6 or -7 provides a good balance between luma and chroma BDR variations. It can be noted that the balance looks also good for class F and TGM content.

For RA and LDB configuration, a LMCSOffset value of -7 provides a better balance between luma and chroma BDR variations. For RA, this is a 1-step change from the reference -6 and provides around 0.2% more gain on PSNRY.

**JVET-Y0126 AHG10: VTM encoder configurations for tests targeting improved coding performance**

In this contribution, a set of encoder-only modifications and settings targeting improved compression efficiency of VTM reference software is proposed. Proposed method is implemented as minor encoder software change on top VTM-14.2 and encoder configuration files for RA and LD. It is reported that under the CTC, proposed method provides the following bd-rate gain vs. VTM-14.2 anchor:

* RA: -2.66%/-6.24%/-6.13% with 168%EncT/96%DecT

It is proposed to adopt SW change to the VTM reference software and include configuration files for tests targeting improved coding performance, for example for benchmarking purposes. No changes to VTM CTC are requested, since those settings are intended for benchmarking purpose where encoding time constraint may be relaxed.

Changes consist in: higher intra QP offset, deeper MTT hierarchy in intra picture, more MTS candidates, DBF in RDO, disabling or changing some fast decisions, activating MCTF for POCs that are multiple of 4, lower strength in DBF, and slight adjustment of GOP-dependent QP offset.

**JVET-Y0152 AHG10: Fast skip of TT split partitioning on top of ECM reference software**

The contribution presents a fast method of conditionally skipping ternary-tree (TT) split on top of ECM-3.0 reference software. The performance loss on top of ECM, for RA and PSNRY is between 0.09% (with ‑6% encoding time) and 0.13% (with -9% encoding time). This work is derived from JVET-W0086, where a fast partitioning method was proposed on top of ECM to speed up the search for UBT and UQT; the method is also applicable to TT.

***Preprocessing***

**JVET-Y0155 AHG10: Fixes and clean up for temporal prefilter**

This contribution proposes bug or inconsistency fixes to the temporal prefilter (MCTF), mostly related to boundary conditions at the beginning and end of sequences and to the computation of a noise estimate. It reports BD-rate improvements of 0.2% for RA and 0.1% for LD configuration.

***RPR and scalability***

**JVET-Y0048 AHG10: study of layer bitrate allocation for spatial scalability in VTM**

This contribution is related to JVET-Y0047 the reports results of VVC spatial scalability visual performance tests, with equal bitrate for HD base layer and UHD enhancement layer. Since it was noticed that the bitrate balance has a significant effect on coding performance, in favor of more bitrate in the base layer, this contribution reports experiments with various bitrate allocations, and reports gains and lower encoding time over single layer in some cases.

**JVET-Y0061 EE1-2.1: Super Resolution with existing VVC functionality**

This contribution reports the performance of RPR tool in VVC if tested for providing super-resolution functionality (test 2.1 in EE1). Experiments are limited to scaling factor 2.0 only. In average under AhG11 test conditions gains (in Y-PSNR) of 1.2 % in random access and 0.6% in all intra configuration can be achieved. Gain is higher in MS-SSIM metric (2.9% in random access and 1.9% in all intra configuration). It is mentioned that these results were obtained by multi-pass encoding (encoding with and w/o RPR, with different QPs and choosing the best).

***Other methods***

**JVET-Y0101 AHG-7/AHG-10: Depth motion based fast Multi-Type Tree Splitting**

This contribution presents a fast method that uses depth and motion information (e.g. from a game engine, as in the new class of gaming content described in JVET-Y0041) to determine the Multi-Type Tree (MTT) split on the encoder side without expensive RDO. The proposed non-normative method is applied under the low delay/Low Latency Controlled Complexity (LLCC) configurations where the depth and motion information is used to achieve coding gains with negligible complexity increase.

Gain in LDP for PSNRY is:

* 0.92% with +13% encoding time on top of VTM-14.0 with MTT disabled
* 1.48% with +22% encoding time on top of VVenC lowdelay faster preset

**JVET-Y0077 AHG10: Block importance mapping**

This contribution proposes an updated version of the method in contribution JVET-V0057.

The proposed algorithm for VTM and HM signals CTU QP delta values in pictures that will be used for reference. The QP value to use for each CTU is based on the estimated importance of a given CTU for future pictures and the QP selected is in the range of −2 to +2 relative to the picture QP. Here importance of CTU means how well the quality is supposed to propagate to other pictures, which is estimated from metrics given by the MCTF (motion-compensated SSD and variance) for adjacent pictures.

Gains of around 2% in RA CTCs are reported for VTM and HM, and adoption in VTM and HM software is proposed.

***Rate control***

**JVET-Y0105 AHG10: An improved VVC rate control scheme**

This contribution presents some improvements based on the current R-lambda rate control algorithm. Three changes were made: CTU-level bit allocation, different for skip and non-skip, smooth window size base on GOP size and intra period, and extension to low frame rate of quality dependency factor from JVET-M0600.

Reported objective gain is 0.65% for LDB, and not yet fully available for RA at the time of writing this report, but should be similar or better.

**JVET-Y0118 AHG10: On Temporal-Layer-Based ChromaQP Coding**

This contribution proposes to improve coding efficient through a temporal-layer based chroma QP adjustment, for both VTM and ECM. Full simulation results were not yet available at the time of this writing. The contribution also suggests creating an ad-hoc group that covers both VTM and ECM, to study temporal-layer based settings (partitioning, deblocking, and chroma-QP) along with VPS and SPS bit efficiency related to the topic.

**Recommendation**

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

[JVET-Y0011](https://jvet-experts.org/doc_end_user/current_document.php?id=11368) JVET AHG report: Neural network-based video coding (AHG11) [E. Alshina, S. Liu, A. Segall, J. Chen, F. Galpin, J. Pfaff, S. S. Wang, Z. Wang, M. Wien, P. Wu, J. Xu]

1. **Activities**

The AHG used the main JVET reflector, [jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de), for email exchange with AHG11 included in the subject lines. Three emails were exchanged on the reflector.

***EE Coordination***

The AHG finalized, conducted and discussed the EE on NN based video coding. The final version of the EE description was uploaded to the document repository on November 2, 2021.

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

|  |  |  |
| --- | --- | --- |
| JVET-X0023 | EE1: Summary of Exploration Experiments on Neural Network-based Video Coding | E. Alshina, S. Liu, W. Chen, F. Galpin, Y. Li, Z. Ma, H. Wang |

***CTC Refinement and Support***

The AHG refined and released the CTC test conditions on November 3, 2021.

As agreed in the previous meeting, the CTC test conditions included the following updates:

1. Clarification of the super-resolution test conditions
2. Clarification of the units for reporting multiply-accumulate operations
3. Improvements in the reporting template

***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 coordinated with SC29/AG5 to prepare a viewing procedure for EE contributions. With close coordination with SC29/AG5, remote viewing sessions have been performed on Jan 10th and 11th to understand the visual benefit of the approaches as input to the 25th meeting.

***Technical Evaluation***

The AHG made meaningful progress on the mandate to evaluate and quantify potential NN based video coding technologies. A summary of the non-EE contribution provided as input to the 24th meeting is provided below:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Title** | **Common Test Conditions** | **Results** | | | **Training Data** | |
| **RA** | **LDB** | **AI** | **CTC** | **Additional** |
| **Loop Filter** | | | | | | | |
| JVET-Y0046 | AHG11: ALF improvement for NNVC | Yes | Yes | No | Yes | BVI-DVC | DIV2K |
| JVET-Y0052 | AHG11: CNN post-processing filter based on depthwise separable convolution and attention mechanism | Yes | Class C, D | Class C, D | Class C, D | BVI-DVC | DIV2K |
| JVET-Y0081 | AHG11: Transformer based in-loop filtering | Yes | No | No | Yes |  | DIV2K |
| JVET-Y0086 | AHG11: A Unet-Based Deep In-Loop Filter | Yes | No | No | Yes | BVI-DVC (4K) | DIV2K |
| JVET-Y0079 | EE1-1.1-related: the result of neural network based in-loop filter on ECM | No (ECM) | Partial | Partial | Yes | BVI-DVC, TVD | DIV2K |
| JVET-Y0080 | EE1-1.1-related: alternative filter designs | Yes | Yes | Yes | Yes | BVI-DVC, TVD | DIV2K |
| JVET-Y0098 | EE1-related: Combination of VVC deblocking and NN loop filtering | No | Yes | Yes | Yes | BVI-DVC | DIV2K |
| **Post Filtering** | | | | | | | |
| JVET-Y0059 | AHG11: Content-adaptive post-processing filter | Yes | Yes | No | No | BVI-DVC | DIV2K |
| **Super-Resolution** | | | | | | | |
| JVET-Y0068 | EE1-2.1-related: RPR encoder with multiple scale factors | No  (VTM-13.0) | Yes | No | Yes | - | - |
| JVET-Y0087 | AHG11: An Improved CNN-based Super Resolution Method | Yes | Class A1, A2 | - | Class A1, A2 | BVI-DVC (4K) |  |
| **Inter-Prediction** | | | | | | | |
| JVET-Y0090 | AHG11: Neural Network Based Motion Compensation Enhancement for Video Coding | Yes | Class B, C | No | No | BVI-DVC |  |
| JVET-Y0096 | AHG11: NN-based Reference Frame Interpolation for VVC Hierarchical Coding Structure | Yes | Yes | No | No | BVI-DVC | Vimeo |
| **Combined Coding** | | | | | | | |
| JVET-Y0111 | AhG11: Hybrid Conventional/Deep-learning-based image coding | No (ECM3.1) | - | - | - | - | - |
| **End-to-End** | | | | | | | |
| JVET-Y0051 | AHG11: Deep omnidirectional video compression in YUV domain | No  (360) | - | - | - |  | VQA-DOV |

1. **Input contributions**

There are 32 input contriubtions related to the AHG mandates. Fourteen of the contributions are part of the EE activity, while the remaining 18 contributions are related to AHG11 but not part of the EE. The list of input contributions is provided below.

***EE Input Contributions***

|  |  |  |
| --- | --- | --- |
| **Reporting** | | |
| JVET-Y0023 | EE1: Summary of Exploration Experiments on Neural Network-based Video Coding | E. Alshina, S. Liu, W. Chen, F. Galpin, Y. Li, Z. Ma, H. Wang |
| **EE Technology** | | |
| JVET-Y0061 | EE1-2.1: Super Resolution with existing VVC functionality | E. Alshina, J. Sauer (Huawei) |
| JVET-Y0069 | EE1-2.3: CNN-based Super Resolution for Video Coding Using Decoded Information | C. Lin, Y. Li, K. Zhang, L. Zhang (Bytedance) |
| JVET-Y0070 | EE1-2.4: CNN-based Super Resolution for Video Coding Using Separate Networks for Chroma Components | C. Lin, Y. Li, K. Zhang, L. Zhang (Bytedance) |
| JVET-Y0078 | EE1-1.1: neural network based in-loop filter with constrained storage and low complexity | L. Wang, X. Xu, S. Liu (Tencent) |
| JVET-Y0082 | EE1-3.1: Intra prediction using neural networks | T. Dumas, F. Galpin, P. Bordes, E. François (Interdigital) |
| JVET-Y0084 | EE1-1.3: A Deep In-Loop Filter | X. Zhang, D. Jiang, J. Lin, C. Fang, S. Peng (Dahua) |
| JVET-Y0143 | EE1-1.2: Test on Deep In-Loop Filter with Adaptive Parameter Selection and Residual Scaling | Y. Li, K. Zhang, L. Zhang (Bytedance), H. Wang, K. Reuze, A.M. Kotra, M. Karczewicz (Qualcomm) |
| **Cross Checks** | | |
| JVET-Y0166 | Crosscheck of EE1-1.2.1 from JVET-Y0143 (EE1-1.2: Test on Deep In-Loop Filter with Adaptive Parameter Selection and Residual Scaling) | J. Ström (Ericsson) |
| JVET-Y0169 | Crosscheck of JVET-Y0069 (EE1-2.3: CNN-based Super Resolution for Video Coding Using Decoded Information) | Johannes Sauer |
| JVET-Y0173 | Crosscheck of JVET-Y0143 (EE1-1.2: Test on Deep In-Loop Filter with Adaptive Parameter Selection and Residual Scaling) | Johannes Sauer |
| JVET-Y0186 | Cross-check of JVET-Y0143 (Test 1.2.2): EE1-1.2: Test on Deep In-Loop Filter with Adaptive Parameter Selection and Residual Scaling | Z. Dai(OPPO) |
| JVET-Y0187 | Cross-check of JVET-Y0082 (Test 3.1.2): EE1-3.1: Intra prediction using neural networks | L. Xu, Y. Yu, Z. Dai(OPPO) |
| JVET-Y0188 | Cross-check of JVET-Y0078 (Test 1.1): EE1-1.1: Neural network based in-loop filter with constrained storage and low complexity | Z. Xie(OPPO) |

***Non-EE Input Contributions***

|  |  |  |
| --- | --- | --- |
| **Reporting** | | |
| JVET-Y0011 | JVET AHG report: Neural network-based video coding (AHG11) | E. Alshina, S. Liu, A. Segall, J. Chen, F. Galpin, J. Pfaff, S. S. Wang, Z. Wang, M. Wien, P. Wu, J. Xu |
| JVET-Y0045 | AhG11/EE1 viewing preparation report | E. Alshina, M. Wien, A. Segall |
| JVET-Y0212 | AHG4: REV Result for AHG11/EE1 and AHG10/Deblocking | M. Wien (RWTH) |
| **Loop Filtering** | | |
| JVET-Y0046 | AHG11: ALF improvement for NNVC | W. Zou, Y. Zhou (Xidian Univ.), C. Huang, Y. X. Bai (ZTE) |
| JVET-Y0052 | AHG11: CNN post-processing filter based on depthwise separable convolution and attention mechanism | H. Zhang, C. Jung (Xidian University), D. Zou, M. Li (OPPO) |
| JVET-Y0081 | AHG11: Transformer based in-loop filtering | T. Ouyang, H. Wang, H. Zhu, Z. Chen (Wuhan University) |
| JVET-Y0086 | AHG11: A Unet-Based Deep In-Loop Filter | X. Zhang, D. Jiang, J. Lin, C. Fang, S. Peng (Dahua) |
| JVET-Y0079 | EE1-1.1-related: the result of neural network based in-loop filter on ECM | L. Wang, X. Xu, S. Liu (Tencent) |
| JVET-Y0080 | EE1-1.1-related: alternative filter designs | L. Wang, X. Xu, S. Liu (Tencent) |
| JVET-Y0098 | EE1-related: Combination of VVC deblocking and NN loop filtering | K.Andersson, J. Ström, D. Liu, R. Sjöberg (Ericsson) |
| **Post Filtering** | | |
| JVET-Y0059 | AHG11: Content-adaptive post-processing filter | M. Santamaria, J. Lainema, F. Cricri, R. G. Youvalari, H. Zhang, G. Rangu, H. R. Tavakoli, H. Afrabandpey, M. M. Hannuksela (Nokia) |
| **Super Resolution** | | |
| JVET-Y0068 | EE1-2.1-related: RPR encoder with multiple scale factors | J. Nam, S. Yoo, J. Lim, S. Kim (LGE) |
| JVET-Y0087 | AHG11: An Improved CNN-based Super Resolution Method | S. Peng, D. Jiang, J. Lin, C. Fang, X. Zhang (Dahua) |
| **Inter Prediction** | | |
| JVET-Y0090 | AHG11: Neural Network Based Motion Compensation Enhancement for Video Coding | C. Ma, R.-L. Liao, Y. Ye (Alibaba) |
| JVET-Y0096 | AHG11: NN-based Reference Frame Interpolation for VVC Hierarchical Coding Structure | Z. Liu, X. Xu, S. Liu (Tencent), Y. Guo, Z. Chen (Wuhan Univ.) |
| **Combined Coding** | | |
| JVET-Y0111 | AhG11: Hybrid Conventional/Deep-learning-based image coding | F. Galpin, T. Dumas, P. Bordes, F. Racapé, E. François (InterDigital), Y. Li, Kai Zhang, Li Zhang (Bytedance), H. Wang, K. Reuze, A.M. Kotra, M. Karczewicz (Qualcomm), |
| **End-to-End** | | |
| JVET-Y0051 | AHG11: Deep omnidirectional video compression in YUV domain | Qipu Qin, Cheolkon Jung (Xidian University), Dan Zou, Ming Li (OPPO) |
| **Software** | | |
| JVET-Y0110 | AHG11: Small Ad-hoc Deep-Learning Library (SADL) update | F. Galpin, F. Mom, T. Dumas, P. Bordes, P. Nikitin, E. François (InterDigital) |

1. **Recommendations**

The AHG recommends:

* Review all input contributions.
* Discuss if DIV2K should be formally included in the training set.
* Continue investigating neural network-based video coding tools, including coding performance and complexity.

Number of cross-checks significantly increased compared to previous meetings.

[JVET-Y0012](https://jvet-experts.org/doc_end_user/current_document.php?id=11369) JVET AHG report: Enhanced compression beyond VVC capability (AHG12) [M. Karczewicz, Y. Ye, L. Zhang, B. Bross, X. Li, K. Naser, H. Yang]

1. **Activities**

The Common Test Conditions were updated to fix mislabelled sequences in the excel sheet and clarify configurations used for SCC (JVET-X2017). The primary activity of the AHG was the “Exploration experiment on enhanced compression beyond VVC capability” (JVET-X2024). The indication of the achievable improvement over VVC are the results of the EE2 software base (ECM-3.0). The combined improvement of the ECM-3.0 over VTM-11.0 with the improved MCTF from JVET-V0056 for AI, RA and LB configurations are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | All Intra Main10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -6.98% | -12.44% | -16.63% | 311% | 241% |
| Class A2 | -6.44% | -13.13% | -11.99% | 298% | 232% |
| Class B | -6.05% | -13.75% | -14.21% | 338% | 257% |
| Class C | -7.06% | -10.21% | -10.67% | 322% | 252% |
| Class E | -7.58% | -11.55% | -12.41% | 322% | 294% |
| Overall | -6.75% | -12.28% | -13.16% | 320% | 255% |
| Class D | -5.86% | -8.45% | -8.23% | 316% | 265% |
| Class F | -10.80% | -15.69% | -15.66% | 244% | 298% |
| Class TGM | -15.62% | -18.73% | -18.25% | 232% | 293% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Random Access Main 10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -14.95% | -16.42% | -21.85% | 367% | 466% |
| Class A2 | -15.84% | -20.38% | -20.42% | 364% | 511% |
| Class B | -13.57% | -19.94% | -19.32% | 384% | 493% |
| Class C | -15.35% | -17.27% | -16.98% | 387% | 427% |
| Class E |  |  |  |  |  |
| Overall | -14.77% | -18.61% | -19.42% | 377% | 473% |
| Class D | -16.10% | -17.65% | -16.99% | 378% | 459% |
| Class F | -13.54% | -18.02% | -17.76% | 345% | 407% |
| Class TGM | -14.65% | -18.96% | -19.20% | 348% | 304% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Low Delay B Main 10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -12.28% | -24.89% | -24.11% | 302% | 331% |
| Class C | -12.51% | -18.81% | -18.94% | 315% | 294% |
| Class E | -12.11% | -17.24% | -19.15% | 298% | 296% |
| Overall | -12.31% | -20.95% | -21.14% | 305% | 310% |
| Class D | -14.09% | -20.17% | -19.36% | 303% | 296% |
| Class F | -12.15% | -18.99% | -18.67% | 295% | 324% |
| Class TGM | -14.02% | -21.79% | -22.13% | 269% | 255% |

1. **Contributions**

In addition to 18 EE2 contributions, 42 related contributions were received which can be subdivided as follows:

***Partitioning (1)***

JVET-Y0174, "EE2-1.1-related: additional tests on partitioning flexibility", F. Urban, K. Naser, F. Galpin (Interdigital), K. Zhang, L. Zhang, Z. Deng, N. Zhang, Y. Wang (Bytedance)

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

JVET-Y0148, "Non-EE2: Spatial-Temporal Adaptive Loop Filter", W. Yin, K. Zhang, Y. Li, H. Liu, L. Zhang (Bytedance)

***Intra (11)***

JVET-Y0055, "AHG12: Slope adjustment for CCLM", J. Lainema, A. Aminlou, P. Astola, R. G. Youvalari (Nokia)

JVET-Y0092, "Non-EE2: On chroma intra prediction mode", X. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba)

JVET-Y0097, "AhG12: Removed DIMD from MPM list of TIMD", K. Naser, T. Dumas, Y. Chen, F. Galpin (InterDigital)

JVET-Y0109, "AhG12: Neural Network-based intra prediction", T. Dumas, F. Galpin, P. Bordes, F. Mom, E. Francois (InterDigital)

JVET-Y0119, "EE2-related: On Extended MRL Intra Prediction", K. Sato, Y. Yu, H. Yu, Z. Xie, L. Xu, F. Wang, H. Huang, J. Gan, D. Wang (OPPO)

JVET-Y0130, "EE2-related : Unification of availability check for intra mode coding", S. Yoo, H. Jang, J. Nam, J. Choi, J. Lim, S. Kim (LGE)

JVET-Y0131, "EE2-related : Clean-up on DIMD", S. Yoo, H. Jang, J. Nam, J. Choi, J. Lim, S. Kim (LGE)

JVET-Y0140, "EE2-related: On the LCU boundary processing by intra-prediction tools", A. Filippov, V. Rufitskiy, D. Ruiz Coll (Ofinno)

JVET-Y0144, "Non-EE2: DIMD Flag Signalling Clean-up", J. Zhao, S. Kim (LGE)

JVET-Y0149, "EE2-related: Modifications of the extended MRL candidate list", A. Filippov, V. Rufitskiy, E. Dinan (Ofinno)

JVET-Y0203, "EE2-related: a combination of the extended MRL candidate list (JVET-Y0149) modifications with the extended LCU boundary processing area (JVET-Y0140)", A. Filippov, V. Rufitskiy, D. Ruiz Coll, E. Dinan (Ofinno)

***Inter and RPR (12)***

JVET-Y0076, "Non-EE2: Template Matching-based OBMC Design",Z. Lv, C. Zhou, J. Zhang (vivo)

JVET-Y0089, "Non-EE2: DMVR with BCW enabled", P. Bordes, A. Robert, Y. Chen, F. Galpin (Interdigital)

JVET-Y0091, "Non-EE2: MVP refinement for regular AMVP mode", C. Zhou, Z. Lv, J. Zhang (vivo)

JVET-Y0120, "EE2-related: Non-adjacent temporal MVP", F. Wang, Z. Xie, Y. Yu, H. Yu, L. Xu, K. Sato, J. Gan, D. Wang (OPPO)

JVET-Y0125, "AHG12: Enhanced bi-directional motion compensation", Y.-W. Chen, C.-W. Kuo, N. Yan, W. Chen, X. Xiu, X. Wang (Kwai Inc.)

JVET-Y0129, "Non-EE2: MVD and merge index signaling of AMVP-merge mode", Z. Zhang, H. Huang, C.-C. Chen, V. Seregin, M. Karczewicz (Qualcomm)

JVET-Y0135, "Non-EE2: 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-Y0139, "Non-EE2: On the extended number of active reference pictures and reference picture reordering", H. Huang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-Y0151, "Non-EE2: Adaptive re-ordering of merge candidates with refined motion", Y. Wang, K. Zhang, N. Zhang, Z. Deng, L. Zhang (Bytedance)

JVET-Y0154, "AHG12: Bilinear Interpolation Filtering for ARMC", W. Chen, X. Xiu, H. Gao, Y.-W. Chen, X. Wang (Kwai)

JVET-Y0161, "EE2-3.12-related: Extensions of history-parameter-based affine model inheritance", K. Zhang, L. Zhang, Z. Deng, N. Zhang, Y. Wang (Bytedance)

JVET-Y0172, "Non-EE2: Long tap interpolation filtering on chroma components", X. Xie, K. Zhang, L. Zhang, Junru Li, Meng Wang, Shiqi Wang (Bytedance)

***RPR (2)***

JVET-Y0095, "Non-EE2: RPR with luma-only re-scaling", P. Bordes, F. Galpin, E. Francois (Interdigital)

JVET-Y0128, "Non-EE2: fixing issues for RPR enabling and non-CTC configuration in ECM", Z. Zhang, H. Huang, C.-C. Chen, V. Seregin, M. Karczewicz (Qualcomm), P.Bordes (InterDigital)

***Entropy and Coefficient Coding (5)***

JVET-Y0114, "Non-EE2: Dependent quantization with 4 states for chroma components", Y. Chen, E. Francois, F. Galpin, P. de Lagrange (InterDigital)

JVET-Y0121, "EE2-4.2-related: On adaptive sign prediction position selection", L. Xu, Y. Yu, H. Yu, Z. Xie, F. Wang, D. Wang (OPPO)

JVET-Y0141, "EE2-4.3 related: More combined test results for sign prediction", J. Chen, Y. Ye, R.-L. Liao, X. Li (Alibaba), X. Xiu, Y.-W. Chen, N. Yan, C.-W. Kuo, H.-J. Jhu, W. Chen, X. Wang (Kwai)

JVET-Y0157, "AHG12: Improved probability estimation for CABAC", X. Xiu, Y.-W. Chen, W. Chen, H. Gao, H.-J. Jhu, C.-W. Kuo, X. Wang (Kwai)

JVET-Y0181 "AHG12: CABAC initialization from previous inter slice", V. Seregin, J. Dong, N. Hu, M. Karczewicz (Qualcomm)

***Transform Coding (1)***

JVET-Y0159, "EE2-related: inter MTS refinement on adaptive intra MTS (EE2-4.4)", T. Hashimoto, T. Ikai (Sharp)

***Screen Content Coding (5)***

JVET-Y0062, "Non-EE2: Cross-component palette coding", B. Vishwanath, K. Zhang, L. Zhang (Bytedance)

JVET-Y0088, "EE2-related: IBC with Template Matching", A. Robert, K. Naser, T. Poirier, Y. Chen, F. Galpin (InterDigital)

JVET-Y0124, "Non-EE2: Intra Block Copy with An Extended Reference Area", J. Xu (ByteDance)

JVET-Y0133, "EE2-related: BVP candidate adjustment based on IBC reference region implemented on top of test EE2-3.13", D. Ruiz Coll, A. Filippov, V. Rufitskiy, T.M. Bae (Ofinno)

JVET-Y0160, "EE2-3.13-related: Enlarged HMVP table for IBC", N. Zhang, K. Zhang, L. Zhang (Bytedance)

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

JVET-Y0102, "On the balance of ECM coding gains between luma and chroma", F. Le Léannec, P. Andrivon, E. Thomas(Xiaomi)

JVET-Y0113, "Adjusting luma/chroma BD-rate balance in ECM", Y. Chen, E. Francois, P. Nikitin (InterDigital)

JVET-Y0108, "AhG3/AhG12 Modification of JVET CTC for environmental considerations", F. Galpin, M. Radosavljevic, E. Francois (InterDigital)

JVET-Y0117, "AhG12: ECM coding performance for HDR/WCG content and suggested common test conditions", T. Lu, F. Pu, P. Yin, S. McCarthy, W. Husak, T. Chen (Dolby), N. Hu, V. Seregin, M. Karczewicz (Qualcomm)

1. Recommendations

The AHG recommends to:

* To review all the related contributions.
* To encourage contribution of new test sequences, especially 8k resolution, which might be included in future test conditions.

[JVET-Y0013](https://jvet-experts.org/doc_end_user/current_document.php?id=11370) JVET AHG report: Film grain technologies (AHG13) [W. Husak, M. Radosavljević, W. Wan, D. Grois, A. Tourapis]

**Related contributions**

Five contributions related to AHG13 were identified as of 01/11/2022:

* Two were related to the AHG and the deliverables:
  + JVET-Y0013 JVET AHG report: Film grain technologies (AHG13)
  + JVET-Y0158 AHG13: Proposed Draft Film Grain Technical Report Text
* One related to a bug fix:
  + JVET-Y0107 AHG9: Text improvement for the film grain SEI
* Two related to proposing improvements to the existing film grain blending process:
  + JVET-Y0053 AHG9/AHG13: Film grain blending process for film grain characteristics SEI message
  + JVET-Y0165 Crosscheck of JVET-Y0053 (AHG9/AHG13: Film grain blending process for film grain characteristics SEI message)

***Contributions***

**JVET-Y0158 AHG13: Proposed Draft Film Grain Technical Report Text**

This document is the draft text for the proposed technical report. The document is in Working Draft form and will be reviewed during the 25th JVET meeting.

**JVET-Y0107 AHG9: Text improvement for the film grain SEI**

This contribution suggests a text improvement for the description text and formula of the film grain SEI message. In the formula, it appears that the value ranges for the variable x and y may not be correct for all the values of c from 0 to 2.

**JVET-Y0053 AHG9/AHG13: film grain blending process for film grain characteristics SEI message**

This contribution proposes a film grain blending process for film grain characteristics SEI message. The blending process is identical to the process specified in SMPTE RDD5 except the modifications to reduce line buffers and on-chip memory storage.

The contribution proposes two independent modifications:

1. The grain block size is changed from fixed 8x8-size in RDD5 to 8x1 for picture resolutions less than or equal to 1920x1080, 16x1 for resolution greater than 1920x1080 and up to 3840x2160, and 32x1 for resolutions greater than 4K including 8K. Using Nx1 block size eliminates the need of line buffers to compute block average in the existing methods.
2. Add note to limit the number of film grain patterns per picture to 1 to reduce on-chip memory requirement.

**Recommendation**

The AHG recommends that the related input contributions are reviewed and to further continue the study of film grain technologies in JVET.

Is there any objective metric for evaluating film grain? No.

# Project development (46)

## Deployment and advertisement of standards (2)

Contributions in this area were discussed in session 20 at 2215–2230 UTC on Wednesday 19 Jan. 2022 (chaired by JRO).

Formularende[JVET-Y0020](https://jvet-experts.org/doc_end_user/current_document.php?id=11243) Deployment status of the HEVC standard [G. J. Sullivan]

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. Following changes are relevant relative to JVET-X0020-v6 of October 2021.

As of August 2021, a developer survey (conducted from 7 July to 22 August 2021) by Bitmovin with 538 respondents from 65 countries (primarily with technical roles) reported [30]:

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

b. 25% of video developers planning to add HEVC in the next 12 months

Microsoft Azure Media Services began supporting HEVC encoding in its Premium Encoder as of September 2017, and offered general availability in its Standard Encoder as of March 2021. Video encoding support for hardware-accelerated HEVC encoding in DirectX 12 for Windows 11 was announced in December 2021.

Nikon Z9: Profession-grade mirrorless camera with all-electronic shutter and full-frame sensor, supports HEVC with 8K 10 bit video at 30 fps and 4K and Full HD at 120 fps (announced October 2021, availability December 2021)

ETSI/DVB Technical Specification 101 154: Approval of HEVC for 8K service was reported in November 2021.

[JVET-Y0021](https://jvet-experts.org/doc_end_user/current_document.php?id=11244) Deployment status of the VVC standard [G. J. Sullivan]

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

Changes relative to JVET-X0021-v2 of October 2021 are as follows.

Version 1.3.1 of VVenC was released in December 2021. Some of its changes include 10% speed-up of the “slower” mode, coding efficiency improvement of the “faster” mode (5% for HD and 3% for UHD), rate control improvements, improved QP adaptation for subjective quality, tile parallelism support, and MCTF improvement (ported from VTM). Version 1.3.0 of VVdeC was released in December 2021. Main changes reported for VVdeC include improved tune-in behaviour, 3× memory reduction, MinGW support, general cleanups, and WebAssembly improvements (support for 32-bit browsers and theoretically UHD decoding). In addition, an example implementation of a web-based player was released to GitHub that enables the use of VVdeC for VVC playback in a web browser.

In an IEEE journal paper of October 2021, the Sharp decoder was reported to be capable of decoding 8K bitstreams at up to 60 Mbps at more than 110 fps.

In January 2022, a successful trial deployment of the Ali266 decoder was reported in Alibaba’s Youku video streaming service for mobile app usage. The trial used movies that were encoded by the Ali266 encoder (see item 18) at 1080p and 720p at bit rates 35% to 55% below that of HEVC encoding, reportedly with satisfactory decoding time and stability metrics.

Bytedance provided updated information about the BVC decoder in January 2022, reporting that the decoder had been optimized for heterogeneous CPU plus GPU systems, operating on the Apple AArch64 platform with GPU acceleration of the VVC decoding process. For decoding of 4K 8-bit CTC bitstreams on a Mac mini (with the Apple M1 chip), it was reported that the updated decoder achieved 39 fps on average with a single thread, which was 10x faster than the VTM-11.0 reference decoder, and that 72 fps could be achieved with two threads. For decoding of 4K 8-bit CTC bitstreams on an iPhone 13 (with the Apple A15 chip), it was reported that the BVC decoder achieved 35 fps on average with a single thread, which was 15x faster than the VTM-11.0 reference decoder.

AstroDesign announced a media player application called the SP-5020 VVC Viewer supporting VVC in November 2021. The AstroDesign player is based on the Sharp software decoder (discussed in item 5). This included support of the Main 10 profile and 4K capability at 60 fps on a recent-generation, 8-core CPU.

MediaTek announced the Pentatonic 2000 chipset for televisions in November 2021, featuring VVC support with 8K resolution and up to 120 Hz frame rate.

In December 2020, a demonstration of live 4K@60 fps transmission over a 4G mobile network was reported by KDDI research. In April 2021, additional capability was reported, including a real-time VVC encoder with 8K@30 fps capability along with a demonstration of 8K@30 fps real-time transmission over network lines.

Further updated information was provided in January 2022 by Alibaba. The Ali266 encoder was enhanced with four predefined quality/speed presets called “slow”, “medium”, “fast”, and “faster”, to meet different application needs. With the “faster” preset, 10-bit real time encoding for 1080p video-on-demand (VOD) content was reported and the coding efficiency of this preset had also been improved as well. Experiments reportedly demonstrated BD rates for {YUV-PSNR, YUV-SSIM} of {−42.01%, −42.90%} over the widely used open-source HEVC software encoder x265 at its “medium” preset, compared to the previously reported results of {−38.10%, −41.10%} in JVET-X0104 over the same anchor.

In Brazil, following its phase 2 testing conducted during July–December 2021, the SBTVD Forum announced in January 2022 the decision for VVC to be used for the main video stream for the upcoming next-generation TV 3.0 broadcasting standard in Brazil, both for over-the-air and over-the-Internet transmission.

## Text development and errata reporting (4)

Contributions in this area were discussed in session 5 at 1300–1400 UTC on Thursday 13 Jan. 2022, in session 9 at 1300–13450 UTC on Friday 14 Jan. 2022, and in session 14 at 1530–1555 UTC on Monday 17 Jan. 2022 (chaired by JRO).

[JVET-Y0049](https://jvet-experts.org/doc_end_user/current_document.php?id=11241) AHG2/AHG8: On the range extensions GCI flags [Y.-K. Wang (Bytedance)]

When gci\_present\_flag is equal to 0, no GCI syntax elements exist, including the new range extensions GCI flags. In that case, the general\_constraint\_info( ) syntax structure does not impose any constraint. However, when gci\_present\_flag is equal to 1, although all the GCI syntax elements specified in VVC edition 1 are surely present, the newly added range extensions GCI flags may or may not be present, depending on whether gci\_num\_additional\_bits (which was gci\_num\_reserved\_bits in VVC edition 1) is greater than 0.

An asserted problem arises when gci\_present\_flag is equal to 1 and gci\_num\_additional\_bits is equal to 0, in which case the new range extensions GCI flags are not present while their values are not inferred, and thus it is unclear whether the constraints specified by the flags equal to 1 apply or not.

To address the above asserted problem, this contribution proposes to specify that the value of each of the six newly added GCI flags (listed below) is inferred to be equal to 0 when gci\_present\_flag is equal to 1 and gci\_num\_additional\_bits is equal to 0:

1. gci\_all\_rap\_pictures\_constraint\_flag
2. gci\_no\_extended\_precision\_processing\_constraint\_flag
3. gci\_no\_ts\_residual\_coding\_rice\_constraint\_flag
4. gci\_no\_rrc\_rice\_extension\_constraint\_flag
5. gci\_no\_persistent\_rice\_adaptation\_constraint\_flag
6. gci\_no\_reverse\_last\_sig\_coeff\_constraint\_flag

Additionally, some asserted editorial improvement changes are proposed to the semantics of gci\_present\_flag, gci\_sixteen\_minus\_max\_bitdepth\_constraint\_idc, and gci\_three\_minus\_max\_chroma‌\_format\_constraint\_idc.

See further notes under JVET-Y0190

It is pointed out that another option would be to globally define that whenever a specific GCI flag is not present, its value would be inferred to be zero. This would avoid that the problem happens again in future extensions and would not be specific to range extensions.

[JVET-Y0050](https://jvet-experts.org/doc_end_user/current_document.php?id=11242) AHG2/AHG9: On the alpha channel information SEI message [Y.-K. Wang (Bytedance)]

In the semantics the alpha channel information SEI message, the alpha\_channel\_incr\_flag is inferred to be equal to 0 when not present. It is asserted that this inference seems unnecessary or even not correct. The flag is not present only when alpha\_channel\_cancel\_flag is equal to 1, in which case no information applies. It is further asserted that, if the absence of the SEI message is also considered, then the value of the flag should depend on the persistency scope of the previous SEI message of this type, and then this inference can be incorrect.

It is therefore proposed to remove the inference rule. Note that this asserted issue applies to both VSEI and HEVC.

Decision: Adopt JVET-Y0050, both for V2 of VSEI and HEVC JVET-Y1005/CDAM

[JVET-Y0190](https://jvet-experts.org/doc_end_user/current_document.php?id=11397) AHG2/AHG8: Suggestions for the operation range extensions GCI [J. Gan, Y. Yu, H. Yu (OPPO)] [late]

It is asserted that the current text for general constraints information (GCI) in VVC version 2 contains two bugs. For the first bug, it is asserted that incorrect semantics and handling of gci\_num\_additional\_bits lead to GCI decoding behaviour in VVC version 2 incompatible with VVC version 1 decoders. For the second bug, it is asserted that setting gci\_present\_flag to 1 and gci\_num\_additional\_bits to 0 leads to undefined constraint behaviour for the 6 coding tools which are governed by the 6 additional GCI flags introduced in VVC version 2.

It is additionally asserted that VTM 15.0 contains a bug related to decoding of GCI as reported by JVET-Y0005.

The same issue is raised as in JVET-Y0049 (aspect 2).

Beyond that, it is pointed out (aspect 1) that a problem might occur when the gci\_num\_additional\_bits has another value than 0 or 6 in version 2. The solution proposed appears appropriate.

Aspect 3 points out a software bug. Aspect 1 and 3 together are likely causing the problem with the conformance bitstreams.

New conformance streams need to be generated with all 6 GCI flags.

In the discussion, it is also mentioned that it might be desirable to rename the syntax element gci\_reserved\_zero\_bit, as it could in have values of 0 or 1. This may however not be necessary.

Side activity with relevant experts, developing an integrated text to fix the bugs.

Follow-up discussion in session 9 Fri 14 Jan. No consensus yet on the question if a decoder should expect that only values 0 or 6 would appear in a bitstream, i.e. if a decoder action for values 1…5 should be defined. If such action would not be defined, values 1…5 would need be reserved forever.

It is mentioned that the suggested solution “aspect 1a” implies values 1..5 to be ignored.

Both methods should not be difficult to describe.

One argument is to retain decoder conformance of main profiles consistent across versions. Minhua Zhou will provide a text that accommodates this and further discuss with other experts.

[JVET-Y0237](https://jvet-experts.org/doc_end_user/current_document.php?id=11446) Integrated specification text for JVET-Y0049 and JVET-Y0190 [J. Gan (OPPO), Y.-K. Wang (Bytedance)] [late]

Following presentation of input contributions JVET-Y0005 [1], JVET-Y0049 [2], and JVET-Y0190 [3], it was asserted that several bugs exist in the VVC version 2 specification relating to the general constraints information syntax, two of which were the cause of errors in conformance testing reported by JVET-Y0005. Relevant experts were asked to develop an integrated text to fix the bugs. Although consensus was reached on the majority of the text, differences of opinion still remain on one section of the text, namely the semantics for gci\_num\_additional\_bits. The differences have been captured in three options.

It is proposed to adopt the integrated text for one of options A, B or C. An associated software patch for the chosen option will be provided for VTM 16.0.

Option A would also allow values 1..5 in future, which is undesirable

Option B requires a decoder to ignore the values of GCI flags, in case that a value 1..5 would appear, and values >6.

Option C requires a a decoder to ignore the values of GCI flags, in case that a value >6 would appear (as values 1..5 could never happen). If values 1..5 would appear, this would be a non-conforming bitstream.

It is agreed that option C is the most logical solution.

Decision: Adopt JVET-0237 option C.

[JVET-Y0242](https://jvet-experts.org/doc_end_user/current_document.php?id=11451) AHG8: On SPS extension syntax [Y.-K. Wang (Bytedance)] [late]

JVET-Y0237 provides three text options on GCI extension for the JVET to choose, for fixing some asserted bugs reported in JVET-Y0049 and JVET-Y0190.

This contribution presents a conditional proposal of a change to the SPS extension syntax, depending on JVET's decision on JVET-Y0237 text options. If JVET chooses option A or B in JVET-Y0237, the SPS extension change is proposed, for an asserted alignment of the two extensions following the same design principle, in terms of handling different forward and backward compatibilities in similar manners. If JVET chooses option C in Y0237, the SPS extension change is not proposed, for the same purpose.

No need for consideration, as “option C” was selected for JVET-Y0237.

## Test conditions (6)

Contributions in this area were discussed in session 15 at 2100–2240 UTC on Monday 17 Jan. 2022 (chaired by JRO).

[JVET-Y0102](https://jvet-experts.org/doc_end_user/current_document.php?id=11296) On the balance of ECM coding gains between luma and chroma [F. Le Léannec, P. Andrivon, E. Thomas (Xiaomi)]

Thanks to the introduction of new coding tools, in particular cross-component tools, ECM coding gain over VVC is larger in chroma than in luma.

This contribution proposes to modify ECM encoder configurations in AI, RA and LB to balance the chroma and luma coding gains, through modified chroma QP mapping tables. Two encoder configurations are tested, leading to the following results (for Y/Cb/Cr) over VTM-11.0+V0056.

* ECM-3.1: AI: -6.8% / -12.3% / -13.2% ; RA: -14.8%/-18.6%/-19.4% ; LB : -12.3% / -21% / -21%
* config. 1: AI: -8.1% /-2.2% /-3.4% ; RA: -15.6% / -11.5% / -12.8% ; LB: -13% / -6.3% / -6.2%
* config. 2: AI: -7.3%/-7.9%/-8.7%; RA: -15.2% / -17.1% / -18.1% ; LB : -12.9% / -9.8% / -9.8%

It is proposed to adopt configuration 1 for AI and RA, and configuration 2 for LB.

Chroma QP offset tuned differently over various QP points

How consistent is the bit rate shift over the various QP/rate points? Was not investigated in detail.

Conf. 1 may be too aggressive.

[JVET-Y0178](https://jvet-experts.org/doc_end_user/current_document.php?id=11385) Crosscheck of JVET-Y0102 (On the balance of ECM coding gains between luma and chroma) [Y. Chen (InterDigital)] [late]

[JVET-Y0108](https://jvet-experts.org/doc_end_user/current_document.php?id=11302) AHG3/AHG12: Modification of JVET CTC for environmental considerations [F. Galpin, M. Radosavljević, E. François (InterDigital)]

In this contribution, new CTCs are proposed, aiming at reducing the environmental footprint of some of the tests conducted by the JVET contributors. As the overall complexity of the reference software steadily increases, required total computational power increases accordingly. This computational power is directly impacting the environmental footprint. In this contribution, a specific focus on the CO2e (Carbon Dioxide Equivalent) footprint is made, and an attempt to estimate this footprint for current reference software on the CTC is presented.

It is proposed to require simulations with only half length of sequences in the “pre experiment” stage. Rigid Testing in experiments not to be changed.

It is pointed out that companies might use the resources to increase the number of contributions instead.

We already have only 5 s sequence length in LB, 2.5s would be too short.

Different test conditions for EE and non-EE would be undesirable, as it becomes more difficult to compare.

No agreement that action should be taken, as the suggested approach might not help, Perhaps we should rather increase the bar for consideration of proposals with low efficiency.

[JVET-Y0112](https://jvet-experts.org/doc_end_user/current_document.php?id=11306) [AHG3] Merged VTM and HM CTC for SD 4:2:0 10-bit video [K. Sühring, F. Bossen, X. Li, V. Seregin, K. Sharman]

JVET-W0152 outlined possible merging of the various JVET CTC documents for HM and VTM reference software to simplify use and maintenance. AHG3 was mandated to propose merged documents. The attached CTC document is a merge of the CTC for SDR 4:2:0 10-bit video from JVET-T2010 (VTM) and JVET-U1100 (HM).

While editing the documents, the following issues were found:

1. In JVET-T2010 the MD5 hash of the CatRobot sequence belongs to a 600-frame version of the sequence which was referred to as “CatRobot1” in the JVET CfP. The hash of a 300-frame version of this sequence was used in JVET-U1100 (HM), referred to as “CatRobot” with “\_jvet\_” in the file name. A difference in coding results may be caused by the temporal filter, which can use additional frames beyond the coded frames, if available. The 300-frame version in located in the “jvet/ctc” folder of the test sequences FTP site, is referred to in the config files included in HM and VTM and was used in anchor generation. Thus, it was decided to refer to the 300-frame version of the sequence in the merged CTC document and use the hash of that version.
2. An informal section exists describing coding of the random-access conditions in segments. The number of additional frames in temporal filtering is described there, because these frames need to be available for each segment to match sequential coding. It is suggested to refer to a configuration parameter in the software that specifies the number of frames to be used to avoid changing the CTC, when this value changes. A software update is pending to implement these parameters.
3. It was noted that some of the test sequences contain an additional frame with copyright information. That frame should not be used for temporal filtering. The reference software should be updated with parameters that specify the first or last frame that may be used for filtering beyond the coded frames.
4. VTM CTC do not contain test cases for 8-bit coding. The 8-bit test cases from HM were not added to the merged CTC.

It is pointed out that the item 1. also affects ECM CTC.

It is discussed whether completely giving up 8 bit would be desirable. In particular, it might also be useful for VVC to ensure that other bit depths than native 10 bits are supported. This might e.g. be done for certain classes only, such as C and D. This is left for further study, should be based on results.

HM CTC also includes range extensions. This should be retained as separate output document.

Decision(CTC): Adopt the merged CTC from JVET-Y0112.

Items 1 and 3 also have impact on ECM CTC

[JVET-Y0113](https://jvet-experts.org/doc_end_user/current_document.php?id=11307) Adjusting luma/chroma BD-rate balance in ECM [Y. Chen, E. François, P. Nikitin (InterDigital)]

No need to present, see notes under JVET-Y0223.

[JVET-Y0117](https://jvet-experts.org/doc_end_user/current_document.php?id=11311) AHG12: ECM coding performance for HDR/WCG content and suggested common test conditions [T. Lu, F. Pu, P. Yin, S. McCarthy, W. Husak, T. Chen (Dolby), N. Hu, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution presents HDR/WCG coding results for ECM3.1 with two encoder settings. The first setting follows the VTM HDR/WCG common test conditions. The second setting adds encoder configuration optimizations recently added for SDR coding with ECM and the CTU size of HLG sequences is 256. The results of ECM3.1 HDR coding compared with VTM11.0 with MCTF are as follows:

First setting (VTM HDR/WCG CTC):

For PQ content, the BD rate savings for wPSNRY/U/V/DE100/L100 are:

AI { -6.19% -18.43% -22.07% -12.31% -6.65% }

RA { -11.19% -25.44% -28.46% -19.21% -11.73% }

For HLG content, the BD rate savings for PSNRY/U/V are:

AI { -5.88% -12.24% -16.13% }

RA {-9.84% -18.70% -21.28% }

Second setting (VTM HDR/WCG CTC + ECM encoder optimization + CTU 256 for HLG):

For PQ content, the BD rate savings for wPSNRY/U/V/DE100/L100 are:

AI { -6.19% -18.43% -22.07% -12.31% -6.65% }

RA { -11.59% -27.82% -31.98% -22.98% -12.41% }

For HLG content, the BD rate savings for PSNRY/U/V are:

AI { -6.02% -11.86% -16.05% }

RA { -10.79% -20.84% -24.42% }

It is recommended to establish ECM HDR/WCG common test conditions corresponding to the second setting.

Decision(CTC): Include in CTC document for ECM/EE2, HDR as optional class.

Als change VTM CTC for HDR (encoder optimization aligned with the ECM approach).

It is further suggested to seek for more HDR test material, in particular 4K for PQ would be desirable (e.g. content from VVC verification test).

[JVET-Y0167](https://jvet-experts.org/doc_end_user/current_document.php?id=11374) Cross-check report of JVET-Y0117 on suggested ECM common test conditions for HDR/WCG content [E. François (InterDigital)] [late]

[JVET-Y0223](https://jvet-experts.org/doc_end_user/current_document.php?id=11432) On luma/chroma BD-rate balance in ECM: tests combining JVET-Y0102 and JVET-Y0113 [Y. Chen, E. François, P. Nikitin (InterDigital), F. Le Léannec, P. Andrivon, E. Thomas (Xiaomi)] [late]

Thanks to the introduction of new coding tools, in particular cross-component tools, ECM BD-rate coding gain over VVC is larger in chroma than in luma using current CTCs. JVET-Y0102 and JVET-Y0113 explore alternative encoder settings based on ECM-3.1 to obtain alternate repartitions of BD-rate variations between luma and chroma. This contribution reports results from JVET-Y0102 and JVET-Y0113 plus additional experiments combining tuning from JVET-Y0102 and JVET-Y0113.

Beyond chroma QP offset, LMCS offset used as an additional means to control the chroma quality.

Both SDR and HDR included

Behaviour over the different QP points not checked in detail.

Cross-checking party expresses concerns if the bit rate shift might be too aggressive, and suggests that it should be checked if artifacts occur.

No need to hurry – subjective quality should be checked. Primarily high QP values should be checked. Potentially also look at additional “colorful” content.

Further study highly recommended.

[JVET-Y0227](https://jvet-experts.org/doc_end_user/current_document.php?id=11436) Crosscheck of JVET-Y0223 (On luma/chroma BD-rate balance in ECM: tests combining JVET-Y0102 and JVET-Y0113) [T. Lu (Dolby)] [late]

## Verification testing (1)

Contributions in this area were discussed in session 13 at 1400–1445 UTC on Monday 17 Jan. 2022 (chaired by JRO).

[JVET-Y0047](https://jvet-experts.org/doc_end_user/current_document.php?id=11239) AHG4: First report of spatial scalability verification tests [P. de Lagrange, F. Urban, E. François (InterDigital), W. Hamidouche (INSA)]

This contribution reports the results a formal visual quality assessment of VVC in spatial scalability mode, following preparation work described in JVET-X0202. With equal bitrate for base HD and enhancement UHD layer, PSNR-Y based performance is ‑24% vs simulcast, while MOS-based performance is around ‑15%; other layer bitrate allocations are tested in JVET-Y0048. Part of those results were used as a subjective evaluation report of VVC scalable profile in response to the TV 3.0 project Call for Proposals from the Brazilian Digital Terrestrial TV Forum (see JVET-U0128, JVET-V0167 and JVET-X0202). Beyond that, these tests can be used as a basis for JVET verification tests of VVC in scalable mode.

2x spatial scalability is investigated.

Performance comparison vs. simulcast and single-layer

Unlike SHVC verification test, which used 0.5 s IRAP period, 1 s was used

MCTF preprocessing was not found effective in scalable coding

Base layer and low res in simulcast almost identical

Quality/rate of base vs. enhancement layer was equivalent to SHVC verification test.

From sequences that were previously used in SHVC verification tests, it is concluded that gain against simulcast are similar (with PSNR based criteria).

Base layer quality might not be chosen in an optimum way over the different rate points. needs to be improved. It is also suggested that one possibility would be including an upsampled base layer to determine appropriate settings. Cases should be investigated where scalability has demonstratable benefit. Also, it should be avoided that quality saturates at the high rate end.

Base layer should be viewable.

At highest rate points, saturation occurs.

Should a verification test include SHVC, or would it be sufficient to demonstrate the benefit vs. simulcast?

It is suggested to further investigate the definition of a scalable VVC verification test, in particular with regard to the optimum rate/quality settings for base and enhancement layers. Put this as mandate of AHG4.

JVET-Y0048 is also related to this.

## Test material (3)

Contributions in this area were discussed in session 13 at 1445–1520 UTC on XXday X Jan. 2022 (chaired by XXX).

JVET-Y0042 (section 4.10) is also related.

[JVET-Y0041](https://jvet-experts.org/doc_end_user/current_document.php?id=11233) AHG7: Proposed new class of gaming sequences with depth and optical flow information [G. Martin-Cocher, M. Badawi, T. Poirier, S. Puri, K. Naser (InterDigital)]

This contribution proposes to create a new class of content (DMV – Depth & Motion Vector) corresponding to the cloud gaming and game casting scenario. The proposed sequences include depth and optical flow information.

10 sequences were created using the Unity game engine to load and play several games. We captured the frame buffers using Unity Recorder and custom post-processing shaders in both the built-in render pipeline and the High-Definition Render Pipeline. We captured the color buffer into an RGBA16 Signed Float texture as well as the depth and optical flow in a second similar texture (with the linear depth info encoded into the R channel, and the [X,Y] coordinates of the motion vectors into the [G,B] channels). The textures were recorded at 60FPS in 1080p and 4K resolutions and written into uncompressed EXR 16 bit images. Two YUVs sequences were then generated. The depth is stored in the Y channel and converted to 14 bits to avoid high bit-depth coding, vertical and horizontal motion vectors are stored at quarter pixel precision in U/V channels for the corresponding texture frame.

Additional gaming content is more than welcome.

Licensing OK, could be put on ftp in candidate folder.

Does the rendering simulate motion blur? Yes, in some of the sequences.

Was the suitability for viewing tests considered?

Expert viewing (including coded material) as an AHG4 activity until next meeting.

[JVET-Y0071](https://jvet-experts.org/doc_end_user/current_document.php?id=11265) New Test Content for Video Conferencing Applications [Z. Sinno, G. Desgouttes, A. M. Tourapis, D. Singer (Apple)] [late]

This document presents 4 new video conferencing sequences that could be used for the development of new Video Coding standards by MPEG and the JVET. The sequences included were captured on a mobile device at 30fps and at a 1080x1920 resolution. They are provided in both an HDR (Non Constant Luminance YCbCr 420 BT.2100 HLG) and an SDR HDR (Non Constant Luminance YCbCr 420 BT.709) representation.

4 sequences in total, up to 25 s.

Creative Commons license BY-SA 4.0 (which should include the consent of the person for public availability)

Some camera artifacts are also present (exposure/color changes), but no compression artifacts.

No coding investigations so far, this would be necessary for further judgement.

Sequences will be made available soon.

[JVET-Y0123](https://jvet-experts.org/doc_end_user/current_document.php?id=11317) On Test Sequences [J. Xu, L. Zhang (ByteDance), M. Martin-Cocher (InterDigital)]

Was presented 0010 UTC on Jan 21 in session 24

This contribution describes several video industry related trends related and propose to have new test sequences to reflect those trends.

Vertical video, gaming, and screen sharing content are mentioned.

It is proposed to 1) call for new test sequences to reflect industry trends; and 2) establish an AhG on collecting recent video industry information and suggest typical test sequences based on the collected information.

It is not fully clear what is requested.

Following review of JVET-Y0041 and JVET-Y0071, an effort of investigating new test material is already planned in AHG4. A call for new test material could be issued as a follow-up. For that, a more precise specification about the characteristics of such test material would be needed. It is suggested to contribute to AHG4 in this direction.

## Quality assessment (3)

Contributions in this area were discussed in session 9 at 1420–1510 UTC on Friday 14 Jan. 2022 (chaired by JRO).

[JVET-Y0045](https://jvet-experts.org/doc_end_user/current_document.php?id=11237) AHG11/EE1 viewing preparation report [[E. Alshina](mailto:elena.alshina@huawei.com), M. Wien, A. Segall]

No need for presentation – was included in AHG activity, and result reported in JVET-Y0212.

[JVET-Y0083](https://jvet-experts.org/doc_end_user/current_document.php?id=11277) AHG4/AHG10: Report of Teleconference on Viewing Session Preparation for Deblocking [M. Wien, H. Zhang, X. Li]

No need for presentation – was included in AHG activity, and results reported in JVET-Y0212.

[JVET-Y0212](https://jvet-experts.org/doc_end_user/current_document.php?id=11420) AHG4: REV Result for AHG11/EE1 and AHG10/Deblocking [M. Wien (RWTH)]

Remote expert viewing tests were conducted during the 25th JVET meeting for the exploration activity on EE1/DNN-based coding tools and deblocking filter tuning. The tests were performed using mp4 files provided by the proponents with a total of 60 comparisons for EE1 and 36 comparisons in the deblocking category. Sequences of 720p and 1080p contents were evaluated for both EE1 and deblocking, and an additional set of UHD sequences was evaluated for EE1 only. Calls for participation in the subjective viewing were issued on the JVET reflector. Overall, 25 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.

In the deblocking filter experiment the figures indicate that the visual performance impact of the proposal is negligible for both, the VTM and the ECM test cases. In the ECM case, two visually significant differences are reported on the ArenaOfValor sequence, with the indication of a significant benefit on the one QP point and the indication of the significant disadvantage on the other QP point. It is noted that JVET-Y0085 reports REV results performed by the proponents themselves. It is found that there is a significant congruency between the two experiments.

For the EE1-1.1 experiment, a significant visual benefit is reported for 7 out of 18 test cases. In one additional case (KristenAndSara QP34), the confidence interval is only very slightly touching the zero line. In all other cases, the zero line is included in the CI and therefore, comparable visual performance is suggested.

In the EE1-1.2 experiment, significant visual benefit is reported for 10 out of 26 test cases. The benefits seem to be more prominent on the UHD sequences. In two cases (Kimono1 QP37 and FourPeople QP42), a significant disadvantage is observed.

In the super-resolution category (EE1-2.x), the largest visual benefits are observed compared to the base line VTM. The gains appear to be higher by tendency for the NN-based method.

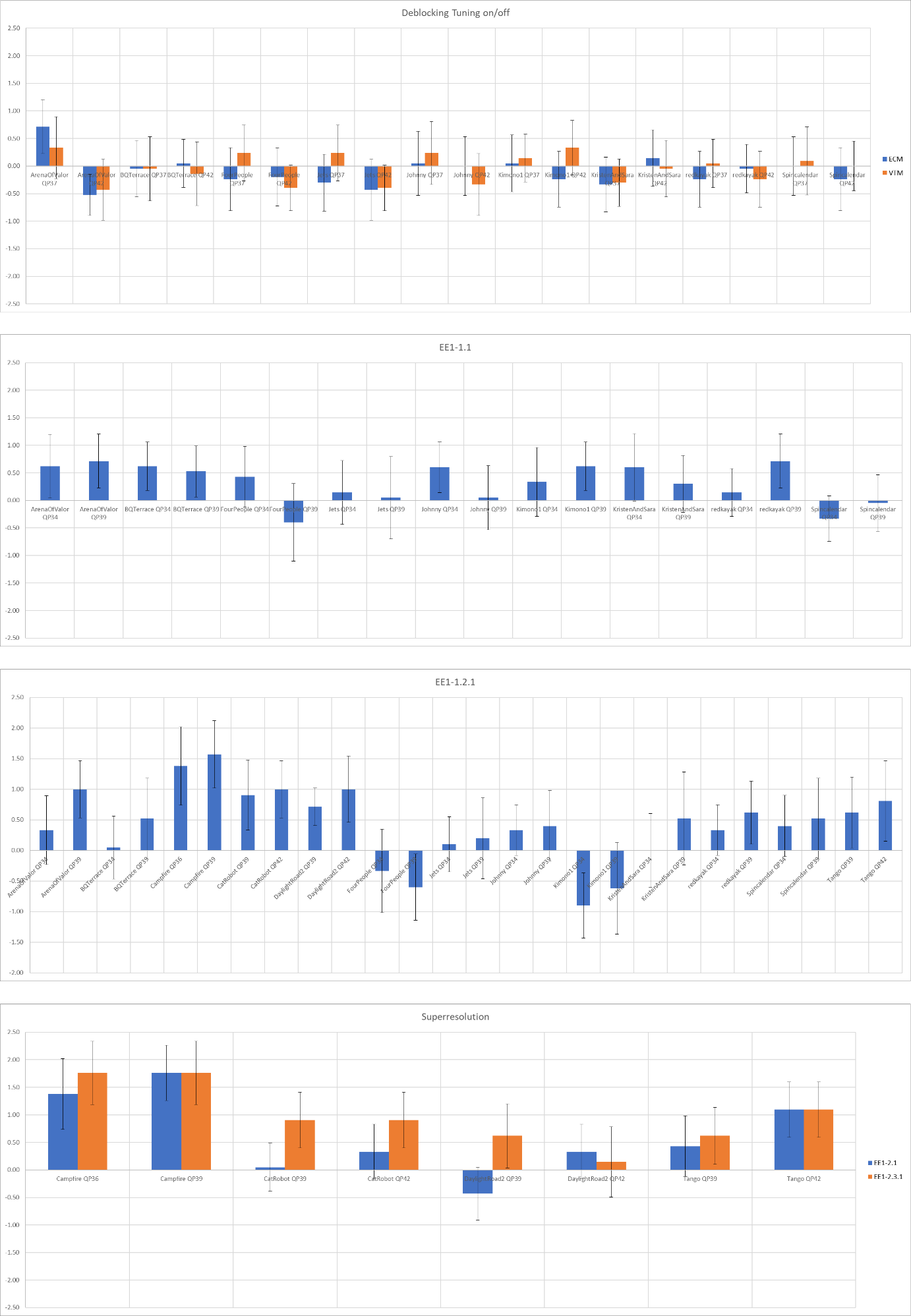


Figure : MOS plots

Based on the reported MOS values, the following conclusions are drawn:

In the deblocking filter experiment the figures indicate that the visual performance impact of the proposal is negligible for both, the VTM and the ECM test cases. In the ECM case, two visually significant differences are reported on the ArenaOfValour sequence, with the indication of a significant benefit on the one QP point and the indication of the significant disadvantage on the other QP point.

For the EE1-1.1 experiment, a significant visual benefit is reported for 7 out of 18 test cases. In one additional case (KristenAndSara QP34), the confidence interval is only very slightly touching the zero line. In all other cases, the zero line is included in the CI and therefore, comparable visual performance is suggested.

In the EE1-1.2 experiment, significant visual benefit is reported for 10 out of 26 test cases. The benefits seem to be more prominent on the UHD sequences. In two cases (Kimono1 QP37 and FourPeople QP42), a significant disadvantage is observed.

In the super-resolution category (EE1-2.x), the largest visual benefits are observed compared to the base line VTM. The gains appear to be higher by tendency for the NN-based method.

No evidence of subjective difference for the deblocking filter experiment

For EE1,

* Much larger diversity of sequences than ever before
* Very clear indication about the visual benefit of NN based loop filters in many cases, but the few cases where NN performs worse should be more closely investigated
* For superresolution, tendency for some cases that NN performs better than RPR, would be good to test with more sequences

It was further commented that it might be good to use a larger number of scales (e.g. 7). Not allowing “0” probably increases the standard deviation when there were many cases with almost equal quality

## Conformance test development (1)

Contributions in this area were discussed in session 20 at 2240–2250 UTC on Wednesday 19 Jan. 2022 (chaired by JRO).

[JVET-Y0127](https://jvet-experts.org/doc_end_user/current_document.php?id=11321) AHG5: Editors update on conformance testing for VVC operation range extensions [D. Rusanovskyy, T. Hashimoto, H.-J. Jhu, I. Moccagatta, M. G. Sarwer, Y. Yu]

This document is draft 2 of the conformance testing specification for VVC operation range extensions, for testing the new coding tool features being added in the second edition of VVC (Rec. ITU-T H.266 | ISO/IEC 23090-3).

Changes are made to resolve the issues raised in the CDAM ballot comments.

DoCR to be reviewed in session 22 on Thursday

## Software development (2)

Contributions in this area were discussed in session 20 at 2250–2315 UTC on Wednesday 19 Jan. 2022 (chaired by JRO).

[JVET-Y0110](https://jvet-experts.org/doc_end_user/current_document.php?id=11304) AHG11: Small Ad-hoc Deep-Learning Library (SADL) update [F. Galpin, F. Mom, T. Dumas, P. Bordes, P. Nikitin, E. François (InterDigital)]

This contribution presents an update on the SADL library already described in JVET-W0181.

- Change the input format of the converter to ONNX for more flexible and universal support

- Added counters to assess accurately MAC/pix for a given model,

- Minor fixes

Conversion to ONNX can be done from various frameworks, including tensorflow and pytorch.

It is reported that a conversion of one of the loop filters from EE1 (test 1.2) is currently under investigation.

Not necessarily optimized for speed, but deemed useful for understanding the behaviour of networks. Also very useful for straightforward integerization.

Include in EE, in particular extend its usage to technology other than intra prediction, investigate for cross-environment cross-checking, investigate and quantify mismatches, and impact of encoder/decoder mismatches/drift (when using different platforms). Note that at this moment it is not mandatory to use the SADL package, but this exercise shall be applied to representative proposals from different domains of EE exploration. Integerization should also be a target.

[JVET-Y0248](https://jvet-experts.org/doc_end_user/current_document.php?id=11457) AHG3: per-picture configuration for VTM [P. de Lagrange (InterDigital)] [late]

Was presented in session 24 0030 UTC

This contribution proposes to add a per-picture configuration capability to VTM. An implementation with update support of a few parameters is available in merge request !2032. Per-picture configuration takes the form of an update file where each line contains a POC and a command line.

The configuration update file consists in a sequence of <poc> : <parameters> lines. Parameters are limited to those supported for per-picture update (currently a few PH/SH related parameters.

Changing parameters in SPS and PPS is not currently supported because of the current state of the VTM, which allocates fixed SPS/PPS. Supporting SPS/PPS change would require more work.

It is asked what would be useful cases to be tested, and evidence about advantages.

One example given is a change of QP at a certain frame.

To be further studied.

## Implementation studies (4)

Contributions in this area were discussed in session 22 at 1520–1720 UTC on Thursday 20 Jan. 2022 (chaired by Frank Bossen).

[JVET-Y0054](https://jvet-experts.org/doc_end_user/current_document.php?id=11248) Update on a VVC software decoder, BVC, for heterogeneous CPU plus GPU systems [L. Li, H. Yin, L. Zhang, Y. Zhang (Bytedance)]

In JVET-V0128, a VVC software decoder, the BVC decoder, fully developed by Bytedance was introduced. To fully utilize hardware accelerators on mobile devices, the BVC decoder has been continuously optimized for heterogeneous CPU plus GPU systems. Currently, the BVC decoder supports the Apple AArch64 platform and utilizes GPU to accelerate the VVC decoding process. For decoding of 4K 8-bit CTC bitstreams on Mac mini (with the Apple M1 chip), it is reported that the decoder achieves 39 fps on average with single thread, which is 10x faster than the VTM-11.0 reference decoder, and 72 fps can be achieved with two threads. For decoding of 4K 8-bit CTC bitstreams on iPhone 13 (with the Apple A15 chip), it is reported that the BVC decoder achieves 35 fps on average with single thread, which is 15x faster than the VTM-11.0 reference decoder.

Q: is there a plan to support 10-bit decoding? Currently not planning to but would not expect major difference in decoding performance.

It was mentioned that a lot of content is encoded with the bit depth set to 8.

Q: how long would a battery last? Hasn't been tested.

Q: what parts of VVC are GPU friendly? ALF is the most friendly.

[JVET-Y0066](https://jvet-experts.org/doc_end_user/current_document.php?id=11260) Update on Ali266, the optimized VVC encoder implementation by Alibaba [X. Dong, S. Fang, Z. Huang, J. Liu, S. Xu, R. Yang, L. Yu, J. Chen, R.-L. Liao, Y. Ye (Alibaba)]

This contribution reports updated information on runtime and coding efficiency of Ali266, an optimized software VVC encoder developed by Alibaba. Ali266 was first presented in JVET-W0127, it supported two presets, with its fast preset aimed at real-time encoding applications. In JVET-W0127, Ali266 was able to achieve encoding speed of 37.7 fps for 720p sequences.

In JVET-X0104, we reported further improvement in Ali266’s encoding speed at its fast preset. Under the configuration of fast preset with 8-bit internal bit depth, we reported that a 1.9x encoding speedup had been achieved compared with that in JVET-W0127 for video sequences at 1920x1080 or 1080x1920 resolutions. The average encoding speeds achieved 36.3 fps for classes 1920x1080 / 1080x1920 and 63.4 fps for classes 1280x720 / 720x1080. Based on such data, we concluded that Ali266 was a VVC encoder capable of real-time encoding for 1080p sequences.

Further updates to Ali266 are reported. First, Ali266 was enhanced with 4 predefined quality/speed presets, slow, medium, fast, and faster, to meet different application needs. Furthermore, we have not only achieved 10-bit real time encoding for 1080p video-on-demand (VOD) content with the “faster” preset, but also improved the coding efficiency. Our experiments demonstrate that Ali266 achieves BD rates for {YUV-PSNR, YUV-SSIM} of {-42.01%, -42.90%} over the widely used open-source HEVC software encoder x265 at the medium preset, compared to the previously reported results of {-38.10%, -41.10%} in JVET-X0104 over the same anchor.

[JVET-Y0122](https://jvet-experts.org/doc_end_user/current_document.php?id=11316) Ali266 @ Youku: Trial deployment of VVC for video streaming [Y. Jia, Y. Zhang, F. Hu, M. Li, W. Jiang (Youku), Z. Huang, J. Liu, J. Chen, Y. Ye (Alibaba)]

Youku is a premium online video streaming service wholly owned by Alibaba, serving hundreds of millions of consumers through its namesake app. Ali266 is a software-based VVC encoder and decoder implementation developed by Alibaba Cloud Intelligence. Recently, Youku teamed up with Alibaba Cloud Intelligence to deploy Ali266 on the Youku app. It is claimed to be the first such effort that aims at bringing the latest VVC standard to streaming customers at scale. In this contribution we report preliminary deployment status of Ali266 on Youku.

Trial with 720p and 1080p content. Reported bit rates 35% to 55% lower than with HEVC. Playback on mobile devices. Further study of battery life is planned.

Will bring demos to next F2F meeting.

[JVET-Y0136](https://jvet-experts.org/doc_end_user/current_document.php?id=11330) Update on open, optimized VVC implementations VVenC and VVdeC [A. Wieckowski, J. Brandenburg, C. Bartnik, V. George, J. Güther, G. Hege, C. Helmrich, A. Henkel, T. Hinz, C. Lehmann, C. Stoffers, I. Zupancic, B. Bross, H. Schwarz, D. Marpe, T. Schierl (HHI)] [late]

This document provides updated information on features, coding efficiency and runtime for version 1.3.1 of the open VVC software encoder VVenC released in December 2021 and version 1.3.0 of the open VVC software decoder VVdeC released in December 2021. In addition, an example implementation of a web-based player that enables to use VVdeC for VVC playback in a web browser has been made available on GitHub.

Main changes for VVenC v1.3.1 since version 1.1.0 include:

* Improved presets:
  + Major speedup for slower (10%), minor speedups for others
  + Major efficiency improvement for faster (−5.2% for HD, −3.1% for UHD PSNRYUV BD-rate)
* Improved rate-control:
  + Separate encoding of first and second pass in 2-pass rate control (using a statistics data file)
  + Improvements to single picture and low-rate rate control
  + Added 1-pass rate control with pre-analysis (experimental)
* Improved QP adaptation for subjective quality at low rates
* Added features:
  + Support for packed 10 bit YUVs
  + Frame rate parametrization and signaling using nominator and denominator
  + Tile parallelism
* Ported MCTF improvements from VTM
* Various bugfixes, improvements and memory reductions

Without QP adaptation for subjective optimization and 8 threads the following PSNR-based YUV BD-rates compared to HM-16.24 (GOP16+MCTF) as well as speedup factors compared to HM-16.24 and VTM-14.2 (GOP32+MCTF) are reported for different presets and JVET class B (HD), class A (UHD) as well as both (HD4K) test sequences:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Preset | HD |  |  | UHD |  |  | HD4K |  |  |
| PSNRYUV BD-rate vs. HM‑16.24 | Speedup vs. HM‑16.24 | Speedup vs. VTM-14.2 | PSNRYUV BD-rate vs. HM‑16.24 | Speedup vs. HM‑16.24 | Speedup vs. VTM-14.2 | PSNRYUV BD-rate vs. HM‑16.24 | Speedup vs. HM‑16.24 | Speedup vs. VTM-14.2 |
| FASTER | −7.7% | 170x | 1200x | −12.4% | 200x | 1400x | −10.2% | 180x | 1300x |
| FAST | −21.2% | 94x | 700x | −23.9% | 110x | 780x | −22.7% | 100x | 740x |
| MEDIUM | −31.5% | 25x | 180x | −34.5% | 33x | 240x | −33.2% | 29x | 210x |
| SLOW | −35.1% | 8.7x | 65x | −37.9% | 12x | 88x | −36.6% | 10x | 76x |
| SLOWER | −38.6% | 1.9x | 14x | −41.4% | 2.6x | 19x | −40.2% | 2.3x | 17x |

With QP adaptation for subjective optimization and 8 threads, the following MS-SSIM-based YUV BD-rates compared to HM-16.24 (GOP16+MCTF) as well as speedup factors compared to HM-16.24 and VTM-14.2 (GOP32+MCTF) are reported for different presets:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Preset | HD |  |  | UHD |  |  | HD4K |  |  |
| MS-SSIMYUV BD-rate vs. HM-16.24 | Speedup vs. HM-16.24 | Speedup vs. VTM-14.2 | MS-SSIMYUV BD-rate vs. HM | Speedup vs. HM-16.24 | Speedup vs. VTM-14.2 | MS-SSIMYUV BD-rate vs. HM-16.24 | Speedup vs. HM-16.24 | Speedup vs. VTM-14.2 |
| FASTER | −16.1% | 150x | 1200x | −14.8% | 190x | 1500x | −15.4% | 170x | 1300x |
| FAST | −28.3% | 90x | 700x | −27.2% | 110x | 840x | −27.7% | 98x | 770x |
| MEDIUM | −36.6% | 24x | 190x | −37.6% | 32x | 250x | −37.1% | 28x | 220x |
| SLOW | −40.0% | 8.7x | 67x | −40.9% | 12x | 95x | −40.5% | 10x | 81x |
| SLOWER | −42.6% | 1.9x | 15x | −44.2% | 2.6x | 20x | −43.4% | 2.2x | 18x |

Main changes for VVdeC v1.3.0 since version 1.2.0 include various fixes and improvements:

* Improved tune-in behavior
* 3x memory reduction
* MinGW support
* Cleanups
* Minor SIMD improvements
* WebAssembly improvements (support for 32-bit browsers and theoretically UHD decoding)

Q about memory. Memory use increases with each thread. Memory reduction was needed to fit within 2GB limit.

## AHG7: Low latency and constrained complexity (7)

Contributions in this area were discussed in session 15 at 2240–2305 UTC and in session 16 at 2325–0150+1 UTC on Monday 17 Jan. 2022 (chaired by JRO).

[JVET-Y0041](https://jvet-experts.org/doc_end_user/current_document.php?id=11233) AHG7: Proposed new class of gaming sequences with depth and optical flow information [G. Martin-Cocher, M. Badawi, T. Poirier, S. Puri, K. Naser (InterDigital)]

See section 4.5

[JVET-Y0042](https://jvet-experts.org/doc_end_user/current_document.php?id=11234) AHG7: Modification of and new classes of sequences [G. Martin-Cocher (InterDigital)]

As different resolutions (up to 8K), different average bitrates (from few kbps up to 25 Mbps) and applications will make use of low delay/ lower complexity operating points, this contribution proposes to redefine classes of sequences based on the identification of key scenarios and their characteristics as described in [JVET-Y0043](https://jvet-experts.org/doc_end_user/current_document.php?id=11235). These new or redefined classes of sequences should be used in the LLCC/LD configurations.

Proposal to create a new class S (screen), replacing F and TGM. This would be highly welcome, a good replacement has been sought for a long time.

Proposal to modify class E (new sequences, 1080p, vertical content, remove QP22); some concern is expressed about removing QP22; it is suggested to perform visual testing if quality of QP27 is sufficient for this kind of applications. Sequences from JVET-Y0071 could be considered, but more would be desirable.

Other classes: Gaming, Surveillance, Remote control. Some sequences in these categories exit in our data bases.

Tests need to be conducted to investigate the RD behaviour of the sequences proposed. This should also include visual inspection. Coordinated effort of AHG4 and 7.

[JVET-Y0043](https://jvet-experts.org/doc_end_user/current_document.php?id=11235) AHG7: LLCC Scenarios and baseline configurations [G. Martin-Cocher, S. Puri, T. Poirier, K. Naser (InterDigital), J. Xu (Bytedance), D. Nicholson (Ektacom), M. Sychev (Huawei), L. Wang (Nokia), S. Liu, W. Yang (Tencent), J. M. Tiesse (VITEC), M. Karczewicz (Qualcomm)]

This contribution identifies key Low Latency Controlled Complexity (LLCC) scenarios, namely cloud gaming, game casting, video surveillance, remote control of systems, video conferencing and summarize their characteristics in terms of video formats, requirements, constraints on coding tools. This contribution proposes baseline LLCC CTCs as follow:

The VTM LLCC-1 configuration is defined as a modification of LD-P, with flat QP, GDR enabled with the GdrInterval set to 1second and a GdrPeriod = GdrInterval+ 4P frames. LMCS chroma scaling is turned OFF. QP42 is added and BD-rate calculations are proposed to be performed on QP22-37 and on QP 27-42.

For ECM, the LLCC-1 configuration is similar to the VTM LLCC-1 configuration, but instead of GDR, an IDR is set every second, until GDR will be implemented in ECM.

The VTM and ECM LLCC-2 configuration is defined as a modification of LD-B, with a maximum of 2 reference frames. QP42 is added and BD-rate calculations are proposed to be performed on QP22-37 and on QP 27-42. RPR is proposed to be enabled once per clip for class E.

The contribution further proposes baseline principals to refine these CTCs.

A reference software encoding run time to real time ratio (ET/RT ratio) principle for a hypothetical single threaded non optimised reference software is proposed as a reference point to refine the baseline LLCC CTCs. It’s initial value could be set at 1000/1 for 1080p resolution. This ratio is not proposed as an evaluation criterion for individual tools.

LLCC-1 for cloud gaming, remote control, video surveillance (with some similarity to LP)

LLCC-2 for game casting and video conferencing (with some similarity to LB)

This allocation of applications might be not fully justified

Why frequent GDR (proposed is a period of 1s)?

Is RPR a low latency tool? It is rather used for rate control, whereas otherwise the suggested conditions are constant QP

Are GDR and RPR relevant in the context of evaluating tools?

Why still use an I frame in the beginning – better start with GDR?

Why IPPP in LLCC-1? LB is low delay as well, and not necessarily high complex. There are also opinions that many realtime encoders nowadays only use only P frames

ET/RT ratio for single-threaded encoder is proposed as an evaluation criterion of encoder complexity. Would complexity be reasonably judged via run time in a non-optimized codec such as VTM/ECM? Then better use an optimized encoder such as VVEnc? Or have other criteria such as memory usage, etc.?

QP ranges should be defined according to relevant rates and quality, might even be sequence dependent.

Are the configurations suitable to identify encoder-friendly tools? Are certain tools planned to be disabled?

RPR is frequently used in videoconferencing nowadays

To be further refined and discussed according to the aspects above. It is suggested to start from a baseline without GDR and RPR, LB with 2 reference frames, and investigate a set of enabled tools that provide a reasonable tradeoff regarding encoder complexity vs. compression.

[JVET-Y0060](https://jvet-experts.org/doc_end_user/current_document.php?id=11254) AHG7: refined LLCC configurations [G. Martin-Cocher, K. Nasser, T. Poirier, S. Puri (InterDigital)]

This contribution complements the Low Latency Controlled Complexity (LLCC) baseline configurations defined in [JVET-Y0043](https://jvet-experts.org/doc_end_user/current_document.php?id=11235) by enabling and disabling further tools according to their latency and complexity impacts.

The following additional constraints are proposed for both LLCC-1 (LD-P) and LLCC-2 (LD-B): Tools off: ALF, DQ, RDOQ and MTS; Tools on: Sign-hiding, IBC, MTSImplicit.

For LLCC-1 the following additional constraints are proposed: MTT=0, min QT size = 4. For Class B (1080p) the Encoding Time to Real Time (ET/RT) ratio is reduced from 8538/1 to 1524/1 for VTM and from 22392/1 to 10248/1 for ECM. The encoding runtime is reduced to 11% for VTM and to 44% for ECM.

For LLCC-2 the following additional constraints are proposed: MTT=1, min QT size = 4. For Class B (1080p) the Encoding Time to Real Time (ET/RT) ratio is reduced from 12090/1 to 2850/1 for VTM and from 36108/1 to 14088/1 for ECM. The encoding runtime is reduced to 27% for VTM and to 38% for ECM.

A software patch for ECM AffineAMVP is available, enabling to test it independently from Affine. It is further proposed to turn OFF AffineAMVP in the LLCC configurations to further reduce the encoding runtime.

* ECM LD-P (class E,F, DMV) Affine Off: 9,53% Y, 8,33% U, 8,19%V, 87%EncT
* ECM LD-P (class E,F, DMV) AffineAMVP off: 2,88% Y, 2,46% U, 2,70%V, 88%EncT
* ECM LD-B (class E,F, DMV) Affine Off: 9,59% Y, 8,35% U, 8,13%V, 88%EncT
* ECM LD-B (class E,F, DMV) AffineAMVP off: 3,08% Y, 2,77% U, 2,65%V, 89%EncT

VTM and ECM are considered more or less separately here. Would it be possible to find an ECM configuration that disables some tools (regardless if they are VVC or ECM tools) and achieves a better tradeoff complexity vs. compression than VTM?

Continue this investigation (similar as tool-on/tool-off) in AHG. It may be necessary (like in JVET-Y0060) to develop specific software patches allowing disabling of certain tools or parts of tools (in cases where no HLS flag exists for disabling).

Decision(SW): The software patch from JVET-Y0060 to be reviewed by software coordinators for merging (both for VTM and ECM, pure encoder configuration change).

Further refinement of investigating encoder settings for a future LLCC CTC and investigating suitable test sets for this kind of applications should be conducted in parallel.

[JVET-Y0101](https://jvet-experts.org/doc_end_user/current_document.php?id=11295) AHG7/AHG10: Depth motion based fast Multi-Type Tree Splitting [S. Puri, K. Naser, T. Poirier, G. Martin-Cocher (InterDigital)]

A new class of gaming content (DMV) is defined in [JVET-Y0041](https://jvet-experts.org/doc_end_user/current_document.php?id=11233) that includes depth and motion information (motion field or optical flow) retrieved from the frame buffers inside a game engine.

This contribution presents a Depth Motion based Fast Multi-Type Tree (DM\_FastMTT) method that uses depth and motion information (e.g. from a game engine) to determine the Multi-Type Tree (MTT) split on the encoder side without expensive RDO. The proposed non-normative method is applied under the low delay/Low Latency Controlled Complexity (LLCC) configurations where the depth and motion information is used to achieve coding gains with negligible complexity increase. The proposed DM\_FastMTT method is implemented and tested on VTM-14.0 reference software and [vvenc v1.3.0 software](https://github.com/fraunhoferhhi/vvenc).

The overall simulations results on top of VTM-14.0 with MTT disabled and vvenc v1.3.0 lowdelay faster preset are reported as below:

On top of VTM-14.0 with MTT0 (i.e. with MTT splitting disabled)

* LDP: -0.92%/-2.17%/-2.29% with 113% EncT and 95% DecT

On top of vvenc v1.3.0 lowdelay faster preset

* LDP: -1.48%/-1.74%/-1.91% with 122% EncT and 101% DecT

MTT0 loses 8.58% compared to VTM14, with 15% encoder time, this is the anchor of number above (called “test 1” in the contribution) In the contribution, further configurations are tested with higher compression and run time.

Contributors are asked to better refer to VTM14 rather than MTT0 anchor to understand the tradeoff of reduced encoder runtime vs. loss in performance.

Benefit very specific for this type of gaming sequences which come with depth and motion maps, algorithm would not work with other sequences.

Does it also give benefit in LB? Yes, according to observation of contributors

It is pointed out that it may not be suitable for hardware encoders, as more data need to be read. It might be necessary to clarify for which architecture (hardware, software) reduction of complexity is searched for.

Further study.

[JVET-Y0162](https://jvet-experts.org/doc_end_user/current_document.php?id=11356) AHG7: Gradual Decoding Refresh for ECM [S. Hong, L. Wang, K. Panusopone (Nokia), T. Poirier, G. Martin-Cocher (InterDigital)] [late]

This contribution presents some preliminary results of GDR implementation for ECM 3.0. GDR has been implemented in VTM, where constraints are imposed on coding tools to avoid possible leaks. Similar GDR implementation was ported into ECM 3.0. This GDR implementation has been tested with turning off ECM new coding tools, and no leaks have been observed.

Simulations were conducted under LDP configuration specified by LLCC AhG. The results show an overall performance loss against the anchor by 3.40%.

With new ECM tools enabled, the GDR crashes. Basically, disabling the tool converts ECM back to VTM. The purpose of this contribution is to provide information that the VTM GDR tool still works in principle in ECM.

[JVET-Y0163](https://jvet-experts.org/doc_end_user/current_document.php?id=11357) AHG7: New Gradual Decoding Refresh for ECM [L. Wang, S. Hong, K. Panusopone, M. M. Hannuksela, J. Lainema (Nokia)] [late]

This contribution presents a new GDR implementation for ECM. With this new GDR implementation, many constraints that have to be imposed on coding tools under the current VVC standard can be removed, which helps to reduce the code size for ECM significantly. In addition, this new GDR implementation provides great flexibility in future tool development for ECM.

Simulations were conducted under LDP configuration specified by LLCC AhG. The results demonstrate that

1. Without turning on ECM new coding tools, the new GDR performs better than VTM-like GDR by 1.11%, and
2. With ECM coding tools (except new loop-filter related coding tools), the new GDR has a small loss against the anchor by 2.50%.

New loop filter tools of ECM need to be disabled (as they are not using the virtual boundary which causes the problem), other tools can be used with the new GDR. This is probably causing part of the loss.

The method also is a change in decoding, not encoder only.

It is recommended to further investigate the method in the AHG study, and to study the problem of the conflict with the loop filters. However, implementing VB with loop filters does not have highest priority in this phase of exploration. Also in VVC it was implemented later to make ALF practical in the final specification for CTU row-wise processing. VB has hardly impact on compression performance, but is not simple to implement.

[JVET-Y0200](https://jvet-experts.org/doc_end_user/current_document.php?id=11407) Crosscheck of JVET-Y0163 (AHG 7: GDR without encoder constraints for ECM) [T. Poirier, K. Naser (InterDigital)] [late]

## AHG10: Encoding algorithm optimization (11)

Contributions in this area were discussed in session 21 at 2340–0140+1 UTC on Wednesday 19 Jan. 2022 (chaired by JRO) and in session 22 at 1520-1730 UTC on Thursday 20 Jan. 2022 (chaired by Frank Bossen).

[JVET-Y0048](https://jvet-experts.org/doc_end_user/current_document.php?id=11240) AHG10: study of layer bitrate allocation for spatial scalability in VTM [P. de Lagrange, F. Urban, G. Marquant (InterDigital)]

First results of VVC spatial scalability visual performance tests have been reported in JVET-Y0047, with equal bitrate for HD base layer and UHD enhancement layer. However, it was noticed that the bitrate balance has a significant effect on coding performance, in favor of more bitrate in the base layer. This contribution reports experiments with various bitrate allocations, and reports gains and lower encoding time over single layer in some cases.

Optimum bit rate allocation to base and enhancement layers highly sequence dependent.

Interesting study – problem is how to decide for a given sequence whichb is the best tradeoff for base/enhancement rate allocation?

Base layer should have a minimum reasonable quality if it is foreseen to be watchable in an application

Would it be possible to generate RD plots showing overall bit rate, where the upper convex hull points would represent the optimum allocation.

VVC scalable decoder in terms of hardware implementation is relatively simple.

Relationship with SR upscaling?

Further optimization in context of JVET-Y0047 strongly recommended.

[JVET-Y0077](https://jvet-experts.org/doc_end_user/current_document.php?id=11271) AHG10: Block importance mapping [P. Wennersten, J. Enhorn, C. Hollmann, J. Ström (Ericsson)]

This contribution proposes an updated version of the method in contribution JVET-V0057.

The proposed algorithm for VTM and HM signals CTU QP delta values in pictures that will be used for reference. The QP value to use for each CTU is based on the estimated importance of a given CTU for future pictures and the QP selected is in the range of −2 to +2 relative to the picture QP.

Presented SDR CTC results Y/U/V over VTM-15.0 are for RA -1.92%/-2.98%/-2.83%, for LDB -2.24%/-0.07%/-0.45% and for LDP -2.28%/0.12%/-0.31%. Respectively over HM-16.24 presented results for RA are -2.2%/-7.3%/-6.9%, for LDB -2.6%/-4.3%/-3.4% and for LDP -2.4%/-3.9%/-3.1%.

Presented results on HDR CTC RA over VTM-15.0 are DE100 -7.86%, PSNR-L100 -3.55%, wPSNR Y -1.23%, U -8.14%, V -7.09%, PSNR Y -1.33%, U -6.85%, V -6.63%.

It is proposed to adopt the algorithm into VTM and HM software.

QP changed on per-CTU basis. Criteria: Variance of block, and difference with adjacent motion compensated block.

Previous bug in HM has been corrected.

LD performs the MC based criterion only backwards.

DE100 known not to be very reliable.

Visual quality was checked, no problem. In contrary, static areas are preferred, which sometimes increases subjectively. Als on HDR? Was tested in a previous version, not much difference was visible.

The tool introduces kind of rate control. In the discussion, there are diverging opinions if rate control should be applied in CTC, where constant QP is well established.

Decision(SW): Adopt JVET-Y0077, both HM and VTM. Not consider for CTC at this moment. Could be interesting for verification tests.

[JVET-Y0083](https://jvet-experts.org/doc_end_user/current_document.php?id=11277) AHG4/AHG10: Report of Teleconference on Viewing Session Preparation for Deblocking [M. Wien, H. Zhang, X. Li]

See section 4.6

[JVET-Y0085](https://jvet-experts.org/doc_end_user/current_document.php?id=11279) AHG10: Report of Deblocking filter setting for VTM [H. Zhang, J. Jung, X. Li, S. Liu (Tencent)]

This proposal reports the objective quality and internal viewing test result of the new deblocking filter setting for VTM. The proposed method includes a set of encoder only settings for VTM deblocking filter. Basically, the deblocking filter control parameters BetaOffset\_div2 and TcOffset\_div2 are adjusted based on temporal layer ID for both luma and chroma components. The proposed method is implemented by only modifying VTM configuration files. It is reported that under the CTC, compared with VTM-15.0, the overall performance are as follows:

* AI: -0.74%/-0.69%/-0.44% with 100%EncT/99%DecT
* RA: -0.32%/-0.58%/-0.50% with 100%EncT/100%DecT
* LDB: -0.55%/-0.26%/-0.12% with 100%EncT/99%DecT
* LDP: -0.37%/0.00%/-0.05% with 100%EncT/99%DecT

It is reported that an internal remote expert viewing test was conducted by following the SC29/AG5 guidelines to investigate the visual quality impact. It is reported that there is no degradation on visual quality for the proposed method. It is also reported that additional objective metrics, including deep-learning based methods, further confirm this observation. It is reported that the result of internal viewing test is consistent with that of the viewing test conducted by AHG4 [6].

Having identical settings in ECM and VTM appears reasonable. Gain in ECM where the method is already used, is slightly higher.

Would it also be possible applying to HM? Has not been tried, but the proponents believe that the benefit mainly is present when in combination with ALF.

Decision(CTC): Change deblocking parameters as suggested in JVET-Y0085, to align settings of VTM with ECM.

[JVET-Y0202](https://jvet-experts.org/doc_end_user/current_document.php?id=11409) Crosscheck of JVET-Y0085 (AHG10: Report of Deblocking filter setting for VTM) [H.-J. Jhu (Kwai)] [late]

[JVET-Y0101](https://jvet-experts.org/doc_end_user/current_document.php?id=11295) AHG7/AHG10: Depth motion based fast Multi-Type Tree Splitting [S. Puri, K. Naser, T. Poirier, G. Martin-Cocher (InterDigital)]

See under section 4.10

[JVET-Y0105](https://jvet-experts.org/doc_end_user/current_document.php?id=11299) AHG10: An improved VVC rate control scheme [G. Ren, J. Jia, J. Wang, Z. Chen (Wuhan Univ.), Z. Liu (Tencent)]

This contribution presents some improvements based on the current R-lambda rate control algorithm. With the proposed algorithm, when using the anchor bit rate of VTM 14.0 as the target, there are -0.65%/-0.65%/-0.51% and -1.38%/-1.59%/-1.63% for Y/U/V coding efficiency improvements in low delay B and random access configuration.

The proposed method includes three parts:

* CTU-level bit allocation for skip and non-skip CTU.
* set smooth window size based on the GOP size and IntraPeriod.
* the quality dependency factor (QDF) based rate control was extended to low frame rate as well in this contribution.

It was noted that a large is observed for the Campfire sequence.

BD rate loss and lower rate accuracy is observed on Class F. May be mainly driven by the SlideShow sequence which features lot of “scene” changes. Seems to be related to modification 2.

Modifications 2 and 3 affect a few lines of code. Modification 1 is more involved (no more than 100 LOC).

Modifications should replace exisiting methods.

Decision(SW): Include items 2 and 3 in next VTM release.

Further study encouraged on item 1.

[JVET-Y0176](https://jvet-experts.org/doc_end_user/current_document.php?id=11383) Crosscheck of JVET-Y0105 (AHG10: An improved VVC rate control scheme) [Y. Wang (Bytedance)] [late]

[JVET-Y0118](https://jvet-experts.org/doc_end_user/current_document.php?id=11312) AHG10: On Temporal-Layer-Based ChromaQP Coding [K. Sato, Y. Yu, H. Yu, Z. Xie, L. Xu, F. Wang, H. Huang, J. Gan, D. Wang (OPPO)]

With the hierarchical B-picture structure employed in RA CTC condition, the image quality of the lower temporal layers will affect the image quality of the higher temporal layers. Therefore, coding efficiency improvement can be realized to enhance image quality of the lower temporal layers by for example spending more bits. Also, computational cost reduction can be realized to apply coding tools just to lower temporal layers.

This contribution proposes a temporal-layer-based chroma QP setting as a non-normative change. The simulation results show that the following gain is obtained for Y, Cb, Cr with RA test condition:

-0.14%, -1.37%, -1.36% over ECM3.1

-0.21%, -1.43%, -1.34% over VTM15.0

Additionally, the proposed method is tested on top of JVET-Y0126 encoder optimization. RA and LDB conditions are tested. Gain over Y0126 (on top of VTM14.2) is as follows:

-0.17 %, -1.27 %, -1.22 % in RA condition

0.45 %, -8.89 %, -9.12% in LD condition

This contribution also recommends establishing an AHG to study on the VPS and SPS syntax related to temporal-layer based coding, for deblocking filter, partitioning information, and other parameters as well as chroma qp offset.

It was noted that run time numbers are not accurate (heterogenous grid).

Modification is to configuration file only. No changes to code.

Q: how does PSNR vary from frame to frame? Is it more or less consistent? Could be further studied.

Large variations from sequence to sequence. Large luma gain (and chroma loss) on Campfire whereas small luma loss may be common otherwise. Overall luma gain driven by gain on single Campfire sequence.

Q: what is motivation? More compression gain (as opposed to improving luma/chroma balance).

Seems to shift performace towards chroma whereas for balance the opposite direction should be taken.

Existing AHG mandates should cover possible study. No need for new AHG.

Further study encouraged.

[JVET-Y0236](https://jvet-experts.org/doc_end_user/current_document.php?id=11445) Crosscheck of JVET-Y0118 (AHG10: On Temporal-Layer-Based ChromaQP Coding) [H. Zhang (Tencent)] [late]

[JVET-Y0126](https://jvet-experts.org/doc_end_user/current_document.php?id=11320) AHG10: VTM encoder configurations for tests targeting improved coding performance [D. Rusanovskyy, M. Karczewicz (Qualcomm), K. Andersson, R. Sjöberg, L. Litwic (Ericsson), P. Nikitin, G. Martin-Cocher (InterDigital), A. Wieckowski, J. Brandenburg, B. Bross (HHI)]

In this contribution, a set of encoder-only modifications and settings targeting improved compression efficiency of VTM reference software is proposed. Proposed method is implemented as minor encoder software change on top VTM-14.2 and encoder configuration files for RA and LD. It is reported that under the CTC, proposed method provides the following BD-rate gain vs. VTM-14.2 anchor:

SDR:

* RA: -2.66%/-6.24%/-6.13% with 168%EncT/96%DecT.
* LDB: -2.6%/-2.2%/-2.1% with 150%EncT/97%DecT.
* LDP: -3.0%/-2.6%/-2.8% with 171%EncT/101%DecT.

HDR RA:

* PQ: -2.6%/-4.1%/-2.1% (wPSNR) with 167%EncT/110%DecT.
* HLG: -3.3%/-3.9%/-5.3% (PSNR) with 165%EncT/90%DecT.

Against the VTM-15.0 anchor, same RA results are achieved. For LDB and LDP configuration, BD-rate gain of -1.8%/-1.4%/-1.2% and -1.9%/-1.4%/-1.5% are reported, respectively.

It is proposed to modify VTM and:

* Add command-line option to control RDO cost factor for partitioning (default behaviour unchanged)
* Add command-line option to control fast decision for TT from BT (default behaviour unchanged)
* Speed-up RPR handling (bit-exact change)

It is further proposed to:

* include configuration files for tests targeting improved coding performance, for example for benchmarking purposes. Such files would be put into a separate configuration folder.

No changes to VTM CTC are requested.

It was noted that some changes may be applied to ECM as well.

Decision(SW): include proposed VTM changes into next VTM release

Decision(SW): include proposed configuration files in separate folder in next VTM release (no change to CTC)

Further study, including applicability to ECM, is encouraged.

[JVET-Y0194](https://jvet-experts.org/doc_end_user/current_document.php?id=11401) Cross-check of JVET-Y0126 (AHG10: VTM encoder configurations for tests targeting improved coding performance) [J. Lainema (Nokia)] [late]

[JVET-Y0208](https://jvet-experts.org/doc_end_user/current_document.php?id=11415) Crosscheck of JVET-Y0126 (AHG10: VTM encoder configurations for tests targeting improved coding performance) [F. Pu (Dolby)] [late]

[JVET-Y0244](https://jvet-experts.org/doc_end_user/current_document.php?id=11453) Crosscheck of JVET-Y0155 and combination with JVET-Y0126 [K. Sühring (HHI)] [late]

[JVET-Y0152](https://jvet-experts.org/doc_end_user/current_document.php?id=11346) AHG10: Fast skip of TT split partitioning on top of ECM reference software [L.-F. Chen, X. Li, S. Liu (Tencent)]

TBP

[JVET-Y0238](https://jvet-experts.org/doc_end_user/current_document.php?id=11447) Crosscheck of Y0152 (AHG10: Fast skip of TT split partitioning on top of ECM reference software) [C.-W. Kuo (Kwai)] [late]

[JVET-Y0155](https://jvet-experts.org/doc_end_user/current_document.php?id=11349) AHG10: Fixes and clean up for temporal prefilter [F. Bossen (Sharp)] [late]

TBP

[JVET-Y0243](https://jvet-experts.org/doc_end_user/current_document.php?id=11452) Cross-check of JVET-Y0155: AHG10: Fixes and clean up for temporal prefilter [J. Enhorn, K. Andersson, P. Wennersten (Ericsson)] [late]

[JVET-Y0244](https://jvet-experts.org/doc_end_user/current_document.php?id=11453) Crosscheck of JVET-Y0155 and combination with JVET-Y0126 [K. Sühring (HHI)] [late]

[JVET-Y0177](https://jvet-experts.org/doc_end_user/current_document.php?id=11384) AHG 10: Enhanced deblocking settings for VTM CTC [K. Andersson, J. Enhorn, R. Sjöberg, J. Ström, L. Litwic (Ericsson)] [late]

Deblocking can provide a significant subjective benefit for VVC. This contribution proposes setting the deblocking parameter beta offset to -2 in the CTC configuration files. The beta offset controls the amount of spatial activity that is allowed until deblocking is turned off. It is asserted by the proponents that this change improves the objective performance of the deblocking filtering while maintaining the subjective quality benefit.

The objective benefit of the approach is as follows:

AI: -0.28%Y, -0.04%U, -0.03%V

RA: -0.25%Y, -0.09%U, -0.12%V

LDB: -0.43%Y, -0.03%U, -0.07%V

LDP: -0.34%Y, 0.29%U, -0.13%V

No change in encoding or decoding time.

Similar approach as in JVET-0085. However the tcOffset is not changed, and beta is less aggressively changed, as proponents believe that otherwise artifacts might occur.

Were similar settings applied for ECM? Not so far.

The proponent claims that, when turning on RDO with deblocking for both proposals, the performance of JVET-Y0177 gets comparably better than with JVET-0085.

As it was agreed to adopt the parameter settings of JVET-Y0085, no action is taken on JVET-0177.

[JVET-Y0245](https://jvet-experts.org/doc_end_user/current_document.php?id=11454) Cross-check of JVET-Y0177: AHG 10: Enhanced deblocking settings for VTM CTC [A. Segall (Sharp)] [late]

[JVET-Y0212](https://jvet-experts.org/doc_end_user/current_document.php?id=11420) AHG4: REV Result for AHG11/EE1 and AHG10/Deblocking [M. Wien (RWTH)]

See section 4.6

[JVET-Y0240](https://jvet-experts.org/doc_end_user/current_document.php?id=11449) AHG10: Block importance mapping for ECM [P. Wennersten, J. Enhorn, C. Hollmann, J. Ström (Ericsson)] [late]

TBP

## Profile/tier/level specification (5)

Contributions in this area were discussed in session 5 at 1407–1505 UTC on Thursday 13 Jan. 2022 and in session 9 at 1350–1415 UTC on Friday 14 Jan. 2022 (chaired by JRO).

[JVET-Y0056](https://jvet-experts.org/doc_end_user/current_document.php?id=11250) AHG2: High tier for lower levels [S. Keating, A. Browne, K. Sharman (Sony)]

This contribution proposes to remove the defined MinCrBase values for high tier at levels below 4.0 to make clear that high tier is not supported at these levels.

This is asserted to make the spec text clearer.

Decision(ed.): Adopt the modification of text in JVET-Y0056

[JVET-Y0057](https://jvet-experts.org/doc_end_user/current_document.php?id=11251) AHG2: MinCr for still picture profiles [S. Keating, A. Browne, K. Sharman (Sony)]

This contribution proposes defining a minimum compression ratio for still picture profiles. It is asserted that this would be useful for encoder and decoder design to set maximum buffer sizes etc.

It is pointed out that the only constraint imposed by MinCR is decoder timing. Therefore, it is not needed for still picture decoding. Timing is only relevant for video bitstreams, while CPB constraints are still applying in SP profiles in the levels defined.

V2 still picture profiles should follow the same design principles as V1 SP profiles which don’t impose any constraint based on MinCR.

No action necessary.

It was suggested to add a note which clarifies which constraints do or do not apply in case of still picture profiles. Editors should consider if that can be put in some appropriate place of the spec.

[JVET-Y0063](https://jvet-experts.org/doc_end_user/current_document.php?id=11257) AHG2: On Main 10 4:4:4 Still Picture profile for VVC v1 and v2 [M. Pettersson, R. Sjöberg, M. Damghanian (Ericsson)]

In VVC v1 there is a requirement for decoders conforming to the Main 10 4:4:4 Still Picture profile that they shall also be capable of decoding the first picture of a Main 10 4:4:4 bitstream. However, there is no requirement that such decoder shall also be capable of decoding the first picture of a Main 10 bitstream.

The proponents assess that this is an oversight and proposes to add to the errata of VVC v1, and consequently also to VVC v2, that a Main 10 4:4:4 Still Picture decoder shall also be capable of decoding the first picture of a Main 10 bitstream.

It is agreed that this is an obvious modification.

Decision: Adopt JVET-Y0063 (both for v1 and v2 Main10 4:4:4 Still Picture Profiles).

[JVET-Y0072](https://jvet-experts.org/doc_end_user/current_document.php?id=11266) New Levels for HEVC and VVC [A. M. Tourapis, D. Singer, K. Kolarov (Apple)] [late]

Presented in session 9

The currently specified levels for HEVC and VVC limit the usage of these coding specifications to up to video resolutions of 8192x4320 at 120fps. This document requests the creation of new levels that would enable higher resolution applications up to 16384x8640 at 120fps. This document also proposes the extension of the unconstrained level 8.5 for video applications. It is claimed that such an extension could enable content or service providers to create valid bitstreams for new applications that may exceed the constraints of the existing levels, without having to wait for the definition of new appropriate levels.

Tables in the contribution are defining the levels for HEVC

Several experts supported definition of levels for 16K video, as devices are starting to appear.

The proposed level 8.5 (for video) has no constraints in bit rate, picture size, frame rate etc. A note might be useful to add that it may not require guaranteed real-time decoding.

Decision: Adopt JVET-Y0072

Include in HEVC CDAM (and JVET-Y1005), along with previous text of JVET-X1005, and potentially some errata items from JVET-X1004, with the intent of converting to a new edition incuding Amd.1.

It was further agreed (session 20 Wed 19 Jan 2200 UTC) to add a corresponding unconstrained level (probably 15.5) in VVC Amd.1 (WD as output from the current meeting).

[JVET-Y0099](https://jvet-experts.org/doc_end_user/current_document.php?id=11293) VVC level 4.2 [G. Martin-Cocher (InterDigital)]

This contribution proposes to discuss whether a VVC level 4.2 is needed to address use-cases such as Cloud/Online Gaming as results on some non-CTC 1080p sequences indicates that a bitrate above the main and high tier limit is reached. It may be desirable to rely on a “Main Tier” to address this use-case rather than on a “high tier”.

It is pointed out that the proposal would have the problem that level 5.0 main tier would not be capable to decode the proposed 4.2 main tier. A possible solution would be to change the max bitrate in main tier from 50000 to 25000, and changing high tier from 75000 to 100000. This would be the same bitrate as 5.0, but smaller picture size.

Several concerns were expressed if such a new level is needed or would be used practically. Further evidence sought.

## Proposed modification of system interface (0)

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

See the notes on this topic from the joint meeting in section 7.4.

# Low-level tool technology proposals

## AHG8: High bit rate and high bit depth coding for VVC (2)

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

[JVET-Y0049](https://jvet-experts.org/doc_end_user/current_document.php?id=11241) AHG2/AHG8: On the range extensions GCI flags [Y.-K. Wang (Bytedance)]

[JVET-Y0190](https://jvet-experts.org/doc_end_user/current_document.php?id=11397) AHG2/AHG8: Suggestions for the operation range extensions GCI [J. Gan, Y. Yu, H. Yu (OPPO)] [late]

## AHG11: Neural network-based video coding (23)

### Summary and BoG reports

[JVET-Y0023](https://jvet-experts.org/doc_end_user/current_document.php?id=11422) EE1: Summary of Exploration Experiments on Neural Network-based Video Coding [E. Alshina, S. Liu, W. Chen, F. Galpin, Y. Li, Z. Ma, H. Wang]

Presented and discussed in session 2 Wed 12 Jan. 1605-1725 (chaired by JRO)

This contribution provides a summary report for the Exploration Experiment 1 on Neural Network-based Video Coding.Originally 7 NN-based technologies (plus VVC RPR) have been included into EE1. Among them: 3 NN-based filter, 3 NN-based super resolution and one NN-based Intra techniques. One test (super resolution) was dropped, 3 EE related contributions with further development tools under study are submitted to this meeting. Several variants of NN-based in-loop filters with complexity 500...2500 kMAC/pxl providing average bit-saving in a range 2...11% (in random access configuration) have been demonstrated in this round of Exploration Experiment. Several variants of super-resolution (from available in VVC standard functionality to NN-based) demonstrate 3...6% bit-saving in random access configuration for 4K content. NN-based Intra tool improves Intra coding by approx. 3%.

Tests in this Exploration Experiment follow AhG11 testing conditions and complexity assessment methodology [1]. Anchor is VTM11.0 with improved GOP-based temporal filter [2] enabled. In total 5 quality levels (corresponding QP=22, 27, 32, 37, 42) are used for objective performance measurement.

Additionally in this round of EE several proposals have been tested on top of ECM SW. Some proponents used ECM3.0, another ECM3.1 as SW basis/anchor, so very accurate performance comparison is not possible. For best NN-based in-loop filters gain on top of AhG11 anchor is luma 8-10%, but drops to 5-7% on top of ECM. So, overlap in performance between NN-based under study in EE1/AhG11 and “classical” tools under study in EE2/AhG12 is about 2-3%.

AhG11 has agreed before [3] to summarize results in graphic forms, showing compression performance (BD-rate) vs kMAC/pxl and BD-rate vs Total number of model parameters (considering those two not only, but the most important complexity characteristics).

Fig. 1 shows average BD-rate in Random Access configuration vs kMAC/pxl for different EE1 and related proposals. Fig. 2 shows average BD-rate in Random Access configuration vs Total number of parameters for different EE1 and related proposals.

Super-resolution proposals do not appear on those plots. Two proposals ,which are based on VVC RPR (2D separable filter of up-sampling), which is applied adaptively and so show average BD-rate gain, have too low complexity to be show next to NN-based technologies. Two NN-based Super-resolution proposals included into this EE are used unconditionally for all sequences, so show BD-rate drop in average. Merging of two gain bringing aspects (adaptive resolution coding and NN-based up-sampling) is still subject for further study.

**Fig. 1.** Average BD-rate in Random Access configuration vs kMAC/pxl for different EE1 and related proposals. The dash line indicates capability of NVIDIA RTX3080 for 4K@60 processing.

**Fig. 2.** Average BD-rate in Random Access configuration vs Total number of parameters for different EE1 and related proposals.

In total 5 tests in this EE have been cross-checked (thanks Ericsson, OPPO and Huawei for their efforts and comments). Addressing comments from cross-checker should greatly help collaborative work on NN-based tools.

Majority of EE1 participants used PyTorch/libtorch (v1.3, 1.4, 1.6, 1.8, 1.9) for their training/tests and float-point implementation. Conducting cross-checks or/and collaborative development on NN-based tools w/o agreement on particular version of libtorch (if it will decided to use libtorch for future development) is complicated.

Due to the float point arithmetic if (for example) encoding is done on linux, decoding on windows then mismatch is observed.

One proponent used SADL library [4] and int16 implementation. Cross-check for this proposal was successful (matching results, no issue with encoding decoding on different platforms) even at JVET-X meeting.

Results on combination of intra prediction with ECM:

* Simplified (less complex) version was used which may be one reason of less gain
* Results with AI are provided below in EE1-3

EE1-1 (loop filtering)

Two proposals perform significantly better than other in this category. It makes sense to compare them with more details in Table 1.

|  |  |  |
| --- | --- | --- |
| Test | EE1-1.1 (Tencent)  [JVET-Y0078](https://jvet-experts.org/doc_end_user/current_document.php?id=11272) | EE1-1.2 (Bytedance and Qualcomm)  [JVET-Y0143](https://jvet-experts.org/doc_end_user/current_document.php?id=11337) |
| Filter-structure |  |  |
| Attention module | Use prediction and reconstruction | |
| both are as inputs to NN | weight between NN-module input and output  *for F\_out = F\_in f (Rec, Pred, BS, QP) + F\_in*  *f* comprises 2 convolutional layers |
| partitioning map | BS, QP |
| Stride2 convolution | yes | yes |
| Colors | 3 colors processed together | separate processing for Luma and Chroma |
| N  – number of Res Blocks | 24 | 8 |
| M  – number of channels / features | 64 (192)  goes up before and then goes down after the activation layer | 96 |
|  |  |  |
| Position | I-slice: replaces DF, SAO  B-slice: switching between proposed NN-filter and DF & SAO | replaces DF, SAO |
| Control | CTU and slice level on/off | Filter selected from 3 candidates {q, q - 5, q - 10}  on/off control as well as the conditional parameter index signalled per 128x128 |
| Content adaptation | scaling factor signaled in slice header | scaling factor signaled in picture header for each color component |
| Total number of models | 2 | 4 |
| Block extension | 8 samples each side | 8 samples each side |
| kMAC/pxl | 509 | 539 |
| Total Parameters (Millions) | 3.1 | 6.3 |
| Implementation | Torch v1.9 / float32 | Pytorch v1.6 / float32 |
| BD-rate vs VTM | **8.6**% (Luma) / 20% (Chroma) | **9.8**% (Luma) / 21% (Chroma) |
| BD-rate vs ECM | 5.3% (Luma) / 15% (Chroma) | 7.2% (Luma) / 18% (Chroma) |

From this analysis one can see that filters are conceptually very similar. Both were included into subjective quality assessment test. Both have attracted attention of cross-checkers and there are EE1-releated contribution which further improve both filters.

If combined with EE2 tools available in ECM3.1 then JVET-Y0143 can provide 21% (Luma) / 33% (Chroma) gain over VTM (in RA configuration, QP=22,...,42).

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | ECM3.1 over VTM-11.0\_nnvc-1.0 | | | | | ECM3.1 & EE1-1.2 over VTM-11.0\_nnvc-1.0 | | | | |
|  | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU | Y-PSNR | U-PSNR | V-PSNR | EncT | DecT CPU |
| Class A1 | -16.4% | -16.4% | -22.3% | 4.6 | 4.8 | -22.4% | -27.7% | -34.8% | 5.5 | 476 |
| Class A2 | -16.6% | -21.2% | -20.7% | 4.5 | 5.2 | -23.1% | -33.7% | -36.1% | 5.2 | 462 |
| Class B | -13.7% | -20.9% | -20.0% | 4.2 | 4.9 | -19.6% | -35.0% | -35.0% | 4.8 | 422 |
| Class C | -15.0% | -17.2% | -16.4% | 4.0 | 4.7 | -21.2% | -31.6% | -32.1% | 4.3 | 331 |
| Class E |  |  |  |  |  |  |  |  |  |  |
| **Overall** | **-15.2%** | **-19.1%** | **-19.6%** | **4.3** | **4.9** | **-21.3%** | **-32.4%** | **-34.4%** | **4.9** | **413** |
| Class D | -15.4% | -16.9% | -15.8% | 3.9 | 4.9 | -22.7% | -33.0% | -33.4% | 4.1 | 296 |
| Class F | -13.6% | -19.4% | -19.2% | 3.4 | 3.9 | -16.8% | -27.2% | -27.2% | 4.7 | 181 |

Series of tests from Dahua EE1-1.3.1, 2 and 3 show not that good complexity performance trade-off, but provides the group with useful information about performance impact of taking Chroma information for Luma NN-based enhancement and Luma information for Chroma NN-based enhancement:



(a)



(b)

Table below demonstrates that Chroma as extra input for Luma NN-based in-loop filtering provides ~0.3% Luma gain (with negligible increment of complexity). Luma as extra input for Chroma NN-based in-loop filtering brings 11% Chroma gain, while computational complexity roughly doubles.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Test | Description | Total Number of Parameters (Millions) | Worst Case Complexity (kMAC/pixel) | Random Access (CTC) | | |
| Y | Cb | Cr |
| EE1-1.3.1 | Luma only (a) without chroma input | 12.2 | 1902.6 | -2.3% | -0.2% | -0.3% |
| EE1-1.3.2 | Luma only (a) with chroma input | 12.2 | 1904.4 | -2.6% | -0.2% | -0.3% |
| EE1-1.3.3 | (a) + (b) | 20.0 | 2510.6 | -2.7% | -11.8% | -11.2% |

No implementation of filters in SADL library available, but could hypothetically be possible.

The analysis above shows that potentially a convergence of 1-1.1 and 1-1.2 could be possible

Results combined with ECM are different from EE2 conditions due to different QP range (5 QPs) used here. Results with the normal 4 QPs should be added.

Anchor of combined test is VTM 11 (the usual anchor of EE1, which has a minor difference of MCTF).

EE1-2 (super resolution)

There are proposals which use VVC RPR content adaptively and so show gain (or at least no loss) for all sequences. And there are proposals which encode all content in smaller resolution (so in average show drop). Before those two gain-bringing aspects will be combined one can compare technologies in this category only looking on 4K results (Fig. 3 and Fig. 4).

One of EE1-related proposals shows reasonable-speed encoder which decides about content adaptive enabling of RPR, and so allows to obtain compression gain and have no cross of RD-curves w/o multi-pass. It is highly recommended to combine this of similar technology with NN-based super resolution in future EE (if any) to be able eventually compare average (not only 4K) gain of all NN-based tools (including super-resolution).

Subsequent figures need change from v2

**Fig. 3.** Average BD-rate for 4K sequences in Random Access configuration vs kMAC/pxl for different EE1 and related proposals in super resolution category. The dash line indicates capability of NVIDIA RTX3080 for 4K@60 processing.

**Fig. 4.** Average BD-rate for 4K sequences in Random Access configuration vs Total number of parameters for different EE1 and related proposals in super resolution category.

Observations of cross-checkers in this category is as follows. NN-based super-resolution processing is often not performed after decoding (PSRN, MS-SSIM metrics are measured based on encoder reconstruction), and so NN-based processing doesn’t appear in decoding run-time. This must be fixed in a future.

Only 4K is shown here because the NN based methods are not using adaptive operation (unlike the RPR method) and would have loss in lower resolution. However, even for 4K, performance would be better with adaptive operation. There are EE related contributions which resolve that.

Still, there is typically loss in chroma (regardless if RPR or NN).

EE1-3 (intra prediction)

It makes sense to look closely onto test results in “all-intra” configuration while discussing Intra tools. The comparison of NN-filter and NN-Intra tests in this EE is shown in Fig. 5.

Subsequent figure needs change from v2

**Fig. 5.** Average BD-rate in “all intra” configuration vs kMAC/pxl for different EE1 proposals. The dash line indicates capability of NVIDIA RTX3080 for 4K@60 processing.

The only NN-based Intra tool was implemented using SADL library [4] using 16 bits integer implementation. Complexity of NN-based intra tools (measured in kMAC/pxl) is significantly lower than complexity of NN-based filter, while gain is substantial. Likely NN-based Intra and NN-based filters utilize different sources of compression gain. It make sense to test combination.

Insert new results (table) on combination with ECM in v2 of EE report

In the discussion, it is also pointed out that an important complexity aspect to be considered in intra prediction is the lack of parallelization. It needs to be computed sequentially, unlike loop filters and super resolution. The original predecessor of MIP had a similar problem. Practical implementation of NN technology in a hybrid architecture probably needs more study.

It is further commented that the minimum number of computational cycles is equivalent to the number of stages/layers in the network.

Visual test:

Two best performing NN-based filter, VVC RPR and typical representative of NN-based super-resolution technology were recommended to be include into the viewing.

Sequences and QP for viewing were discussed at AhG11/EE1 teleconference conducted at Nov. 16. Details on this discussion are summarized in [5]. Viewing was conducted in 3 different sessions: 2160p, 1080p, 720p. Each video was limited to be 5 seconds only (to shorten overall viewing time). Several sequences out of AhG11 CTC were included into viewing sessions.

Several issues have been identified during viewing preparation:

* ffmpeg version was not the same for different proponents who prepared viewing and naming and organization of mp4 files was inconsistent across test points;
* way of cutting 5 second part of sequences was not the same for different proponents who prepared viewing.

Proposed solutions:

* Specify ffmpeg version in viewing preparation instruction (latest 64bit-version recommended)
* In order to cut 5 sec part of the video first cut then create mp4. Example for creation “last 5 sec” video from Campfire sequence:
  + ffmpeg -s:v 3840x2160 -c:v rawvideo -pix\_fmt yuv420p10le -i Campfire\_3840x2160\_30fps\_bt709\_420\_videoRange.yuv -filter:v select="between(n\, 150\, 299)" temp.yuv
  + ffmpeg -s:v 3840x2160 -c:v rawvideo -pix\_fmt yuv420p10le -r 30 -i temp.yuv -c:v libx265 -crf 15 -tag:v hvc1 -pix\_fmt yuv420p10le Campfire\_3840x2160\_30fps\_10bit\_420.mp4
* Specify a precise naming convention and directory structure for provided mp4 files.
* Specify a time line for provision of test material including crosscheck among participating proponent organizations.

Viewing was conducted on Jan. 10th and Jan. 11th, results are reported in JVET-Y0212.

[JVET-Y0045](https://jvet-experts.org/doc_end_user/current_document.php?id=11237) AHG11/EE1 viewing preparation report [[E. Alshina](mailto:elena.alshina@huawei.com), M. Wien, A. Segall]

See section 4.6

[JVET-Y0212](https://jvet-experts.org/doc_end_user/current_document.php?id=11420) AHG4: REV Result for AHG11/EE1 and AHG10/Deblocking [M. Wien (RWTH)]

See section 4.6

[JVET-Y0239](https://jvet-experts.org/doc_end_user/current_document.php?id=11448) AHG11: “Overview of technologies considered in JVET’s neural network-based video coding exploration” presentation slides from the ISO/IEC JTC1/SC29/AG4 Workshop on Image-related Activities, Current and Future [A. Segall (Sharp)] [late]

For information – no presentation necessary.

[JVET-Y0241](https://jvet-experts.org/doc_end_user/current_document.php?id=11450) AhG11 “Methodologies for evaluation and complexity assessment of neural network-based video coding technology” presentation slides from the ISO/IEC JTC1/SC29/AG4 Workshop on Image-related Activities, Current and Future [E.Alshina (Huawei)] [late]

For information – no presentation necessary.

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

Contributions in this area were discussed in session 6 at XXXX–XXXX UTC on XXday X Jan. 2022 (chaired by XXX).

[JVET-Y0061](https://jvet-experts.org/doc_end_user/current_document.php?id=11255) EE1-2.1: Super Resolution with existing VVC functionality [E. Alshina, J. Sauer (Huawei]

[JVET-Y0214](https://jvet-experts.org/doc_end_user/current_document.php?id=11423) Cross-check of JVET-Y0061: EE1-2.1: Super Resolution with existing VVC functionality [K. Andersson (Ericsson)] [late]

[JVET-Y0069](https://jvet-experts.org/doc_end_user/current_document.php?id=11263) EE1-2.3: CNN-based Super Resolution for Video Coding Using Decoded Information [C. Lin, Y. Li, K. Zhang, L. Zhang (Bytedance)]

[JVET-Y0169](https://jvet-experts.org/doc_end_user/current_document.php?id=11376) Crosscheck of JVET-Y0069 (EE1-2.3: CNN-based Super Resolution for Video Coding Using Decoded Information) [J. Sauer (Huawei)] [late]

[JVET-Y0070](https://jvet-experts.org/doc_end_user/current_document.php?id=11264) EE1-2.4: CNN-based Super Resolution for Video Coding Using Separate Networks for Chroma Components [C. Lin, Y. Li, K. Zhang, L. Zhang (Bytedance)]

[JVET-Y0078](https://jvet-experts.org/doc_end_user/current_document.php?id=11272) EE1-1.1: neural network based in-loop filter with constrained storage and low complexity [L. Wang, X. Xu, S. Liu (Tencent)]

[JVET-Y0188](https://jvet-experts.org/doc_end_user/current_document.php?id=11395) Cross-check of JVET-Y0078 (Test 1.1): EE1-1.1: Neural network based in-loop filter with constrained storage and low complexity [Z. Xie (OPPO)] [late]

[JVET-Y0082](https://jvet-experts.org/doc_end_user/current_document.php?id=11276) EE1-3.1: intra prediction using neural networks [T. Dumas, F. Galpin, P. Bordes, E. François (InterDigital)]

[JVET-Y0187](https://jvet-experts.org/doc_end_user/current_document.php?id=11394) Cross-check of JVET-Y0082 (Test 3.1.2): EE1-3.1: Intra prediction using neural networks [L. Xu, Y. Yu, Z. Dai (OPPO)] [late]

[JVET-Y0084](https://jvet-experts.org/doc_end_user/current_document.php?id=11278) EE1-1.3: A Deep In-Loop Filter [X. Zhang, D. Jiang, J. Lin, C. Fang, S. Peng (Dahua)]

[JVET-Y0224](https://jvet-experts.org/doc_end_user/current_document.php?id=11433) Cross-check of JVET-Y0084 (EE1-1.3: A Deep In-Loop Filter) [K. Takada (Sharp)] [late]

[JVET-Y0143](https://jvet-experts.org/doc_end_user/current_document.php?id=11337) EE1-1.2: Test on Deep In-Loop Filter with Adaptive Parameter Selection and Residual Scaling [Y. Li, K. Zhang, L. Zhang (Bytedance), H. Wang, K. Reuzé, A.M. Kotra, M. Karczewicz (Qualcomm)]

[JVET-Y0173](https://jvet-experts.org/doc_end_user/current_document.php?id=11380) Crosscheck of JVET-Y0143 (EE1-1.2: Test on Deep In-Loop Filter with Adaptive Parameter Selection and Residual Scaling) [J. Sauer (Huawei)] [late]

[JVET-Y0186](https://jvet-experts.org/doc_end_user/current_document.php?id=11393) Cross-check of JVET-Y0143 (Test 1.2.2): EE1-1.2: Test on Deep In-Loop Filter with Adaptive Parameter Selection and Residual Scaling [Z. Dai (OPPO)] [late]

[JVET-Y0166](https://jvet-experts.org/doc_end_user/current_document.php?id=11373) Crosscheck of EE1-1.2.1 from JVET-Y0143 (EE1-1.2: Test on Deep In-Loop Filter with Adaptive Parameter Selection and Residual Scaling) [J. Ström (Ericsson)] [late]

### EE1 related contributions: Neural network-based video coding (5)

Contributions in this area were discussed in session 6 at 1520–XXXX UTC on Thursday 13 Jan. 2022 (chaired by JRO).

[JVET-Y0068](https://jvet-experts.org/doc_end_user/current_document.php?id=11262) EE1-2.1-related: RPR encoder with multiple scale factors [J. Nam, S. Yoo, J. Lim, S. Kim (LGE)]

This contribution proposes RPR encoder control using multiple scale factors, such as x2.0 (half size) and x1.5 (2/3 size). At each GOP, the encoder can select a scale factor based on initial QP and a PSNR value which is calculated between the source and down-up sampling pictures for the first picture of each GOP. The experimental results are reportedly shown as below,

* For the range of QP 22 to 42 case,
  + AI: -0.57% / 3.14% / 2.73% / 98% / 83% for Y/U/V/Enc. Time/Dec. Time
  + RA: -0.98% / 2.62% / 2.50% / 92% / 80% for Y/U/V/Enc. Time/Dec. Time
* For the range of QP 27 to 47 case,
  + AI: -1.67% / 6.18% / 4.87% / 102% / 73% for Y/U/V/Enc. Time/Dec. Time
  + RA: -2.42% / 5.82% / 5.03% / 89% / 71% for Y/U/V/Enc. Time/Dec. Time

Performance worse than EE1-2.1, but encoder is less complex (avoids multi-pass coding)

How often is x1.5 selected? Mostly around QP32.

PSNR curves are not crossing (not becoming worse than full res anchor)

Similar approach could be used with any SR method

It is recommended to use this method as a new reference for EE on superresolution, as it allows to compare NN based SR versus RPR with similar degree of encoder optimization.

Include in EE

It is mentioned that this could also be used with ECM.

[JVET-Y0228](https://jvet-experts.org/doc_end_user/current_document.php?id=11437) Cross-check of JVET-Y0068: EE1-2.1-related: RPR encoder with multiple scale factors [Kenneth Andersson (Ericsson)] [late]

[JVET-Y0079](https://jvet-experts.org/doc_end_user/current_document.php?id=11273) EE1-1.1-related: the result of neural network based in-loop filter on ECM [L. Wang, X. Xu, S. Liu (Tencent)]

This contribution reports the result of JVET-Y0078 on ECM. The proposed filter in EE1-1.1.1 is integrated on ECM-3.0, and the corresponding 2 models are both used directly without any other retaining processes or refinements. Compared with ECM-3.0, {5.32%, 13.97%, 14.62%}, {4.48%, 13.03%, 13.73%} and {4.54%, 13.07%, 12.87%} BD-rate savings are reported with RA, LB and AI configurations, respectively. Compared with EE1 anchor, ECM3.0 with the proposed method can achieve {19.64%, 29.89%, 31.22%}, {11.14%, 24.34%, 24.98%} BD-rate savings with RA and AI configurations, respectively.

5 QPs are used as in EE1. Would be interesting to see results with 4 QPs as in EE2.

Filter is added as another stage before deblocking. Deblocking/SAO are still active in ECM (this is different from the EE).

Model is the same as EE1-1.1.1, not re-trained

[JVET-Y0080](https://jvet-experts.org/doc_end_user/current_document.php?id=11274) EE1-1.1-related: alternative filter designs [L. Wang, X. Xu, S. Liu (Tencent)]

In this contribution, two neural network based in-loop filters are proposed as the alternative filter designs for EE1-1.1. Multiple models and single model are trained for the proposed filter 1 and filter 2, respectively. In addition to the reconstruction image, other side information is fed into the network, such as prediction image, partition image. In this contribution, these two filters are both evaluated on EE1 anchor. The proposed filter 1 can reportedly achieve {10.63%, 26.10%, 25.78%}, {9.43%, 21.97%, 20.54%} and {7.53%, 20.45%, 20.14%} BD-rate savings with RA, LB and AI configurations, respectively. The proposed filter 2 can reportedly achieve {8.16%, 19.21%, 20.47%}, {x.xx%, x.xx%, x.xx%} and {6.51%, 15.47%, 16.94%} BD-rate savings with RA, LB and AI configurations, respectively.

Deblock/SAO are replaced for I slices, and for B slices switching is performed at CTU level.

What is the benefit of applying ALF after the filter? It is mainly improving chroma performance.

Performance is much better than for original proposal X0054. Mainly better training (including L2 loss function)

In filter 1, model is selected at CTU level is based on base QP, no input of QP into the network directly

For filter 2, scaling factor (3 parameters) is signalled at slice level

It is mentioned that switching a large model at CTU level could be undesirable. Filter 2 appears a more appropriate practical design.

Investigate filter 2 in EE

[JVET-Y0087](https://jvet-experts.org/doc_end_user/current_document.php?id=11281) AHG11: An Improved CNN-based Super Resolution Method [S. Peng, D. Jiang, J. Lin, C. Fang, X. Zhang (Dahua)]

This contribution presents a convolutional neural network-based super-resolution used as upsampling filter for Y component. Before entering the network, the luma block is downsampled by RPR and then encoded. The proposed method is developed from the previous contribution JVET-X0097. In this contribution, the residual block has more convolutional layers and the input adds side information in this network.

Compared with VTM-11.0-NNVC, the experimental results show 9.16% BD-rate gain and 0.52% BD-rate loss on average for luma, under AI and RA configuration, respectively.



1. Architecture of the proposed CNN model



1. Architecture of the designed ARB
2. Results of the proposed method under AI configuration

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | **AI** | | | | |
|  |  | **BD-rate Over VTM-11.0\_NNVC** | | | | |
|  |  | Y | U | V | EncT | DecT |
| Class A1 | Tango2 | -10.82% | 1.32% | 7.95% | 148% | 33% |
| FoodMarket4 | -5.94% | 5.32% | 8.23% | 137% | 33% |
| Campfire | -12.66% | 100.00% | 32.38% | 78% | 30% |
| Class A2 | CatRobot | -8.69% | 21.72% | 28.89% | 92% | 34% |
| DaylightRoad2 | -2.50% | -0.07% | 2.55% | 87% | 31% |
| ParkRunning3 | -14.34% | 125.44% | 48.30% | 64% | 33% |
| **Overall** | | -9.16% | 42.29% | 21.38% | 101% | 32% |

1. Results of the proposed method under RA configuration

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | **RA** | | | | |
|  |  | **BD-rate Over VTM-11.0\_NNVC** | | | | |
|  |  | Y | U | V | EncT | DecT |
| Class A1 | Tango2 | -4.51% | -1.99% | 7.67% | 81% | 31% |
| FoodMarket4 | -3.11% | 3.27% | 4.44% | 91% | 30% |
| Campfire | -4.56% | 97.41% | 27.98% | 55% | 33% |
| Class A2 | CatRobot | 6.31% | 36.81% | 52.42% | 74% | 30% |
| DaylightRoad2 | 16.37% | 6.39% | 12.32% | 73% | 31% |
| ParkRunning3 | -7.38% | 265.83% | 87.07% | 52% | 29% |
| **Overall** | | 0.52% | 67.95% | 31.98% | 71% | 31% |

Why is the performance so different in AI and RA?

2 Models for I slice and B slice

Adaptation would be beneficial

Further study recommended.

[JVET-Y0098](https://jvet-experts.org/doc_end_user/current_document.php?id=11292) EE1-related: Combination of VVC deblocking and NN loop filtering [K. Andersson, J. Ström, D. Liu, R. Sjöberg (Ericsson)]

NN loop filtering as in EE1-1.6 has shown to be capable of promising bitrate reductions while also mostly producing good visual quality. One concern is that when NN loop filter is not selected, no deblocking filtering at all is performed.

This contribution enables the use of VVC deblocking in combination with NN loop filtering. The NN loop filtering is from EE1-1.6 but with some notable differences. Instead of signalling which QP to use for the model from a fixed set of QPs, only one QP is used. Also, the maximum block size (excluding border extension for input) for luma is restricted to 128x128 for inter pictures for resolutions below 4K. The output is a convex combination of the VVC deblocked samples and the NN-processed samples. Since both VVC deblocking and NN loop filtering provide a deblocking effect, the output samples will always have seen some form of deblocking. The VVC deblocking is performed in RDO and a deblocking beta offset of -2 is also used.

The BDR effect and the impact on encoding and decoding time of the proposal compared to the EE1-1.6 is as follows:

AI: 0.12% Y, -0.79% U, -1.05% V, EncTime: 105%, DecTime: 96%

RA: -0.27% Y, -0.66% U, -0.31% V, EncTime: 76%, DecTime: 90%

LDB: -0.69% Y, -2.13% U, -1.60% V, EncTime: 71% DecTime: 84%

A notable encoder speedup can be seen for RA and LDB due to omitting multiple runs of the NN model with different QPs as in EE1-1.6. It is also asserted, by visual inspection, that the proposal makes sure that sufficient deblocking always is applied.

In v2 we also include results compared to the NNVC anchor:

AI: -7,27%Y, -19,80%U, -20,12%V, EncTime: 158%, DecTime: 16345%

RA: -10,04%Y, -21,63%U, -21,39%V, EncTime: 132%, DecTime: 28001%

LDB: -9,15%Y, -15,92%U, -16,03%V, EncTime: 139% DecTime: 25044%

It is reported that without using deblocking, there are some cases of blocking artifacts observed in particular in class E, happening when the NN filter is switched off.

Operation in parallel, and then a weighted combination of NN and deblocking output is used. Weight is optimized by MSE minimization, and signalled at slice level.

Combination to be considered in EE

### Other NN technology related contributions (10)

Contributions in this area were discussed in session 6 at 1645–1720 UTC on Thursday 13 Jan. 2022 and in session 10 at 1530–1730 UTC on Friday 14 Jan. 2022 (chaired by JRO).

[JVET-Y0046](https://jvet-experts.org/doc_end_user/current_document.php?id=11238) AHG11: 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, which indicate ALF enabled/disable, are introduced for luma and chroma components, respectively. It is reported that, with JVET-X0065 as anchor and test platform, the overall performance are as follows:

* AI: 0.03%/-0.96%/-1.04% with 100%EncT/100%DecT,
* RA: 0.05%/-0.30%/-0.41% with 100%EncT/100%DecT.

Interesting study showing that ALF is less useful (and better disabled) for chroma when NN filter is used

It is pointed out that the situation might change if some of the chroma gain would be shifted to luma, as currently the chroma quality may be too high to make ALF efficient. Would be useful to test with larger chroma QP offset.

It is asked whether it might be better to use ALF when NN filter is disabled locally (which is impossible when switching is done at slice level).

Further investigation in EE

[JVET-Y0195](https://jvet-experts.org/doc_end_user/current_document.php?id=11402) Cross-check of JVET-Y0046 (AHG11: ALF improvement for NNVC) [C. Lin (Bytedance)] [late]

[JVET-Y0051](https://jvet-experts.org/doc_end_user/current_document.php?id=11245) AHG11: Deep omnidirectional video compression in YUV domain [Q. Qin, C. Jung (Xidian University), D. Zou, M. Li (OPPO)]

This contribution presents deep omnidirectional video compression in the YUV domain, named DOVC-YUV. DOVC-YUV is an extension of the original DOVC framework to the YUV domain, which further saves bitrate using sub-sampled YUV 4:2:0 format. All modules of DOVC-YUV are jointly optimized by weighted rate distortion to minimize the geometric distortion of the sphere-to-plane projection. Compared with HM-16.16 under RA configuration, DOVC-YUV achieves average BD-rate reductions of {34.10% (Y), 35.87% (U), and 38.09% (V)} and average BD-WSPSNR (end-to-end) gains of {1.15dB (Y), 0.79dB (U), and 0.83dB (V)} on 360-degree video datasets provided by JVET. Compared with VTM-12.0 under RA configuration, DOVC-YUV achieves slightly worse performance with average BD-rate reductions of {-3.83% (Y), -6.54% (U), and -5.67% (V)} and average BD-WSPSNR (end-to-end) gains of {-0.10dB (Y), -0.26dB (U), and -0.30dB (V)} on the same dataset. Furthermore, DOVC-YUV takes advantage of GPU parallel processing and thus the average encoding time of DOVC-YUV is only 0.0425 times that of HM-16.16 and 0.0118 times that of VTM-12.0.

Different from previous proposal JVET-X0043, compression is performed in YUV domain (not RGB)

BPG used for key frames, whereas intermediate frames are coded by end-to-end-approach

Distance between key frames (intra period) is 9.

Which percentage of data rate is used for the BPG-coded keyframes.

Could this be applied to regular 2D video (not 360)? This would be an interesting test case for a future EE. Further study to investigate with 2D video highly recommended.

If BPG would be replaced by VVC intra coding, this would also likely improve the performance. Furthcr study in this direction is also recommended.

For further study, VTM should be used as anchor.

[JVET-Y0052](https://jvet-experts.org/doc_end_user/current_document.php?id=11246) AHG11: CNN post-processing filter based on depthwise separable convolution and attention mechanism [H. Zhang, C. Jung (Xidian University), D. Zou, M. Li (OPPO)]

This contribution proposes weakly connected dense attention neural network for compression artifact removal, called WCDANN. WCDANN is a post-processing filter based on convolutional neural networks (CNNs) that does not need any change of codecs. WCDANN consists of several weakly connected dense attention blocks (WCDABs), which is based on residual learning by taking the compressed video after codecs as the input. We use depthwise separable convolution as the basic convolution unit for WCDANN to generate a lightweight model. Moreover, we introduce attention mechanism into the proposed filter to emphasize important features. Compared with VTM-11.0-NNVC, WCDANN achieves average 4.78%, 4.06% and 4.33% BD-rate reductions for Y channel on C and D classes, under AI, RA and LDP configurations, respectively.

Two models for lower/higher QP ranges

Training uses L1 loss first, then switches to L2

Only applied on Y

Depth-wise separable convolution is the main difference from other approaches – has it been studied which is its benefit? Not specifically in terms of performance, but amount of computations is lower.

In previous EE, no post filter has been investigated. A previous post filter approach (W0057) had slightly worse performance.

It is mentioned that interest in post filters was somehow decreased because it performs worse than loop filter. In particular, NN based filters are better positioned before ALF.

Investigate in EE

It is suggested to also compare against the performance when operating the filter as a loop filter (without retraining).

[JVET-Y0059](https://jvet-experts.org/doc_end_user/current_document.php?id=11253) AHG11: Content-adaptive post-processing filter [M. Santamaria, J. Lainema, F. Cricri, R. G. Youvalari, H. Zhang, G. Rangu, H. R. Tavakoli, H. Afrabandpey, M. M. Hannuksela (Nokia)]

This contribution presents a content-adaptive post-processing filter based on Convolutional Neural Networks (CNNs). The filter is trained offline on general video sequences and the content-adaptation is achieved by means of finetuning the filter on the test video sequence. This contribution is developed on top of two previous contributions. Firstly, JVET-X0111 where only the bias terms of the CNN are finetuned and the weight-update is compressed with the MPEG Neural Network compression and Representation (NNR) standard. Secondly, the CNN in-loop filter presented in JVET-W0131 is used as post-processing filter. The proposed approach was implemented on top of VVC Test Model 11.0 with Neural Network for Video Coding (NNVC) technology 1.0. It is reported that the overall BD-rate in Random Access (RA) configuration is -4,56% (Y), -10,36% (U), -6,54% (V). It is also reported that content adaptation improves the overall BD-rate gains in RA configuration by -2,36% (Y), -7,77% (U), -4,66% (V).

Filter from JVET-W0131 was not re-trained, also same models and model switching criteria are used.

Encoding time increased to 126%, decoding time roughly 4000% (CPU times).

Weight update computation (for updating subset of bias terms) is not included in encoding time above, is done on GPU. Done for each test sequence and QP, run times per all test cases of each class are included in table 1 of the doc. 10 iterations are used currently in the training.

No interest in joining EE. Further study.

[JVET-Y0081](https://jvet-experts.org/doc_end_user/current_document.php?id=11275) AHG11: Transformer based in-loop filtering [T. Ouyang, H. Wang, H. Zhu, Z. Chen (Wuhan University)] [late]

This contribution reports the experimental results of transformer based in-loop filtering, which employs attention modules for image restoration instead of the convolutional neural network as backbone. The performance is evaluated using PSNR and MS-SSIM based BD-rate reduction in VTM-11.0 + new MCTF, which reports an average gain of {-6.77%, -14.62%, -14.97%} on PSNR and {-6.88%, -17.24%, -17.50%} on MS-SSIM under all intra configuration.

Only AI tested so far

Input taken before DBF, output switched with SAO output, picture level and CTU level switching

Swin Transformer was originally developed for super resolution, and same loss function was used as in the original paper, which is MSE plus regularization. Further study on modifying the loss function, e.g. using plain MSE?

Number of parameters somewhat higher than in other approaches, due to large number of fully-connected layers.

It is also suggested to investigate more training iterations.

No interest in joining EE. Further study.

[JVET-Y0086](https://jvet-experts.org/doc_end_user/current_document.php?id=11280) AHG11: A Unet-Based Deep In-Loop Filter [X. Zhang, D. Jiang, J. Lin, C. Fang, S. Peng (Dahua)]

This contribution presents a convolutional neural network-based in-loop filtering method with QP based models. In this contribution, a Unet-based neural network is utilized. The luma and chroma components both have 5 different CNN models, corresponding to 5 QPs. The CNN model is selected according to colour components and QP.

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.

Better performance from filters investigated in EE, with even less complexity. One reason might be that UNet is not using attention mechanisms.

Would be interesting to see how the performance in RA would be.

Further study recommended.

[JVET-Y0090](https://jvet-experts.org/doc_end_user/current_document.php?id=11284) 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 square coding blocks with neural network. The proposed method is mandatorily performed on all inter-coded square coding blocks. Experimental results demonstrate that, compared with VTM-11.0-nnvc-1.0, the proposed method achieves 1.23%, 1.20%, and 2.95% BD-rate reductions for Y component on Class B, C, and D under RA configuration, respectively.

25 different model for different block sizes, only square blocks so far.

Network targets reducing the MSE. Used for both uni and bi prediction in same way in the very end, after BCW, DMVR

Has it been tried to make it switchable? Yes, but performance was worse.

Could it be applied to intra? Has not been investigated.

Has it been investigated together with loop filter? No.

Further study recommended.

[JVET-Y0096](https://jvet-experts.org/doc_end_user/current_document.php?id=11290) AHG11: NN-based Reference Frame Interpolation for VVC Hierarchical Coding Structure [Z. Liu, X. Xu, S. Liu (Tencent), Y. Guo, Z. Chen (Wuhan Univ.)]

This contribution presents a deep-neural-network-based reference frame generation method for VVC hierarchical coding structure. During the encoding and decoding process, the frame interpolation network receives two reconstructed frames from the reference picture list as inputs and generates a new frame. Then a cascade filter network will process this frame consequently. The generated frame will be put into the last position of the reference list for the current picture. It is reported that this method can bring 2.01%/7.01%/6.44% bitrate savings on Y/U/V component respectively, following the JVET-NNVC CTC RA configuration.

Have other positions in the RP list been used? This could possibly improve the performance. Was not investigated so far

Encoding time increased approx. 2x, due to generation of the new reference picture and the additional search.

Only applied in two highest temporal layers.

Further study recommended.

[JVET-Y0110](https://jvet-experts.org/doc_end_user/current_document.php?id=11304) AHG11: Small Ad-hoc Deep-Learning Library (SADL) update [F. Galpin, F. Mom, T. Dumas, P. Bordes, P. Nikitin, E. François (InterDigital)]

see section 4.8

[JVET-Y0111](https://jvet-experts.org/doc_end_user/current_document.php?id=11305) AHG11: Hybrid Conventional/Deep-learning-based image coding [F. Galpin, T. Dumas, P. Bordes, F. Racapé, E. François (InterDigital), Y. Li, Kai Zhang, Li Zhang (Bytedance), H. Wang, K. Reuzé, A.M. Kotra, M. Karczewicz (Qualcomm)]

This contribution demonstrates the compression performance of a hybrid conventional/deep learning-based image coding approach, compared to fully conventional and end-to-end trained methods. Neural Network based Intra Prediction (as described in JVET-X0118) and Neural Network based Post-filtering (as described in JVET-X0066) are implemented on top of ECM-3.1. The compression performance is evaluated on the Kodak image dataset and compared to other state-of-the-art approaches. Average PSNR BD-rate variations of -11.7%, -27.5 and -26.6% on Y, U and V components, respectively, are reported on the Kodak dataset compared to the VTM-15.0 anchor.

Contribution for information.

### NN related HLS signalling (4)

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

[JVET-Y0073](https://jvet-experts.org/doc_end_user/current_document.php?id=11267) AHG9: Colour component description for post-filter purpose SEI message [T. Chujoh, Y. Yasugi, K. Takada, T. Ikai (Sharp)]

[JVET-Y0074](https://jvet-experts.org/doc_end_user/current_document.php?id=11268) AHG9: Data conversion description for NNR post-filter SEI message [Y. Yasugi, T. Chujoh, K. Takada, T. Ikai (Sharp)]

[JVET-Y0075](https://jvet-experts.org/doc_end_user/current_document.php?id=11269) AHG9: Complexity description for NNR post-filter SEI message [K. Takada, Y. Yasugi, T. Chujoh, T. Ikai (Sharp)]

[JVET-Y0115](https://jvet-experts.org/doc_end_user/current_document.php?id=11309) AHG9: On post-filter SEI [M. M. Hannuksela, M. Santamaria, F. Cricri, E. B. Aksu, H. R. Tavakoli (Nokia)]

## AHG12: Enhanced compression beyond VVC capability (56)

### Summary and BoG reports

[JVET-Y0024](https://jvet-experts.org/doc_end_user/current_document.php?id=11419) 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]

Presented in sessions 3 and 4 Wed 12 Jan 2100-2305 and 2330-0120+1

This document provides a summary report of Exploration Experiment on Enhanced Compression beyond VVC capability. The tests are categorized as partitioning, inter prediction, and in-loop filtering tests.

The software basis for this EE is ECM-3.1, released at [https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tags/ECM-3.1](https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tags/ECM-3.1.). ECM-3.1 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-x-ee2/ECM/-/branches>.

Test and cross-check results can be found in the input JVET contributions and are located at https://vcgit.hhi.fraunhofer.de/ecm/jvet-x-ee2/simulation-results/.

1. **List of tests**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Tests** | **Tester** | **Cross-checker** |
| **1 Partitioning** | | | |
| 1.1a | ABT | ByteDance  [Kai Zhang](mailto:zhangkai.video@bytedance.com)  InterDigital  [Fabrice Urban](mailto:fabrice.leleannec@interdigital.com)  [JVET-Y0150](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0150-v1.zip) | Xiaomi  Fabrice Le Léannec  [JVET-Y0180](https://jvet-experts.org/doc_end_user/current_document.php?id=11387) |
| 1.1b | UQT | ByteDance  [Kai Zhang](mailto:zhangkai.video@bytedance.com)  InterDigital  [Fabrice Urban](mailto:fabrice.leleannec@interdigital.com)  [JVET-Y0150](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0150-v1.zip) | Xiaomi  Fabrice Le Léannec  [JVET-Y0180](https://jvet-experts.org/doc_end_user/current_document.php?id=11387) |
| 1.1c | ABT + UQT | ByteDance  [Kai Zhang](mailto:zhangkai.video@bytedance.com)  InterDigital  [Fabrice Urban](mailto:fabrice.leleannec@interdigital.com)  [JVET-Y0150](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0150-v1.zip) | Huawei  Johannes Sauer |
| **2 Intra prediction** | | | |
| 2.1a | Extended MRL candidate list | Qualcomm  [Keming Cao](mailto:kemicao@qti.qualcomm.com)  [JVET-Y0116](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0116-v1.zip) | OPPO  [Kazushi Sato](mailto:kazushi.sato@oppo.com)  [JVET-Y0171](https://jvet-experts.org/doc_end_user/current_document.php?id=11378)  Ofinno  [Vasily Rufitskiy](mailto:vrufitskiy@ofinno.com)  [JVET-Y0204](https://jvet-experts.org/doc_end_user/current_document.php?id=11411) |
| 2.1b  2.1c | Extended MRL with the reduced number candidates | Qualcomm  [Keming Cao](mailto:kemicao@qti.qualcomm.com)  [JVET-Y0116](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0116-v1.zip) | Ofinno  [Vasily Rufitskiy](mailto:vrufitskiy@ofinno.com)  [JVET-Y0204](https://jvet-experts.org/doc_end_user/current_document.php?id=11411) |
| **3 Inter prediction** | | | |
| 3.1a | GPM with intra and inter prediction | KDDI  [Yoshitaka Kidani](mailto:yo-kidani@kddi.com)  LGE  [Hyeongmun Jang](mailto:hm.jang@lge.com)  Bytedance  [Zhipin Deng](mailto:zhipin.deng@bytedance.com)  [JVET-Y0065](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0065-v1.zip) | OPPO  [Kazushi Sato](mailto:kazushi.sato@oppo.com)  JVET-Y0182  Alibaba  [Ru-Ling Liao](mailto:ruling.lrl@alibaba-inc.com) |
| 3.1b | 3.1a + DIMD and neighboring mode based IPM derivation + combination of GPM-intra and GPM-MMVD | KDDI  [Yoshitaka Kidani](mailto:yo-kidani@kddi.com)  LGE  [Hyeongmun Jang](mailto:hm.jang@lge.com)  Bytedance  [Zhipin Deng](mailto:zhipin.deng@bytedance.com)  [JVET-Y0065](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0065-v1.zip) | Alibaba  [Ru-Ling Liao](mailto:ruling.lrl@alibaba-inc.com) |
| 3.1c | 3.1b + TIMD-based IPM derivation | KDDI  [Yoshitaka Kidani](mailto:yo-kidani@kddi.com)  LGE  [Hyeongmun Jang](mailto:hm.jang@lge.com)  Bytedance  [Zhipin Deng](mailto:zhipin.deng@bytedance.com)  [JVET-Y0065](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0065-v1.zip) | Alibaba  [Ru-Ling Liao](mailto:ruling.lrl@alibaba-inc.com) |
| 3.1d | 3.1c + combination of GPM-intra and GPM-TM | KDDI  [Yoshitaka Kidani](mailto:yo-kidani@kddi.com)  LGE  [Hyeongmun Jang](mailto:hm.jang@lge.com)  Bytedance  [Zhipin Deng](mailto:zhipin.deng@bytedance.com)  [JVET-Y0065](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0065-v1.zip) | Alibaba  [Ru-Ling Liao](mailto:ruling.lrl@alibaba-inc.com) |
| 3.2 | Modified pairwise candidates | Canon  [Guillaume Laroche](mailto:guillaume.laroche@crf.canon.fr)  [JVET-Y0100](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0100-v1.zip) | Bytedance  Lei Zhao |
| 3.3a | TMVP improvement method 1 | Alibaba  [Ru-Ling Liao](mailto:ruling.lrl@alibaba-inc.com)  [JVET-Y0093](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0093-v1.zip) | OPPO  [Zhihuang](mailto:xiezhihuang@oppo.com) Xie  JVET-Y0183 |
| 3.3b | TMVP improvement method 1 + method 2 | Alibaba  [Ru-Ling Liao](mailto:ruling.lrl@alibaba-inc.com)  [JVET-Y0093](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0093-v1.zip) | OPPO  [Zhihuang](mailto:xiezhihuang@oppo.com) Xie  JVET-Y0183 |
| 3.4 | MV candidate type-based ARMC | Qualcomm  [Yao-Jen Chang](mailto:yjchang@qti.qualcomm.com)  Alibaba  [Ru-Ling Liao](mailto:ruling.lrl@alibaba-inc.com)  [JVET-Y0134](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0134-v1.zip) | Kwai  [Xiaoyu Xiu](mailto:xiaoyuxiu@kwai.com) |
| 3.5a | Template matching based merge candidate list construction | Bytedance  Lei Zhao  [JVET-Y0134](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0134-v1.zip) | Canon  [Guillaume Laroche](mailto:guillaume.laroche@crf.canon.fr) |
| 3.5b | Template matching based merge candidate list construction (Test 3.5a) + Modified pairwise candidates (Test 3.2) | Bytedance  Lei Zhao  Canon  [Guillaume Laroche](mailto:guillaume.laroche@crf.canon.fr)  [JVET-Y0134](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0134-v1.zip) |  |
| 3.6a | MV candidate type-based ARMC (Test 3.4) + TMVP improvement from Test 3.3 | Qualcomm  [Yao-Jen Chang](mailto:yjchang@qti.qualcomm.com)  Alibaba  [Ru-Ling Liao](mailto:ruling.lrl@alibaba-inc.com)  [JVET-Y0134](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0134-v1.zip) | Kwai  [Xiaoyu Xiu](mailto:xiaoyuxiu@kwai.com) |
| 3.6b | MV candidate type-based ARMC (Test 3.4) + modified pairwise candidates (Test 3.2) | Qualcomm  [Yao-Jen Chang](mailto:yjchang@qti.qualcomm.com)  Canon  [Guillaume Laroche](mailto:guillaume.laroche@crf.canon.fr) | withdrawn |
| 3.6c | MV candidate type-based ARMC (Test 3.4) + MV candidates from Test 3.5 | Qualcomm  [Yao-Jen Chang](mailto:yjchang@qti.qualcomm.com)  Bytedance  Lei Zhao | withdrawn |
| 3.6d | MV candidate type-based ARMC (Test 3.4) + TMVP improvement from Test 3.3 + Modified pairwise candidates from Test 3.2 + MV candidates from Test 3.5  Test 3.2 part is withdrawn from the combination | Qualcomm  [Yao-Jen Chang](mailto:yjchang@qti.qualcomm.com)  Canon  [Guillaume Laroche](mailto:guillaume.laroche@crf.canon.fr)  Bytedance  Lei Zhao  Alibaba  [Ru-Ling Liao](mailto:ruling.lrl@alibaba-inc.com)  [JVET-Y0134](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0134-v1.zip) | InterDigital  [Fabrice Urban](mailto:Fabrice%20Urban%20%3cFabrice.Urban@InterDigital.com%3e) |
| 3.7 | Increased number of TM merge candidates | Qualcomm  [Yao-Jen Chang](mailto:yjchang@qti.qualcomm.com)  [JVET-Y0132](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0132-v1.zip) | Alibaba  [Ru-Ling Liao](mailto:ruling.lrl@alibaba-inc.com) |
| 3.8 | Alternative template matching | Qualcomm  [Han Huang](mailto:hanhuang@qti.qualcomm.com) | withdrawn |
| 3.9a | TM based reordering for MMVD | Bytedance  Mehdi Salehifar  [JVET-Y0067](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0067-v1.zip) |  |
| 3.9b | TM based reordering for MMVD and affine MMVD | Bytedance  Mehdi Salehifar  [JVET-Y0067](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0067-v1.zip) | Kwai  [Xiaoyu Xiu](mailto:xiaoyuxiu@kwai.com) |
| 3.9c | TM based reordering for MMVD (Test 3.9a) + MVD sign prediction (Test 3.10) | Bytedance  Mehdi Salehifar  Qualcomm  Yan Zhang  [JVET-Y0067](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0067-v1.zip) |  |
| 3.9d | TM based reordering for MMVD and affine MMVD (Test 3.9b) + MVD sign prediction (Test 3.10) | Bytedance  Mehdi Salehifar  Qualcomm  Yan Zhang  [JVET-Y0067](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0067-v1.zip) | Kwai  [Xiaoyu Xiu](mailto:xiaoyuxiu@kwai.com) |
| 3.10 | MVD sign prediction | Qualcomm  Yan Zhang  [JVET-Y0067](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0067-v1.zip) | Kwai  Han Gao  Ofinno  [Vasily Rufitskiy](mailto:vrufitskiy@ofinno.com) |
| 3.11 | Non-adjacent spatial neighbor for affine merge mode | Kwai  [Wei Chen](mailto:chenwei06@kwai.com)  [JVET-Y0153](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0153-v2.zip) | Qualcomm  Yan Zhang  Bytedance  Kai Zhang  [JVET-Y0207](https://jvet-experts.org/doc_end_user/current_document.php?id=11414) |
| 3.12a | History-parameter-based affine model inheritance | Bytedance  Kai Zhang  [JVET-Y0145](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0145-v1.zip) | Alibaba  [Jie](mailto:m.sarwer@alibaba-inc.com) Chen |
| 3.12b | History-parameter-based affine model inheritance (Test 3.12a) + non-adjacent spatial neighbors for affine merge mode (Test 3.11) | Bytedance  Kai Zhang  Kwai  Wei Chen  [JVET-Y0146](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0146-v1.zip) | Alibaba  [Jie](mailto:m.sarwer@alibaba-inc.com) Chen |
| 3.12c | 3.12b with TM disabled | Bytedance  Kai Zhang  Kwai  Wei Chen  [JVET-Y0146](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0146-v1.zip) |  |
| 3.13 | Modifications of IBC merge/AMVP list construction | Bytedance  Na Zhang  [JVET-Y0058](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0058-v1.zip) | Ofinno  [Damian Ruiz Coll](mailto:druizcoll@ofinno.com)  [JVET-Y0206](https://jvet-experts.org/doc_end_user/current_document.php?id=11413) |
| **4 Transform and coefficients coding** | | | |
| 4.1a | QIdx based signs selection | Alibaba  [Jie](mailto:m.sarwer@alibaba-inc.com) Chen  [JVET-Y0094](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0094-v1.zip) | OPPO  [Luhang Xu](mailto:xuluhang@oppo.com)  JVET-Y0185 |
| 4.1b | Extension of sign prediction area | Alibaba  [Jie](mailto:m.sarwer@alibaba-inc.com) Chen | withdrawn |
| 4.1c | Qidx based signs selection + extension of sign prediction area | Alibaba  [Jie](mailto:m.sarwer@alibaba-inc.com) Chen  [JVET-Y0094](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0094-v1.zip) | OPPO  [Luhang Xu](mailto:xuluhang@oppo.com)  JVET-Y0185 |
| 4.2a | Sign prediction with adaptive coefficient selection | Kwai  [Xiaoyu Xiu](mailto:xiaoyuxiu@kwai.com)  [JVET-Y0137](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0137-v1.zip) | OPPO  [Luhang Xu](mailto:xuluhang@oppo.com)  JVET-Y0184 |
| 4.2b | Sign prediction with adaptive coefficient selection + extension of sign prediction to LFNST | Kwai  [Xiaoyu Xiu](mailto:xiaoyuxiu@kwai.com)  [JVET-Y0137](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0137-v1.zip) | OPPO  [Luhang Xu](mailto:xuluhang@oppo.com)  JVET-Y0184 |
| 4.3a | Test 4.1c + extension of sign prediction to LFNST (Test 4.2b) | Alibaba  [Jie](mailto:m.sarwer@alibaba-inc.com) Chen  Kwai  Xiaoyu Xiu  [JVET-Y0138](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0138-v1.zip) | Bytedance  Yang Wang |
| 4.3b | Extension of sign prediction area (4.1b) + Sign prediction with adaptive coefficient selection with extension of sign prediction to LFNST (4.2b) | Alibaba  [Jie](mailto:m.sarwer@alibaba-inc.com) Chen  Kwai  Xiaoyu Xiu  [JVET-Y0138](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0138-v1.zip) | Tencent  L.-F. Chen |
| 4.4a | Adaptive intra MTS | Qualcomm  [Bappaditya Ray](mailto:bray@qti.qualcomm.com)  [JVET-Y0142](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0142-v1.zip) | Sharp  [Tomonori Hashimoto](mailto:tomonori.hashimoto@sharp.co.jp) |
| 4.4b | Adaptive intra MTS – encoder only (non-normative) | Qualcomm  [Bappaditya Ray](mailto:bray@qti.qualcomm.com)  [JVET-Y0142](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0142-v1.zip) | Sharp  [Tomonori Hashimoto](mailto:tomonori.hashimoto@sharp.co.jp) |
| 4.4c | Adaptive intra MTS – with parameter signalling | Qualcomm  [Bappaditya Ray](mailto:bray@qti.qualcomm.com)  [JVET-Y0142](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0142-v1.zip) | Sharp  [Tomonori Hashimoto](mailto:tomonori.hashimoto@sharp.co.jp) |
| 4.4d | Intra MTS without adaptivity (fixed number of candidates) | Qualcomm  [Bappaditya Ray](mailto:bray@qti.qualcomm.com)  [JVET-Y0142](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0142-v1.zip) | Sharp  [Tomonori Hashimoto](mailto:tomonori.hashimoto@sharp.co.jp) |
| **5 In-loop filtering** | | | |
| 5.1 | Edge-based classifier for CCSAO | Qualcomm  [Anand Meher Kotra](mailto:akotra@qti.qualcomm.com)  Kwai  [Che-Wei Kuo](mailto:cheweikuo@kwai.com)  [JVET-Y0106](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0106-v1.zip) | Alibaba Xinwei Li  [JVET-Y0196](https://jvet-experts.org/doc_end_user/current_document.php?id=11403) |
| 5.2a | Adaptive filter shape selection for ALF | Bytedance  [Wenbin Yin](mailto:yinwenbin.hit@bytedance.com)  Qualcomm  [Nan Hu](mailto:nanh@qti.qualcomm.com)  Alibaba  [Mohammed Golam Sarwer](mailto:m.sarwer@alibaba-inc.com)  [JVET-Y0147](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0147-v1.zip) | Ericsson  Jacob Ström |
| 5.2b | Adaptive filter shape selection for ALF with reduced filter length | Bytedance  [Wenbin Yin](mailto:yinwenbin.hit@bytedance.com)  Qualcomm  [Nan Hu](mailto:nanh@qti.qualcomm.com)  Alibaba  [Mohammed Golam Sarwer](mailto:m.sarwer@alibaba-inc.com)  [JVET-Y0147](https://jvet-experts.org/doc_end_user/documents/25_Teleconference/wg11/JVET-Y0147-v1.zip) | Ericsson  Jacob Ström |

Test 1: Partitioning

**Test 1.1a: Asymmetric binary tree (ABT)**

Four new asymmetric binary tree splitting modes are added to the multi-type tree structure of VVC shown on Figure 1.

|  |  |
| --- | --- |
|  |  |
| ABT-H1 | ABT-H2 |
|  |  |
| ABT-V1 | ABT-V2 |

**Figure 1. ABT partitions**

CU can be divided into 2 sub-CUs with 1/4 and 3/4 sizes. 6-point, 12-point, 24-point, and 48-point transforms/inverse-transforms are introduced. Intra prediction is also modified to accommodate non-power-2 block sizes.

**Test 1.1b: Unsymmetric quad tree (UQT)**

Four types of unsymmetric quad tree partitions with 1/8, 1/2, 1/4, and 1/8 sizes as shown on Figure 2 are tested.

|  |  |
| --- | --- |
|  |  |
| UQT-H1 | UQT-H2 |
|  |  |
| UQT-V1 | UQT-V2 |

**Figure 2. UQT partitions**

**Test 1.1c: Asymmetric binary tree (ABT) and Unsymmetric quad tree (UQT)**

This test is a combination of the tests 1.1a and 1.1b

Results (AI, RA, LB; each column Y/U/V/enc/dec)



Similar tradeoff as in last EE, which was not asserted as beneficial. Encoder-only method that was adopted last time seems sufficient at this point.

Compared to benefit of improved partitioning introduced in HEVC and VVC, the additional gain appears marginal. One expert points out that at least UQT would be very simple to implement at the decoder. It is also pointed out that some more evidence about the benefit of modified partitioning is coming from EE related contributions.

Test 2: Intra prediction

**Test 2.1: Extended MRL candidate list**

MRL candidate list is extended from lines {1, 2} to {1, 3, 5, 7, 12} in test 2.1a, to {1, 3, 5, 7} in test 2.1b or {1, 3, 5} in test 2.1c. For template-based intra mode derivation (TIMD), lines {1, 3} are used.

Chart

Description automatically generated

Figure 3. Extended MRL candidate list

**Table 1 Usage frequency of each line in MRL candidate list {1, 3, 5, 7, 12}**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | line 1 | line 3 | line 5 | line 7 | line 12 |
| Class A1 | 36% | 15% | 15% | 13% | 21% |
| Class A2 | 54% | 19% | 11% | 7% | 9% |
| Class B | 43% | 21% | 16% | 10% | 11% |
| Class C | 53% | 19% | 11% | 8% | 9% |
| Class E | 46% | 21% | 12% | 8% | 12% |
| **Overall** | 46% | 19% | 14% | 9% | 12% |
| Class D | 45% | 19% | 13% | 9% | 13% |
| Class F | 57% | 18% | 10% | 7% | 9% |
| Class TGM | 38% | 17% | 20% | 10% | 15% |

Results (AI, RA; each column Y/U/V/enc/dec)



Small benefit, but also relatively small change. Compression benefit (AI) close to 0.1%. Generally, AI performance of ECM is significantly less compared to VTM than for RA (ECM2 to ECM3 improved by roughly 0.2%, most likely from filtering as no new intra tools were adopted).

No additional line buffers relative to current ECM.

.

It was later agreed (Wed. 19 Jan. in session20) to adopt 2.1a which gives best gain compared to the others, and is complexity-wise practically equivalent

Decision: Adopt JVET-Y0116 version 2.1a (extension of MRL list to 5 lines 1,3,5,7,12)

Test 3: Inter prediction

**Test 3.1: GPM with intra and inter prediction**

In the tests, the prediction is generated by weighting inter predicted samples and intra predicted samples for each GPM-separated region. Inter part is derived by using existed prediction in GPM, while intra part is derived by intra prediction mode indicated by an index into the intra modes candidate list.

The intra candidate list has size of 3, which has the parallel to GPM edge angular mode, the perpendicular angular mode, and planar mode as shown **Figure 4**.

グラフ, ダイアグラム

中程度の精度で自動的に生成された説明

Figure 4 GPM with intra and inter prediction.

This base mode is tested in Test 3.1a.

In the Test 3.1b, the following two modifications are introduced on top of Test 3.1a:

* DIMD and neighboring mode based IPM derivation are added after parallel mode in the intra mode list. Neighbor intra modes are derived from at most 5 locations but they are restricted based on the GPM angle as detailed in Table 2
* Combination of GPM-intra and GPM-MMVD

**Table 2 : The positions of available neighboring blocks for intra candidate derivation based on the angle of GPM block boundary. A and L denotes the above and left side of the prediction block.**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Angle of GPM | 0 | 2 | 3 | 4 | 5 | 8 | 11 | 12 | 13 | 14 |
| 1st partition | A | A | A | A | L+A | L+A | L+A | L+A | A | A |
| 2nd partition | L+A | L+A | L+A | L | L | L | L | L+A | L+A | L+A |
| Partition angle | 16 | 18 | 19 | 20 | 21 | 24 | 27 | 28 | 29 | 30 |
| 1st partition | A | A | A | A | L+A | L+A | L+A | L+A | A | A |
| 2nd partition | L+A | L+A | L+A | L | L | L | L | L+A | L+A | L+A |

In the Test 3.1c, on top of Test 3.1b template-based intra mode derivation (TIMD) is additionally utilized for intra mode candidates, the order of the modes is parallel mode, TIMD candidates, DIMD candidates, and neighboring intra modes. The list size is kept as 3.

In the Test 3.1d, template matching for GPM is added on top of Test 3.1c.

**Test 3.2: Modified pairwise candidates**

In the test, the pairwise candidate is removed from the regular merge and BM merge candidate derivation. It is added, when it is not a duplicate, during the adaptive reordering of merge candidates with template matching (ARMC-TM) as depicted in Figure 5. The pairwise is built with the two first reordered candidates.

In addition, each merge candidate, in the non-reordered subgroup, is replaced by a pairwise between the first candidate and this candidate if the created pairwise is not a duplicate.

Text

Description automatically generated with medium confidence

Figure 5. Pairwise in the merge candidate reordering ARMC-TM and additional pairwise candidates after reordering.

**Test 3.3: TMVP improvement**

In Test 3.3a, TMVP candidate is conditionally added to the list according to the motion information of spatial neighboring blocks, which are bottom left, above right, and above left.

TMVP is derived from the bottom right position, if it is inter coded and if none of the spatial neighboring blocks has similar motion or, at least one of the spatial neighboring blocks has similar motion and the spatial neighboring blocks use TMVP.

If TMVP is not derived, the same derivation process is repeated for the center position.

In checking the similarity of the temporal motion, if the inter prediction direction of the two temporal motion are the same and the motion vector differences in horizontal and vertical direction are both smaller than 5, the two temporal motion are treated as similar to each other.

The target reference picture of the TMVP candidate can be selected from any one of reference picture in the list according to scaling factor. The scaling factor of each reference picture in the reference picture list is calculated, and the one whose scaling factor is the closest to 1 is used as the target reference picture of the TMVP.

Icon

Description automatically generated with medium confidence

Test 3.3b is implemented on top of Test 3.3a. When a CU is coded using template matching, AMVP-merge and regular merge modes and the current picture is a non-low delay picture, the TMVP candidate is refined using the template matching method. The template is used to decide whether the TMVP is L0-predicted, L1-predicted or bi-predicted, and is used to decide which one of the three scaling factors {,,} is used to perform temporal motion vector scaling. The tb is defined to be the POC difference between the reference picture of the current picture and the current picture and td is defined to be the POC difference between the reference picture of the co-located picture and the co-located picture.

**Tests 3.4: MV candidate type-based ARMC**

In this test, merge candidates of one single candidate type, e.g., TMVP or non-adjacent MVP (NA-MVP), are reordered based on the ARMC TM cost values. The reordered candidates are then added into the merge candidate list. The TMVP candidate type adds more TMVP candidates with more temporal positions and different inter prediction directions to perform the reordering and the selection. Moreover, NA-MVP candidate type is further extended with more spatially non-adjacent positions.

**Test 3.5: Template matching based merge candidate list construction**

In the test, NA-MVP, HMVP, and PAMVP candidates are put into the merge candidate list in an ascending order of template matching costs. Moreover, non-adjacent MVPs is further extended with more spatial and non-adjacent temporal positions.

**Test 3.6: Combination of the ARMC related methods**

The combination test results of Test 3.2, Test 3.3a, Test 3.4 and Test 3.5a are as follows:

Test 3.5b (Combination of Test 3.5a and Test 3.2): On top of 3.5a software, the PAMVP is removed from the merge list and derived during the reordering process of merge candidate list (ARMC). In addition, each merge candidate, in the non-reordered subgroup, is replaced by a pairwise between the first candidate and this candidate.

Test 3.6a (Combination of Test 3.4 and Test 3.3a): On top of 3.4 software, it adds the following method from Test 3.3a:

* The reference picture of the TMVP candidate is selected from any one of reference pictures in the reference picture list according to a pre-defined rule. The selected reference picture is the one whose scaling factor is the closest to 1.

Test 3.6d (Test 3.3a + Test 3.4 + Test 3.5a): On top of 3.5a software, it adds the following methods from Test 3.4 and Test 3.3a:

* Test 3.4: TMVP candidates reordering based on TM cost values.
* Test 3.3a: The reference picture of the TMVP candidate is selected from any one of reference pictures in the reference picture list according to a pre-defined rule. The selected reference picture is the one whose scaling factor is the closest to 1.

**Test 3.7: Increased number of TM merge candidates**

The number of TM merge candidates is increased from 4 to 7 by modifying the parameter TM\_MRG\_MAX\_NUM\_CANDS in ECM-3.1 software.

**Test 3.9: MMVD extension based on template matching cost**

In Test 3.9a, MMVD offsets are extended for MMVD and affine MMVD modes. Additional refinement positions along k×π/8 diagonal angles are added shown on Figure 6, thus increasing the number of directions from 4 to 16. Second, based on the SAD cost between the template (one row above and one column left to the current block) and its reference for each refinement position, all the possible MMVD refinement positions (16×6) for each base candidate are reordered. Finally, the top 1/8 refinement positions with the smallest template SAD costs are kept as available positions, consequently for MMVD index coding. The MMVD index is binarized by the rice code with the parameter equal to 2.



Figure 6. Additional directions along k×π/8 diagonal angles (red positions are used in the anchor).

In Test 3.9b, on top of the 3.9a, affine MMVD reordering is extended, in which additional refinement positions along k×π/4 diagonal angles are added. After reordering top 1/2 refinement positions with the smallest template SAD costs are kept.

In Test 3.9c (Test 3.9a + Test 3.10), where MMVD MVD sign prediction part of EE2 3.10 is replaced with MMVD extension of Test 3.9a.

In Test 3.9d (Test 3.9b + Test 3.10), where MMVD and affine MVD sign prediction parts of EE2 3.10 are replaced with MMVD and affine MMVD extensions of Test 3.9a.

**Test 3.10: MVD sign prediction**

In this method, possible MVD sign combinations are sorted according to the template matching cost and index corresponding to the true MVD sign is derived and context coded. At decoder side, the MVD signs are derived as following:

1. Parse the magnitude of MVD components
2. Parse context-coded MVD sign prediction index
3. Build MV candidates by creating combination between possible signs and absolute MVD value and add it to the MV predictor
4. Derive MVD sign prediction cost for each derived MV based on template matching cost and sort
5. Use MVD sign prediction index to pick the true MVD sign

MVD sign prediction is applied to inter AMVP, affine AMVP, MMVD and affine MMVD modes.

**Test 3.11: Non-adjacent spatial neighbour for affine merge mode**

Non-adjacent spatial neighbor positions used in the regular merge mode are added for affine merge and affine AMVP modes.

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 7 (a), the checking orders of the neighbors on the left and above sides are bottom-to-up and right-to-left, respectively.



Figure 7. 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 7 (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 8. 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 3.12. 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 7 (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.

**Test 3.12: History-parameter-based affine model inheritance**

In the test, affine model can be inherited from a previously affine-coded block which may not be neighboring to the current block. 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 five reference indices are supported for each reference list in the history table.

A history-affine-parameter-based candidate is derived from a neighbouring 4×4 block denoted as A0, A1, B0, B1 or B2 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.

Figure 9 shows an example of history based affine candidate derivation from block A0. The affine parameters {*a*0, *b*0, *c*0, *d*0} are directly fetched from one entry of category HPTIdx(RefListA0, refIdx0A0) in the history table. The affine parameters from the entry are used to derive CPMV with the center position of A0 as the base position and the MV of block A0 as the base MV.

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 10 in RA and 12 in LB, respectively.

Besides, such candidates are used to derive MVs located at the center of the current block, as regular merge candidates. The size of regular merge candidate list is increased from 10 to 11 for random access configurations to accommodate the newly added regular merge candidates.



Figure 9. Affine candidate derived from the history table

In Test 3.12b, non-adjacent affine candidates (Test 3.11) and history-based affine candidates (Test 3.12a) are all put into the subblock-based merge candidate list.

**Test 3.13: Modifications of IBC merge/AMVP list construction**

The IBC merge/AMVP list construction is modified as follows:

* Only if an IBC merge/AMVP candidate is valid, it can be inserted into the IBC merge/AMVP candidate list.
* Above-right, bottom-left, and above-left spatial candidates and one pairwise average candidate can be added into the IBC merge/AMVP candidate list.
* Template based adaptive reordering (ARMC-TM) is applied to IBC merge list

Results (AI, RA, LB, LP; each column Y/U/V/enc/dec)



Test 3.1 (GPM intra/inter) can be seen to have benefit independent from other tests on MV coding etc.

3.1c has the best tradeoff, no relevant increase in encoder/decoder run time, and re-uses existing pipelines of DIMD, TIMD and GPM-MMVD. Decision: Adopt JVET-Y0065 Test 3.1c, not in LP CTC

Test 3.3a (TMVP improvement) – it is asked why “method1” from the previous proposal JVET-X0091 was not tested standalone but only in combination. It is however noted that such a variant was used in the combination tests 3.6a and 3.6d

Test 3.3a has mainly benefit on LB

Test 3.4 generally performs favorable on RA, less gain on LB

Test 3.5a is conceptually similar to 3.4, but has worse performance on RA, but better on LB

The combination tests 3.6a/d unveil that the benefits of the methods (3.4 for RA, 3.3a and 3.5a for LB are complementary)

Test 3.2 was originally planned to be tested in more combinations, but was finally only included in 3.5b where the additional benefit compared to 5.3a is marginal.

It is mentioned that another version of test 3.4 was used in 3.6d which has an early termination mechanism in the reordering. It is asserted that this might be the reason for the slight decrease in decoder runtime, but also shows a performance drop of 0.05% in RA. 3.6a also has more benefit in classes A. It is only a combination of two modifications rather than three.

Decision: Adopt JVET-Y0134 Test 3.6a (where in this combination Test 3.3a is a modified version without the additional selection process on the temporal position where the MV comes from – this is confirmed by cross-checkers)

Test 3.7 is encoder-only. It shows that potential gain of 0.1% is achieved when increasing the numbr of TM merge candidates from 4 to 7, but also increases encoder runtime slightly.

Test 3.9 uses TM based reordering for MMVD (or MMVD and affine MMVD). 3.10 uses TM for MVD sign prediction. Gains of these methods are somewhat additive (as shown by cmbined test 3.9d), and in combination this gives 0.25% in RA, 0.34% in LB. Only small increase in encoder/decoder run time.

It was asked if a similar method had also been tested on GPM MMVD, but this was not the case.

Decision: Adopt JVET-Y0067 Test 3.9d

Test 3.11 and 3.12 target the improvement of candidates for affine merge. 3.11 uses non-adjacent neighbors whereas 3.12a is history based (table). 3.12b is a combination of both, where it is however mentioned that the method from 3.11 was somewhat modified. Results unveil that the gain of 3.12b is somewat similar to 3.11, from which it appears that 3.12a does not provide additional benefit.

It is claimed that 3.11 does not require an additional buffer, as the non-adjacent neighbors are stored anyway. In the further discussion, it is clarified that this might not be correct, as the control point MVs need to be stored rather than regular MVs. It should be clarified offline how large the additional buffer requirements would be.

It is further pointed out that more extensive template matching is used in 3.11 than in 3.12a. The additional gain may not be due to the different candidates.

It was requested to study offline the complexity impact (memory requirements and amount of processing) to get a better understanding about the differences between the two proposals. It was later concluded that it is more appropriate to further study the two proposals in the next EE again, and make a more thorough analysis about the complexit aspects in that context.

Further investigate in EE: 3.11 (from JVET-Y0153) and 3.12a (from JVET-Y0145).

Test 3.13 is modifying the IBC merge/AMVP list derivation, introducing template matching. Even though the gain is relatively low and only for SC classes (1% in class TGM, 0.2% in class F), this is asserted to be straightforward, as similar approaches are already existing elsewhere in the ECM design.

Decision: Adopt JVET-Y0058 Test 3.13

Test 4: Transform and coefficient coding

**Test 4.1: Sign prediction**

In test 4.1.a, instead of always selecting the first K transform coefficients in raster-scan order of a top-left 4x4 area for sign prediction, the transform coefficients with the largest K qIdx value of the top-left 4x4 area were selected. qIdx value is the transform coefficient level after compensating the impact from the multiple quantizers in DQ. A larger qIdx value will produce a larger de-quantized transform coefficient level. qIdx is derived as follows

qIdx = (abs(level) << 1) - (state & 1);

where level is the transform coefficient level parsed from the bitstream and state is a variable maintained by the encoder and decoder in DQ.

On top of test 4.1.a, the sign prediction area was extended to maximum 32x32 in test 4.1.c. In this test, signs of top-left MxN block are predicted. The value of M and N is computed as follows:

where, *w* and *h* are the width and height of the transform block.

The maximum number of predicted signs is kept unchanged.

**Test 4.2: Enhanced sign prediction**

Two enhancements to sign prediction are tested:

1. Instead of selecting the first *M* coefficients in raster-scan order, the non-zero transform coefficients are selected based on their impacts on the quality of the reconstructed border samples of one TB, by maximizing L1-norm cost of the variations caused by one transform coefficient on the reconstructed border samples:

where represents the transform coefficient at coordinate in the TB; represents the template that corresponds to the transform coefficient, which consists of reconstructed samples at the top and left border of the TB.

1. Sign prediction is extended to LFNST, and only up to 4 coefficients in the top-left CG are allowed to be predicted for one LFNST TB.

**Test 4.3: Combination of sign prediction methods**

Test 4.3.a is on top of test 4.1.c and additionally applies sign prediction to LFNST blocks as follows:

1. The transform coefficients with the largest K qIdx value are selected for sign prediction. qIdx value is the transform coefficient level after compensating the impact from the multiple quantizers in DQ. A larger qIdx value will produce a larger de-quantized transform coefficient level.
2. The sign prediction area is extended. The signs of top-left MxN block are predicted. The value of M and N is computed as , , where *w* and *h* are the width and height of the transform block
3. The sign prediction is also applied to LFNST block. And for LFNST block, a maximum of 4 coefficients in the top-left 4x4 area are allowed to be sign predicted.

Test 4.3.b is on top of test 4.2.b and additionally extends the sign prediction area to maximum 32x32:

1. Adaptively select the signs to be predicted based on their impacts on the reconstructed border samples of one TB.
2. The sign prediction is applied to LFNST blocks and for LFNST block, a maximum of 4 coefficients in the top-left 4x4 area are allowed to be sign predicted.
3. The maximum area for sign prediction is extended. Encoder selects the maximum area size from four allowed values 4, 8, 16 and 32 based on configuration, sequence class and QP, and signaled it in SPS.

**Test 4.4: Adaptive intra MTS**

Number of intra MTS candidates are adaptively selected (between 1, 4 and 6 MTS candidates) depending on the sum of absolute value of transform coefficients. The sum is compared against the two fixed thresholds to determine the total number of allowed MTS candidates:

1 candidate: sum <= th0

4 candidates: th0 < sum <= th1

6 candidates: sum > th1

The existed the non-DCT2 transform kernels (DST7, DCT8, DCT5, DST4, DST1, and identity transform) are used to derive additional transform pairs.

The encoder speedup is done as follows: for a given intra mode, the corresponding sum of absolute coefficients from the DCT2 pass (sumDCT2) is stored. For the MTS pass, the number of MTS candidates for RD cost evaluation is determined as follows:

1 candidate: sumDCT2 <= th0

4 candidates: th0 < sumDCT2 <= th1

6 candidates: sumDCT2 > th1

Additionally, for inter-picture, if the SATD cost, i.e., difference between the original and predicted signals, of all intra modes are higher than a scaled version of the best inter mode SATD cost then MTS RD cost evaluation is skipped.

The following tests have been carried out:

* Test 4.4a: adaptive intra MTS.
* Test 4.4b: only encoder modification.
* Test 4.4c: adaptive intra MTS with thresholds signaled in SPS.
* Test 4.4d: fixed 6 MTS candidates (without adaptivity).

Results (AI, RA, LB; each column Y/U/V/enc/dec)



The approaches of 4.1 and 4.2 (as basic approaches for sign prediction) are competingAdding additional elements (in particular extension to LFNST) increases the gain in a similar way for both

Adaptive area reduces decoder complexity in similar way in both

Selection would need to be done between 4.3a\* and 4.3b

More detailed analysis to be done by proponents on the complexity impact of the two proposals (including impact on parsing, number of computations).

It was found that 4.3a\* has an undesirable property that sign selection is necessary during parsing. 4.3b has an undesirable amount of computations particularly in 32x32 area case. The analysis is included in JVET-Y0141 (combined test, test 3), reviewed in session 20 (Wed. 19 Jan. 2125). See further notes under JVET-Y0141.

- 4.4a Adaptive intra MTS (number of transforms derived from sum of coefficient levels and based on thresholds applied on this). Fixed thresholds are used.

- 4.4b Encoder modification is incurring loss, except for LB

- 4.4c Signalling thresholds does not provide advantage according to results, but it was not adaptive

- 4.4d Fixed candidates slightly less gain, increased encoding time.

Before MTS syntax can be parsed, it is necessary to decode and sum the coefficients. This is not much different from how it is done in VVC, where the MTS syntax also comes after the transform block.

Decision: Adopt JVET-Y0142 Test 4.4a (adaptive intra MTS with fixed threshold).

It was emphasized that performing further study on adapting the threshold (e.g., depending on QP, optimized per sequence) with the purpose of signalling (test 4.4.c) would be interesting. Further study is recommended on this.

Test 5: In-loop filtering

**Test 5.1: Edge-based classifier for CCSAO**

Currently, CCSAO uses only band classifier. In this test, edge-based classifier is added to CCSAO.

Similar to SAO, edge-based classifier uses the four 1-D directional patterns for sample classification: horizontal, vertical, 135° diagonal and 45° diagonal, as shown in the figure below, where each sample is classified based on the sample difference between the luma sample value labeled as “c” and its two neighbor luma samples labeled as “a” and “b” along the selected 1-D pattern.



Figure 10. Four 1-D directional patterns for CCSAO EO sample classification: horizontal (EO class = 0), vertical (EO class =1), 135° diagonal and 45° diagonal

Both the sample differences “a-c” and “b-c” are compared against a pre-defined threshold value (Th) to derive the final “class\_idx” information. The encoder selects the best “Th” value from an array of pre-defined threshold values based on RDO and the index into the “Th” array is signalled.

The Edge-based classifier process is formulated as follows:

Ea=(a-c<0)? (a-c<(-Th)? 0:1) : (a-c<(Th)? 2:3)

Eb=(b-c<0)? (b-c<(-Th)? 0:1) : (b-c<(Th)? 2:3)

= \* 16 + Ea \* 4 + Eb

variable “” in equation (3) is derived as follows.

(or) (or) ,

wherein, sample “cur” is the current sample being processed, col1 and col2 are the collocated samples. When luma samples are processed, col1 and col2are the collocated Cb and Cr samples, respectively. When chroma Cb samples are processed, col1 and col2are the collocated Y and Cr samples, respectively. When chroma Crsamples are processed, col1 and col2are the collocatedand Cb samples, respectively.

Based on RDO, encoder signals one among the samples cur, col1, or col2 used in deriving the band information.

**Test 5.2: Adaptive filter shape selection for ALF**

In the test, two candidate filter shapes: a diamond shape candidate as shown in Figure 11(a), which is currently applied in ECM, and a cross shape candidate as shown in Figure 11(b)/(c), which is newly introduced, can be adaptively selected by the luma filters in ALF. The number of coefficients that need to be trained and signaled for a luma filter is 20 with all the filter shapes.

In Test 5.2a, filter shape selection is performed between shapes (a) and (b), while in Test 5.2b shapes (a) and (c) are used as the shape candidates. The only difference between the two shapes (b) and (c) is the maximum filter length.

|  |  |  |
| --- | --- | --- |
| 图表  低可信度描述已自动生成 | 图表  描述已自动生成 | 图表  描述已自动生成 |
| (a) | (b) | (c) |

Figure 11. Adaptive ALF shapes.

In each APS, a shape index for the online-trained luma filters is signaled to decoder. Each APS contains the luma filters that are associated with the filter shape index.

For each CTB, an APS index is signaled to indicate which luma filter shape is used to filter the current CTB. When filtering a luma sample, the coefficients and clip indices are also rearranged according to the corresponding filter shape.

Results (AI, RA, LB; each column Y/U/V/enc/dec)



5.2a: The modified filter shapes (15x15) would require more line buffers, as ECM is currently not using virtual boundary concept, and the maximum filter size of ECM (for fixed filter) is 13x13 which is the same as for 5.2b (9x9 for adaptive currently)

Both adaptive ALF solutions might have some impact on implementation, as the positions are different from the diamond-shaped filters.

5.2b gives 0.1% gain with 2% encoder runtime increase for RA. Not good tradeoff in terms of performance vs. complexity.

5.1 ia asserted to be a straightforward modification, and gives good gain in chroma (as can be expected as it is for CCSAO), and even 0.2% luma gain in LB.

Decision: Adopt JVET-Y0106

### EE2 contributions: Enhanced compression beyond VVC capability (18)

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

[JVET-Y0058](https://jvet-experts.org/doc_end_user/current_document.php?id=11252) EE2-3.13: Modifications of IBC Merge/AMVP List Construction [N. Zhang, K. Zhang, L. Zhang, J. Xu (Bytedance)]

[JVET-Y0206](https://jvet-experts.org/doc_end_user/current_document.php?id=11413) Cross-check of JVET-Y0058 (EE2-3.13: Modifications of IBC merge/AMVP list construction) [D. Ruiz Coll (Ofinno)] [late]

[JVET-Y0065](https://jvet-experts.org/doc_end_user/current_document.php?id=11259) EE2-3.1: GPM with inter and intra prediction (JVET-X0166) [Y. Kidani, H. Kato, K. Kawamura (KDDI), H. Jang, S. Kim, J. Lim (LGE), Z. Deng, K. Zhang, L. Zhang (Bytedance)]

[JVET-Y0182](https://jvet-experts.org/doc_end_user/current_document.php?id=11389) Cross-check of JVET-Y0065 (Test 3.1a): EE2-3.1: GPM with inter and intra prediction [K. Sato (OPPO)] [late]

[JVET-Y0067](https://jvet-experts.org/doc_end_user/current_document.php?id=11261) EE2-3.9 and EE2-3.10: TM based reordering for MMVD and affine MMVD and MVD sign prediction [M. Salehifar, Y. He, K. Zhang, N. Zhang, L. Zhang (Bytedance), Y. Zhang, B. Ray, H. Huang, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-Y0205](https://jvet-experts.org/doc_end_user/current_document.php?id=11412) Cross-check of JVET-Y0067 (EE2-3.10: MVD sign prediction) [V. Rufitskiy, A. Filippov (Ofinno)] [late]

[JVET-Y0093](https://jvet-experts.org/doc_end_user/current_document.php?id=11287) EE2-3.3: On TMVP improvement [R.-L. Liao, J. Chen, Y. Ye, X. Li (Alibaba)]

[JVET-Y0183](https://jvet-experts.org/doc_end_user/current_document.php?id=11390) Cross-check of JVET-Y0093 (Test 3.3): EE2-3.3: On TMVP improvement [Z. Xie (OPPO)] [late]

[JVET-Y0094](https://jvet-experts.org/doc_end_user/current_document.php?id=11288) EE2-4.1: Test Results on Sign Prediction Improvement J. Chen, Y. Yan, R.-L. Liao, X. Li (Alibaba)]

[JVET-Y0185](https://jvet-experts.org/doc_end_user/current_document.php?id=11392) Cross-check of JVET-Y0094 (Test 4.1): EE2-4.1: Test Results on Sign Prediction Improvement [L. Xu, Y.Yu (OPPO)] [late]

[JVET-Y0100](https://jvet-experts.org/doc_end_user/current_document.php?id=11294) EE2-3.2: Pairwise merge candidate [G. Laroche, P. Onno, R. Bellessort (Canon)]

[JVET-Y0106](https://jvet-experts.org/doc_end_user/current_document.php?id=11300) EE2-5.1: Edge-based classifier for Cross-component Sample Adaptive Offset (CCSAO) [A. M. Kotra, N. Hu, V. Seregin, M. Karczewicz (Qualcomm), C.-W. Kuo, X. Xiu, Y.-W. Chen, H.-J. Jhu, W. Chen, N. Yan, X. Wang (Kwai)]

[JVET-Y0196](https://jvet-experts.org/doc_end_user/current_document.php?id=11403) Crosscheck of JVET-Y0106 (EE2-5.1: Edge-based classifier for Cross-component Sample Adaptive Offset (CCSAO) [X. Li (Alibaba)] [late]

[JVET-Y0116](https://jvet-experts.org/doc_end_user/current_document.php?id=11310) EE2-2.1: Extended MRL candidate list [K. Cao, Y.-J. Chang, B. Ray, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-Y0171](https://jvet-experts.org/doc_end_user/current_document.php?id=11378) Cross-check of JVET-Y0116 (Test 2.1a): EE2-2.1: Extended MRL candidate list [K. Sato (OPPO)] [late]

[JVET-Y0204](https://jvet-experts.org/doc_end_user/current_document.php?id=11411) Cross-check of JVET-Y0116 (EE2-2.1: Extended MRL candidate list) [V. Rufitskiy, A. Filippov (Ofinno)] [late]

[JVET-Y0132](https://jvet-experts.org/doc_end_user/current_document.php?id=11326) EE2-3.7: On the increased number of TM merge candidates [Y.-J. Chang, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-Y0134](https://jvet-experts.org/doc_end_user/current_document.php?id=11328) EE2-3.4, EE2-3.5, EE2-3.6: Experimental results of the MV candidates reordering in candidate types based on template matching costs [Y.-J. Chang, H. Huang, V. Seregin, C.-C. Chen, M. Karczewicz (Qualcomm), R.-L. Liao, J. Chen, Y. Ye, X. Li (Alibaba), L. Zhao, K. Zhang, N. Zhang, L. Zhang (Bytedance), G. Laroche, P. Onno, R. Bellessort (Canon)]

[JVET-Y0215](https://jvet-experts.org/doc_end_user/current_document.php?id=11424) Cross-check of JVET-Y0134: EE2-3.5a, EE2-3.6d: Experimental results of the MV candidates reordering in candidate types based on template matching costs [F. Urban (InterDigital)] [late]

[JVET-Y0137](https://jvet-experts.org/doc_end_user/current_document.php?id=11331) EE2-4.2: Enhanced sign prediction [X. Xiu, Y.-W. Chen, N. Yan, C.-W. Kuo, H.-J. Jhu, W. Chen, X. Wang (Kwai)]

[JVET-Y0184](https://jvet-experts.org/doc_end_user/current_document.php?id=11391) Cross-check of JVET-Y0137 (Test 4.2): EE2-4.2: Enhanced sign prediction [L. Xu, Y. Yu (OPPO)] [late]

[JVET-Y0138](https://jvet-experts.org/doc_end_user/current_document.php?id=11332) EE2-4.3: Combined Test Results of EE2-4.1 and EE2-4.2 on Sign Prediction [J. Chen, Y. Ye, R.-L. Liao, X. Li (Alibaba), X. Xiu, Y.-W. Chen, N. Yan, C.-W. Kuo, H.-J. Jhu, W. Chen, X. Wang (Kwai)]

[JVET-Y0226](https://jvet-experts.org/doc_end_user/current_document.php?id=11435) CrossCheck of JVET-Y0138 (EE2-4.3: Combined Test Results of EE2-4.1 and EE2-4.2 on Sign Prediction) [L.-F. Chen (Tencent)] [late]

[JVET-Y0142](https://jvet-experts.org/doc_end_user/current_document.php?id=11336) EE2-4.4: adaptive intra MTS [B. Ray, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-Y0197](https://jvet-experts.org/doc_end_user/current_document.php?id=11404) Crosscheck of JVET-Y0142 (EE2-4.4: Adaptive intra MTS) [T. Hashimoto (Sharp)] [late]

[JVET-Y0145](https://jvet-experts.org/doc_end_user/current_document.php?id=11339) EE2-3.12a: History-parameter-based affine model inheritance [K. Zhang, L. Zhang, Z. Deng, N. Zhang, Y. Wang (Bytedance)]

[JVET-Y0146](https://jvet-experts.org/doc_end_user/current_document.php?id=11340) EE2-3.12b/c: A joint test of EE-2.3.11 and EE-2.3.12a [K. Zhang, L. Zhang, Z. Deng, N. Zhang, Y. Wang (Bytedance), W. Chen, X. Xiu, Y.-W. Chen, H.-J. Jhu, C.-W. Kuo, X. Wang (Kwai)]

[JVET-Y0147](https://jvet-experts.org/doc_end_user/current_document.php?id=11341) EE2-5.2: Adaptive Filter Shape Selection for ALF [W. Yin, K. Zhang, L. Zhang (Bytedance), N. Hu, V. Seregin, M. Karczewicz (Qualcomm), M. G. Sarwer, R.-L. Liao, J. Chen, Y. Yan, X. Li (Alibaba)]

[JVET-Y0229](https://jvet-experts.org/doc_end_user/current_document.php?id=11438) Crosscheck of JVET-Y0147 (EE2-5.2: Adaptive Filter Shape Selection for ALF) [J. Ström (Ericsson)] [late]

[JVET-Y0150](https://jvet-experts.org/doc_end_user/current_document.php?id=11344) EE2-1: Tests on unsymmetric partitioning methods [K. Zhang, L. Zhang, Z. Deng, N. Zhang, Y. Wang (Bytedance), F. Urban, K. Naser, F. Galpin (InterDigital)]

[JVET-Y0180](https://jvet-experts.org/doc_end_user/current_document.php?id=11387) Cross-check of JVET-Y0150 "EE2-1: Tests on unsymmetric partitioning methods", tests 1.1a and 1.1b [F. Le Léannec (Xiaomi)] [late]

[JVET-Y0246](https://jvet-experts.org/doc_end_user/current_document.php?id=11455) Crosscheck of EE2-1.1c from JVET-Y0150 (EE2-1.1: Tests on unsymmetric partitioning methods) [J. Sauer (Huawei)] [late]

[JVET-Y0153](https://jvet-experts.org/doc_end_user/current_document.php?id=11347) EE2-3.11: Non-adjacent spatial neighbors for affine merge mode [W. Chen, X. Xiu, Y.-W. Chen, H.-J. Jhu, C.-W. Kuo, X. Wang (Kwai)]

[JVET-Y0207](https://jvet-experts.org/doc_end_user/current_document.php?id=11414) Crosscheck of JVET-Y0153 (EE2-3.11: Non-adjacent spatial neighbors for affine merge mode) [K. Zhang (Bytedance)] [late]

[JVET-Y0216](https://jvet-experts.org/doc_end_user/current_document.php?id=11425) Crosscheck of JVET-Y0153 (EE2-3.11: Non-adjacent spatial neighbors for affine merge mode) [Y. Zhang (Qualcomm)] [late]

### EE2 related contributions (13)

Contributions in this area were discussed in session 7 at 2100–2300 UTC and session 8 at 2320–0010+1 UTC on Thursday 13 Jan. 2022 (chaired by JRO).

[JVET-Y0088](https://jvet-experts.org/doc_end_user/current_document.php?id=11282) EE2-related: IBC with Template Matching [A. Robert, K. Naser, T. Poirier, Y. Chen, F. Galpin (InterDigital)]

This contribution proposes to enable Template Matching with IBC. The reported simulation results of the proposed method on top of the ECM-3.1 in SCC test conditions show gains of {-0.53%, -0.44%, -0.43%}, {-0.36%, -0.34%, -0.37%} and {-0.20%, -0.44%, -0.43%} for YUV in AI, RA and LDB respectively with an unchanged complexity for both encoder and decoder. On top of the EE2-3.13 test, the reported results are slightly better especially in LDB with {-0.51%, -0.47%, -0.52%}, {-0.39%, -0.27%, -0.51%} and {-0.38%, -0.69%, -1.04%} for YUV in AI, RA and LDB respectively. The combination of the EE2-3.13 test and the proposed method can achieve, on top of the ECM-3.1, {-1.18%, -1.05%, -1.09%}, {-0.94%, -0.79%, -1.08%} and {-0.77%, -1.24%, -1.34%} for YUV in AI, RA and LDB respectively also with a similar complexity for both encoder and decoder.

Interesting gain over the proposal adopted from EE, in particular for class F.

Investigate in next EE (together and also in combination with other IBC related proposals)

[JVET-Y0209](https://jvet-experts.org/doc_end_user/current_document.php?id=11416) Crosscheck of JVET-Y0088 (EE2-related: IBC with Template Matching) [N. Zhang (Bytedance)] [late]

[JVET-Y0119](https://jvet-experts.org/doc_end_user/current_document.php?id=11313) EE2-related: On Extended MRL Intra Prediction [K. Sato, Y. Yu, H. Yu, Z. Xie, L. Xu, F. Wang, H. Huang, J. Gan, D. Wang (OPPO)]

The intra prediction method of multiple reference line (MRL) has been adopted in the VVC standard and the extended MRL intra prediction is being investigated in EE2. This contribution proposes to restrict the increase of line buffer for the above neighbors with the extended MRL. The proposed method is implemented on top of EE2 Test2.1 and it is reported that the simulation results over EE2 Test2.1a and ECM3.1 are as follows:

xx %; xx % xx % for BD-{Y,Cb,Cr} over the anchor: EE2 Test2.1a

xx %; xx % xx % for BD-{Y,Cb,Cr} over the anchor: ECM3.1

It is requested that the proposed methods be studied with various combinations of the number of extended reference lines both for above and left in an EE.

Results incomplete, but by tendency indicate loss compared to EE method.

No action.

[JVET-Y0120](https://jvet-experts.org/doc_end_user/current_document.php?id=11314) EE2-related: Non-adjacent temporal MVP [F. Wang, Z. Xie, Y. Yu, H. Yu, L. Xu, K. Sato, J. Gan, D. Wang (OPPO)]

This contribution proposes to include some non-adjacent temporal MVPs (NATMVP) in not only the merge candidate list but also the AMVP candidate list. The proposed candidate NATMVPs are derived from the blocks that are temporally collocated non-adjacently right-side and down-side to the current block. These NATMVPs are added into the merge candidate list and the AMVP candidate list after the non-adjacent spatial MVPs (NASMVP) of the same distance. It is reported that the simulation results are as follows:

Test 1: RA: -0.02%, -0.05%, -0.05%, 102% 104%; LB: -0.14%, -0.08%, -0.08%, 103% 106% over ECM3.1

Test 2: RA: -x.xx%, -x.xx%, -x.xx%, xx% xx%; LB: -x.xx%, -x.xx%, -x.xx%, xx% xx% over ECM3.1

Test 3: RA: -x.xx%, -x.xx%, -x.xx%, xx% xx%; LB: -x.xx%, -x.xx%, -x.xx%, xx% xx% over ECM3.1 EE2-3.6a

Results incomplete, but so far no significant gain – no action

[JVET-Y0199](https://jvet-experts.org/doc_end_user/current_document.php?id=11406) Crosscheck of JVET-Y0120 (EE2-related: Non-adjacent temporal MVP) [R.-L. Liao (Alibaba) [late]

[JVET-Y0121](https://jvet-experts.org/doc_end_user/current_document.php?id=11315) EE2-4.2-related: On adaptive sign prediction position selection [L. Xu, Y. Yu, H. Yu, Z. Xie, F. Wang, D. Wang (OPPO)]

This contribution proposes a simplification to the sign prediction tool in EE2-4.2a of JVET-X2024. Specifically, the calculation of the energy calculation for transform coefficient selection is simplified for TBs at either the top or left picture boundary. The proposed simplification is implemented on top of EE2-4.2a and EE2-4.3b. It reportedly provides Y, U, V BD rate changes with encoder and decoder runtime as follows:

Simulation results over EE2-4.2a,

AI: {-0.00%, -0.03X%, -0.02%}

RA: {-X%, -X%, -X%}

LDB: {-X%, -X%, -X%}

Simulation results over EE2-4.3b,

AI: {-X%, -X%, -X%}

RA: {-X%, -X%, -X%}

LDB: {-X%, -X%, -X%}

Results incomplete.

This is more at the level of fine-tuning cleanup which would be relevant at final stage of standardization – no action in context of exploration.

[JVET-Y0189](https://jvet-experts.org/doc_end_user/current_document.php?id=11396) Crosscheck of JVET-Y0121: EE2-4.2-related: On adaptive sign prediction position selection [X. Xiu (Kwai)] [late] [miss]

[JVET-Y0133](https://jvet-experts.org/doc_end_user/current_document.php?id=11327) EE2-related: BVP candidate adjustment based on IBC reference region implemented on top of test EE2-3.13 [D. Ruiz Coll, A. Filippov, V. Rufitskiy, T. M. Bae (Ofinno)] [late]

This contribution proposes two techniques are proposed in the IBC Merge/AMVP list construction. (1) the clipping of the invalid candidates BVP pointing outside of the IBC reference region, (2) the replacement of the null candidates included completing the list when there are no sufficient valid candidates, by a set of valid candidates distributed in the IBC reference region.

On top of ECM-3.1, the joint test results for both techniques, on top of the EE2-3.13, for class F and class TGM are reported as follows:

AI: Class F -0.53%, 103%, 103%; Class TGM -1.18%, 100%, 95%

RA: Class F -0.38%, 102%, 99%; Class TGM -1.08%, 101%, 99%

Interesting gain over the proposal adopted from EE.

Investigate in next EE (together and also in combination with other IBC related proposals)

[JVET-Y0230](https://jvet-experts.org/doc_end_user/current_document.php?id=11439) Crosscheck report of JVET-Y0133 (EE2-related: BVP candidate adjustment based on IBC reference region implemented on top of test EE2-3.13) [H. Gao (Kwai)] [late]

[JVET-Y0140](https://jvet-experts.org/doc_end_user/current_document.php?id=11334) EE2-related: On the LCU boundary processing by intra-prediction tools [A. Filippov, V. Rufitskiy, D. Ruiz Coll (Ofinno)] [late]

This contribution presents a modification of the number of reference lines available for intra prediction of a block on the above LCU boundary. Three methods are proposed, enabling 1,3 and 5 lines, respectively. The first methods comprises two subtests. In the first one, the derivation of context for TIMD and DIMD is not modified. In the second subtest a neighbor above block is not available if it is outside the current LCU. The proposed modification of test 1.1 provides the following overall results: 0.49% (Y) / 0.35% (U) / 0.44% (V) and 0.31% (Y) / 0.29% (U) / 0.29% (V) in AI and RA configurations, respectively. The proposed modification of test 1.2 provides the following overall results: 0.48% (Y) / 0.36% (U) / 0.44% (V) and 0.30% (Y) / 0.22% (U) / 0.28% (V) in AI and RA configurations, respectively. The proposed modification of test 2 provides the following overall results: ‑0.05% (Y) / ‑0.12% (U) / -0.08% (V) and ‑0.01% (Y) / -0.01% (U) / 0.04% (V) in AI and RA configurations, respectively. The proposed modification of test 3 provides the following overall results: ‑0.06% (Y) /‑0.08% (U) /-0.06% (V) and ‑0.01% (Y) / 0.04% (U) / 0.07% (V) in AI and RA configurations, respectively.

It is agreed that there is some inconsistency in the way how many line buffers are used at CTU boundary by various tools. This requires more study before taking action. Design consistency alignment may be too much effort at this time of standardization.

It is also pointed out that CCLM and MRLP (where the gain comes from) in their original versions were using more line buffers, so this is reverting back to original proposals.

No action at this stage.

[JVET-Y0225](https://jvet-experts.org/doc_end_user/current_document.php?id=11434) Crosscheck of JVET-Y0140, JVET-Y0149 and JVET-Y0203 [R.-L. Liao (Alibaba)] [late]

Very limited results in v1

[JVET-Y0141](https://jvet-experts.org/doc_end_user/current_document.php?id=11335) EE2-4.3 related: More combined test results for sign prediction [J. Chen, Y. Ye, R.-L. Liao, X. Li (Alibaba), X. Xiu, Y.-W. Chen, N. Yan, C.-W. Kuo, H.-J. Jhu, W. Chen, X. Wang (Kwai)]

This contribution reports more combined test results for sign prediction. EE2-4.3.a[1] combines three improvements for sign prediction proposed in last meeting: 1) the signs to be predicted are selected based on the value of qIdx; 2) the area of the sign prediction is extended to maximum 32x32; 3) sign prediction is applied to LFNST block. To reduce the complexity of EE2-4.3.a, two more combined tests are conducted on top of EE2-4.3.a in this contribution. In test 1, maximum area for sign prediction is not always set to 32x32 but is set by the encoder and signaled in SPS; in test 2, the sign prediction area is not extended (keep 4x4 as the anchor). Results of following two tests are reported:

* Test 1: adaptive maximum sign prediction area + selection of signs to be predicted based on qIdx + apply sign prediction to LFNST blocks
  + AI: -0.33% (Y), -0.20%(U), -0.22%(V), 105%(EncTime), 106% (DecTime)
  + RA: -0.29% (Y), -0.11%(U), -0.19%(V), 102%(EncTime), 100% (DecTime)
  + LB: -0.22% (Y), -0.17%(U), -0.15%(V), 101%(EncTime), 101% (DecTime)
* Test 2: selection of signs to be predicted based on qIdx + apply sign prediction to LFNST blocks
  + AI: -0.31% (Y), -0.17%(U), -0.18%(V), 103%(EncTime), 105% (DecTime)
  + RA: -0.22% (Y), -0.03%(U), -0.16%(V), 101%(EncTime), 100% (DecTime)
  + LB: -0.10% (Y), -0.07%(U), -0.26%(V), 101%(EncTime), 102% (DecTime)

No need for presentation, was already reviewed with EE summary.

Later, a test 3 was added which makes further combination of elements from EE proposals 4.3a\* and 4.3b, however avoiding the complexity bottlenecks of both of them in the combination. This has only marginals worse performance compared to considering the two original proposals standalone (note that the gains cannot be additive, as they are conceptually similar). Crosscheck was also updated.

Y. Kidani inspected the code and reports he found it appropriate and straightforward.

“Official” crosscheckers (Y0175) found no issue and support the adoption, as well as other experts.

K. Naser reports that he has inspected the code and raised concerns about some of its elements, non-normative changes that were additionally included, changes to TIMD/MTS, as well as the results reported in terms of run time. Proponents are asked to provide the code of test 3, K. Naser and V. Seregin are asked to inspect it for possible problems and report back. Follow-up discussion in session 23 on Thursday 20 Jan.

[JVET-Y0250](https://jvet-experts.org/doc_end_user/current_document.php?id=11459) EE2 related: Code Inspection of Sign Prediction for LFNST (JVET-Y0141 and EE2-4.2b) [K. Naser, F. Galpin (InterDigital)] [late]

This contribution provides details about the inspection of the code related to sign prediction for LFNST that is present in EE2-4.2b test and in JVET-Y0141. It is asserted that additional changes in the code are associated with the method, specifically:

- SIMD implementation for TIMD, DIMD, LFNST

- Context retraining for MTS, LFNST, DIMD, TIMD.

- Added LMCS for bilateral filter RDO for certain sequence resolutions (class C and class B).

It is reported that the code was run excluding the re-training, and the gain was lower (only class D tested so far). Further, SIMD was disabled, and runtimes increased. It can be concluded that some of the optimizations made could alos give gain on the ECM anchor without the new elements of the proposal. It is estimated that the actual gain in AI is approximately 0.05% lower. The realistic gain for AI would be around 0.25%, with encoder run time increase of 4%.

Otherwise, the implementation was checked various experts and no problems were found.

Decision: Adopt JVET-Y0141 test 3

The part of encoder optimization shall be implemented in a way that it can be disabled by macro

[JVET-Y0175](https://jvet-experts.org/doc_end_user/current_document.php?id=11382) Crosscheck of JVET-Y0141 (EE2-4.3 related: More combined test results for sign prediction) [Y. Wang (Bytedance)] [late]

[JVET-Y0149](https://jvet-experts.org/doc_end_user/current_document.php?id=11343) EE2-related: Modifications of the extended MRL candidate list [A. Filippov, V. Rufitskiy, E. Dinan (Ofinno)] [late]

This contribution presents modifications of the extended MRL list. Two methods are proposed. In the first method, when TIMD flag is on, selection of the MRL list is performed for left and top side individually in accordance with block width or block height, respectively. In this method, selection of the MRL list is aligned with the template sizes that are used by TIMD for short and long side lengths. In method 2, MRL lists are defined differently for small and larger blocks. The proposed method 1 for the case of 5 elements in the extended MRL list (EE2-2.1a) provides the following overall results: -0.11% (Y) / -0.11% (U) / ‑0.12% (V) and -0.04% (Y) / 0.04% (U) / 0.01% (V) in AI and RA configurations, respectively. The proposed method 1 for the case of 4 elements in the extended MRL list provides the following overall results: -0.09% (Y) / -0.09% (U) / -0.05% (V) and -0.03% (Y) / 0.01% (U) / 0.08% (V) in AI and RA configurations, respectively. The proposed method 2 for the case of 4 elements in the extended MRL list (EE2-2.1b) provides the following overall results: -0.08% (Y) / -0.11% (U) / -0.04% (V) and -0.03% (Y) / 0.07% (U) / 0.11% (V) in AI and RA configurations, respectively. A combination of method1 and method 2 for the case of 4 elements in the extended MRL list (EE2-2.1b) provides the following overall results: ‑0.10% (Y) / ‑0.06% (U) / -0.06% (V) and -0.04% (Y) / 0.02% (U) / 0.02% (V) in AI and RA configurations, respectively.

The combination of test 1 and test 2 provides a slight simplification of the EE2-2.1 method for encoder (5 instead of 6 checks) with very slight gain.

After offline consideration with the contributors of EE2-2.1 it was concluded that this should not be investigated in next EE.

[JVET-Y0225](https://jvet-experts.org/doc_end_user/current_document.php?id=11434) Crosscheck of JVET-Y0140, JVET-Y0149 and JVET-Y0203 [R.-L. Liao (Alibaba)] [late]

[JVET-Y0159](https://jvet-experts.org/doc_end_user/current_document.php?id=11353) EE2-related: inter MTS refinement on adaptive intra MTS (EE2-4.4) [T. Hashimoto, T. Ikai (Sharp)]

This contribution proposes to improve inter MTS aspect of Adaptive intra MTS (AMTS) in EE2-4.4a by restricting AMTS’s candidate number increase in the case of inter MTS. It is noted that AMTS (EE2-4.4a) extends the number of MTS (Multiple transform selection) candidate up to 6:

* VVC 4
* ECM 4 (but the MTS candidate can be chosen from more transforms)
* AMTS 6

It is reported the simulation results of proposed method compared to the ECM-3.1 with inter MTS is:

* RA: BD-rate YUV: -0.06%, -0.12%, -0.02%; EncT: 99%; DecT: 100%
* LDB: BD-rate YUV: -0.23%, -0.22%, -0.47%; EncT: 103%; DecT: 100%

It is recommended to adopt method1 because it is asserted to show best performance without any encoder time increase.

In v2, results are updated.

It is reported that the AMTS software from EE has an encoder/decoder mismatch when enabled for inter, beyond a bug in ECM (for which a ticket was issued)

Method 1 keeps the number of candidates for inter MTS fixed as currently (4).

Question: Which are the 6 candidates in method 3 for inter MTS? Not clear. Probably the selection of candidates would require more investigation. Currently method 3 performs worse than method 1.

Currently inter MTS is not used in CTC

Method 1 provides 0.37%/0.59% in RA/LB, with 20%/30% encoder run time increase.

Decision(SW/BF): Adopt JVET-Y0159 method 1

[JVET-Y0217](https://jvet-experts.org/doc_end_user/current_document.php?id=11426) Crosscheck of JVET-Y0159 (EE2-related: inter MTS refinement on adaptive intra MTS (EE2-4.4)) [B. Ray (Qualcomm)] [late]

Very limited results in v1

[JVET-Y0160](https://jvet-experts.org/doc_end_user/current_document.php?id=11354) EE2-3.13-related: Enlarged HMVP table for IBC [N. Zhang, K. Zhang, L. Zhang (Bytedance)]

In this contribution, the number of HMVP candidates for IBC is increased up to 30, in addition to the modifications of IBC Merge/AMVP list construction tested in EE2-3.13.

On top of the ECM3.1, simulation results of the proposed method together with the modifications in EE2-3.13 are reported as below:

AI: Class F -0.52%, 103%, 100%; Class TGM -2.51%, 103%, 98%

RA: Class F -0.47% , 100%, 101%; Class TGM -2.03%, 100%, 102%

Interesting gain over the proposal adopted from EE.

Investigate in next EE (together and also in combination with other IBC related proposals)

[JVET-Y0193](https://jvet-experts.org/doc_end_user/current_document.php?id=11400) Cross-check of JVET-Y0160 (EE2-3.13-related: Enlarged HMVP table for IBC) [J. Zhao (LGE)] [late]

[JVET-Y0161](https://jvet-experts.org/doc_end_user/current_document.php?id=11355) EE2-3.12-related: Extensions of history-parameter-based affine model inheritance [K. Zhang, L. Zhang, Z. Deng, N. Zhang, Y. Wang (Bytedance)]

Presented in session 8

This contribution presents additional results for history-parameter-based affine model inheritance in test EE2-3.12. Based on the code of EE2-3.12, some extensions are proposed as below:

Aspect #1: The size of sub-block-based merge candidate list is increased to 15, and all the affine merge candidates are involved in the adaptive reordering merge candidates (ARMC) process.

Aspect #2: The number of neighboring blocks, which are used to generate affine merge candidates jointly with affine models stored in tables, are increased from five to seven.

Aspect #3: A second history-parameter table with base MV information is appended. And the history-parameter tables stored in the neighbouring CTU above and above-right to the current CTU can also be used to generate history-affine-parameter-based candidates.

Aspect #4: Pair-wised affine merge candidates are generated by two affine merge candidates which are history-derived or not history-derived.

Two tests are conducted:

Test #1: EE2-3.12 + Aspect #1

Test #2: EE2-3.12 + Aspect #1~#4

On top of ECM-3.1, simulation results of the proposed methods are reported as below:

Test #1 :RA: {-0.26%, -0.19%, -0.15%, 100%, 100%}; LB: {-0.25%, -0.06%, -0.17%, 100%, 100%}.

Test #2 :RA: {-0.33%, -0.27%, -0.20%, 100%, 100%}; LB: {}.

Interesting proposal, but has a number of additional elements

It is requested to perform a similar complexity analysis as for 3.11 and 3.12.

Investigate in EE together with 3.11 and 3.12.

[JVET-Y0235](https://jvet-experts.org/doc_end_user/current_document.php?id=11444) Crosscheck of JVET-Y0161 (EE2-3.12-related: Extensions of history-parameter-based affine model inheritance) [Y. Kidani (KDDI)] [late]

[JVET-Y0174](https://jvet-experts.org/doc_end_user/current_document.php?id=11381) EE2-1.1–related: additional tests on partitioning flexibility [F. Urban, K. Naser, F. Galpin (InterDigital), K. Zhang, L. Zhang, Z. Deng, N. Zhang, Y. Wang (Bytedance)] [late]

To bring more compression gain, it is possible to relax constraints about partitioning, for example by playing with the split depth. Additional gains can be therefore obtained at the expense of increased encoding run-time. Another track consists in adding new partitioning types, such as ABT or UQT.

This contribution reports experiments results based on these two options, and reportedly shows that enabling new partitions, namely ABT or UQT, leads to better compression gain/encoding run-time trade-offs in use cases where encoding run-time is less constrained.

The contribution shows that (from company-internal investigation, and only class B) that at higher encoder complexity the benefit of ABT increases to 0.35% at same encoder runtime, when compared against the existing partitioning.

In the discussion, it is pointed out that these data somehow show that further improvement of the partitioning does not provide as much gain any more as it was the case when defining it as a starting point of a new standard (in HEVC and VVC). Furthermore, it is argued that specifically ABT would have a significant impact on decoder complexity.

The partitioning concepts currently on the table don’t have enough potential to justify them as a starting point for a new codec. Further study encouraged.

[JVET-Y0203](https://jvet-experts.org/doc_end_user/current_document.php?id=11410) EE2-related: a combination of the extended MRL candidate list (JVET-Y0149) modifications with the extended LCU boundary processing area (JVET-Y0140) [A. Filippov, V. Rufitskiy, D. Ruiz Coll, E. Dinan (Ofinno)] [late]

This contribution presents results of a combined modification of extended MRL candidate list (proposed in JVET-Y0149) and the extension of LCU boundary processing area (JVET-Y0140). The proposed combination provides the following overall results for Test 1: -0.17% (Y) / -0.17% (U) / -0.11% (V) and ‑0.07% (Y) / -0.01% (U) / 0.01% (V) in AI and RA configurations, respectively. If to combine just the extension of LCU boundary processing area by 5 reference lines and EE2-2.1a (Test 2), the following overall results are reported: -0.13% (Y) / -0.13% (U) / -0.13% (V) and -0.05% (Y) / -0.06% (U) / 0.06% (V) in AI and RA configurations, respectively. If to perform Test 1 with 3 reference lines in LCU boundary constraints (Test 3) the following overall results are reported: -0.15% (Y) / -0.15% (U) / -0.15% (V) and ‑0.07% (Y) / -0.08% (U) / 0.00% (V) in AI and RA configurations, respectively. If to combine just the extension of LCU boundary processing area by 3 reference lines and EE2-2.1a (Test 4), the following overall results are reported: -0.10% (Y) / -0.14% (U) / -0.11% (V) and x.xx% (Y) / x.xx% (U) / x.xx% (V) in AI and RA configurations, respectively.

Contribution shows that the gains are additive.

[JVET-Y0225](https://jvet-experts.org/doc_end_user/current_document.php?id=11434) Crosscheck of JVET-Y0140, JVET-Y0149 and JVET-Y0203 [R.-L. Liao (Alibaba)] [late]

### ECM modifications beyond EE2 (25)

Contributions in this area were discussed in session 8 at 0010–0120 UTC on Friday 14 Jan. 2022, in session 11 at 2100–2300 UTC on Friday 14 Jan. 2022, and in sessions 17/18 at 2100-0050+1 UTC on Tuesday 18 Jan. 2022 (chaired by JRO).

[JVET-Y0055](https://jvet-experts.org/doc_end_user/current_document.php?id=11249) AHG12: Slope adjustment for CCLM [J. Lainema, A. Aminlou, P. Astola, R. G. Youvalari (Nokia)]

This contribution proposes to signal an adjustment for the slope parameter(s) used in cross-component linear model (CCLM) prediction of ECM. The adjustment is “tilting” the linear function which maps luma values to chroma values with respect to a center point determined by the average luma value of the reference samples. The encoder selectable adjustment is proposed to be an integer value between -4 and 4 in 1/8th sample units and is signalled at the PU level for each CCLM model which uses both top and left reference samples. It is reported that the proposed modification to ECM provides -0.05%, -1.10% and -0.98% BD-rate impact in AI configuration for Y, U and V, respectively. In RA configuration the impact is reportedly -0.01%, -0.44% and -0.48% for Y, U and V, respectively. Impact on encoder and decoder runtimes is asserted not significant.

CCLM slope parameter can be adjusted at PU level, which is combined with an adjustment w.r.t. local mean. Minimal impact w.r.t. processing, interesting gain in chroma.

Investigate in EE.

[JVET-Y0062](https://jvet-experts.org/doc_end_user/current_document.php?id=11256) Non-EE2: Cross-component palette coding [B. Vishwanath, K. Zhang, L. Zhang (Bytedance)]

In this contribution, cross-component palette coding (CC-PLT) is proposed. CC-PLT employs a lookup table to record the corresponding chroma sample value, given a luma sample value. The lookup table is built based on neighboring reconstructed samples from multiple reference lines. For non 4:4:4 sequences, a multi-filter approach is employed to derive co-located luma values.

On top of ECM-3.1, simulation results are reported as:

Class TGM: AI: -1.37%, -1.08%, -1.68%, 105%, 103%; RA: -0.34%, - 0.35%, -0.56%, 103%, 101%

Class F: AI: -0.20%, -0.44%, -0.36%, 105%, 102%; RA: -0.23%, - 0.35%, -0.19%, 104%, 101%

Study in EE along with the IBC proposals. Establish a new category in EE dedicatd to screen content coding.

[JVET-Y0076](https://jvet-experts.org/doc_end_user/current_document.php?id=11270) Non-EE2: Template Matching-based OBMC Design [Z. Lv, C. Zhou, J. Zhang (vivo)]

This contribution presents a method of template matching-based OBMC design, where the prediction value of CU boundary samples derivation approach is decided according to the template matching costs. Instead of directly using the weighted prediction, the proposed method chooses from three approaches for each block, including using current block’s motion information only, or using neighboring block’s motion information as well with one of the two blending modes.

On top of ECM-3.1, the simulation results of the proposed method are reported as below:

Test 1. RA: -0.13%, -0.18%, -0.10%, 114%, 118%.

Test 2. RA: -0.13%, -0.21%, -0.16%, 100%, 102%.

The RA results above are not obtained by correct CTC, but new incomplete results with CTC show practically same performance

The template matching used here is somewhat different from other approaches in ECM, may have some commonality with the template used in AMVR

It is asked if a similar effect (establishing rules whether to perform OBMC or not, method of blending, etc.) really needs template matching, or could be determined by other criteria.

Generally asserted as having interesting aspects.

Investigate Test 2 in EE

[JVET-Y0164](https://jvet-experts.org/doc_end_user/current_document.php?id=11371) Crosscheck of JVET-Y0076 (Non-EE2: Template Matching-based OBMC Design) [Y.-J. Chang (Qualcomm)] [late]

[JVET-Y0089](https://jvet-experts.org/doc_end_user/current_document.php?id=11283) Non-EE2: DMVR with BCW enabled [P. Bordes, A. Robert, Y. Chen, F. Galpin (InterDigital)]

This contribution proposes enabling BCW with DMVR. It is reported the simulation results of the proposed methods on top of the ECM-3.1 is {-0.03%, -0.01%, -0.04%} for YUV in RA configuration and the complexity (encoding/decoding times) is unchanged.

Simple modification, though it does not give relevant benefit in compression it might be interpreted as resolving a conflict between DMVR and BCW.

Decision: Adopt JVET-Y0089

[JVET-Y0218](https://jvet-experts.org/doc_end_user/current_document.php?id=11427) Cross-check of JVET-Y0089 (Non-EE2: DMVR with BCW enabled) [Z. Zhang (Qualcomm)] [late]

[JVET-Y0091](https://jvet-experts.org/doc_end_user/current_document.php?id=11285) Non-EE2: MVP refinement for regular AMVP mode [C. Zhou, Z. Lv, J. Zhang (vivo)]

This contribution presents a method of MVP refinement for regular AMVP mode, in which before using template-based searching to refine the MVP, two changes have been made. First, two out of five candidates are chosen from the candidate list by comparing the template matching costs. Second for each of the two selected candidates, template-based searching process is operated with a search pattern indicated by MMVD to derive an initial refined MVP with minimum template matching cost. Then the template-based searching process in ECM is conducted to the initial refined MVP. On top of ECM-3.1, the simulation results of the proposed method are reported in Test 1.

To reduce the complexity, two more tests with simplified methods were conducted and the performance are shown as Test 2 and Test 3 as below:

Under CTC

Test 1: RA

Test 2: RA

Test 3: RA

Non-CTC (without using per-class configuration provided in the cfg/per-class folder for random access condition)

Test 2: RA: -0.12%, -0.22%, -0.26%, 10x% 10x%

Non-CTC (and partially available CTC) results indicate 0.1% coding efficiency, but encoder run time is significantly increased (up to 10%), and also decoder run time increases due to additional template matching. No good tradeoff at this point – further study is recommended to improve.

[JVET-Y0210](https://jvet-experts.org/doc_end_user/current_document.php?id=11417) Crosscheck of JVET-Y0091: Non-EE2: MVP refinement for regular AMVP mode [J. Chen (Alibaba)] [late]

[JVET-Y0092](https://jvet-experts.org/doc_end_user/current_document.php?id=11286) Non-EE2: On chroma intra prediction mode [X. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba)]

In this contribution, two aspects of chroma intra prediction modification are proposed. For the first aspect, a new chroma intra prediction mode that uses the DIMD derivation method to derive the chroma intra prediction mode based on the collocated reconstructed luma samples is proposed. For the second aspect, it is proposed that a non-LM mode can be fused with a LM mode. The non-LM mode can be one of the proposed DIMD chroma mode, DM and four default chroma intra modes. It is reported that on top of ECM-3.1, when both two aspects are enabled, the overall coding performance impact for {Y, U, V, EncT, DecT} is {-0.11%, -1.49%, -1.39%, 105%, 100%} for AI and {0.00%, -0.71%, -0.62%, 102%, 99%} for RA. When only the proposed DIMD chroma mode can be fused with the LM mode in non-I slices, the overall coding performance impact for {Y, U, V, EncT, DecT} is {-0.03%%, -0.57%, -0.47%, 101%, 99%} for RA.

Study in EE. In this context, the benefit of the two elements should also be studied separately. Also, a complexity analysis should be performed, and potential to reduce processing should be explored.

[JVET-Y0221](https://jvet-experts.org/doc_end_user/current_document.php?id=11430) Cross-check of JVET-Y0092 (Non-EE2: On chroma intra prediction mode) [R. G. Youvalari (Nokia)] [late]

[JVET-Y0095](https://jvet-experts.org/doc_end_user/current_document.php?id=11289) Non-EE2: RPR with luma-only re-scaling [P. Bordes, F. Galpin, E. François (InterDigital)]

This contribution proposes extending RPR feature of picture size adaptation to chroma format adaptation. Doing so, one can re-scale luma only but not chroma. It reports this luma-only RPR provides BD-rate gains while chroma qp-offset alignment is more consistent across sequences and configurations.

Chroma is not downsampled, a QP+6 offset is applied to chroma. Alternatively, QP can be decreased when chroma is downsampled, but the method without downsampling is slightly better.

Luma-only downsampling is not possible with current VVC syntax, as RPR is always applied identical to all three components, and chroma format cannot be changed from 4:2:0 to 4:4:4.

It is pointed out that, as no dynamic resolution change is applied in the experiment, the same might be just achieved by pre and post processing. It is not truly “reference picture resampling”. If it is done at sequence level, this could likewise be done with high-level syntax. In this case, the decoder would need to support both 4:2:0 and 4:4:4.

Could this be done at block level? Might be difficult.

Generally interesting, further study beneficial. Could also be used in scalability and dynamic RPR.

The method of decreasing chroma QP should be studied in EE1 for the superresolution experiments.

[JVET-Y0168](https://jvet-experts.org/doc_end_user/current_document.php?id=11375) Cross-check of JVET-Y0095 "Non-EE2: RPR with luma-only re-scaling" [F. Le Léannec (Xiaomi)] [late]

[JVET-Y0097](https://jvet-experts.org/doc_end_user/current_document.php?id=11291) AhG12: Removed DIMD from MPM list of TIMD [K. Naser, T. Dumas, Y. Chen, F. Galpin (InterDigital)]

In ECM software, DIMD is required to performed prior to TIMD in the decoding process in order to fill in the TIMD-MPM list. It is argued that this process is associated with high complexity at the decoder side. This contribution proposes removing DIMD modes from the TIMD MPM list while keeping them in the default MPM list.

On top of ECM-3.1, the following results are obtained:

AI: 0,01% -0,03% 0,03% EncT 100% DecT 97%

RA: 0,01% 0,00% 0,07% EncT 100% DecT 100%

LDB: 0,01% -0,03% -0,08% EncT 100% DecT 100%

It is noted that run times may not be reliable. The effect may not be large.

Generally a good idea, but at this stage of exploration, such optimizations are not of high importance.

[JVET-Y0222](https://jvet-experts.org/doc_end_user/current_document.php?id=11431) Cross-check of JVET-Y0097 (AhG12: Removed DIMD from MPM list of TIMD) [K. Cao (Qualcomm)] [late]

[JVET-Y0109](https://jvet-experts.org/doc_end_user/current_document.php?id=11303) AHG12: Neural Network-based intra prediction [T. Dumas, F. Galpin, P. Bordes, F. Mom, E. François (InterDigital)]

This contribution proposes a new intra prediction mode based on neural network algorithm in ECM. It is based on the algorithm already tested in EE1 and reported in JVET-X0118 but contains algorithmic improvements for further complexity reduction. Following results are reported:

* RA: -0.71%, -0.16% and -0.23% BD-rate PSNR Y, U and V gains, respectively for a complexity of 104% and 111% at encoder and decoder,
* AI: -1.77%, -1.33% and -1.13% BD-rate PSNR Y, U and V gains, respectively, for a complexity of 136% and 330% at encoder and decoder,

It is proposed to include this tool in the next ECM version.

Compared to EE1, the model was further sparsified to reduce runtime (“unstructured” sparsity). Full integer implementation.

Worst complexity case for 4x4 blocks, 6.2 kMAC/pixel.

How is dependency of other tools from intra modes (e.g. LFNST, MTS) handled? A corresponding intra prediction mode is inferred.

Has training been disclosed/cross-checked?

Concern was expressed about complexity (in terms of number of worst-case computations, in terms of the dependency from reconstruction of neighbored blocks, and relative high increase of decoder runtime in AI).

Further investigation in the context of EE1, where EE1 establishes a part which exercises NN based tools on top of ECM (as was already the case in the previous EE1).

[JVET-Y0211](https://jvet-experts.org/doc_end_user/current_document.php?id=11418) Crosscheck of JVET-Y0109 (AHG12: Neural Network-based intra prediction) [Y. Li (Bytedance)] [late]

[JVET-Y0114](https://jvet-experts.org/doc_end_user/current_document.php?id=11308) Non-EE2: Dependent quantization with 4 states for chroma components [Y. Chen, E. François, F. Galpin, P. de Lagrange (InterDigital)]

This contribution proposes to apply dependent quantization with 4 states for chroma components in stead of using the current dependent quantization with 8 states in ECM-3.1, keeping luma component using dependent quantization with 8 states. It aims to reduce the implementation complexity for the encoder.

The following average results are reported for dependent quantization with 4 states for chroma components relative to ECM-3.1 (dependent quantization with 8 states for chroma components) under common test conditions:

AI: 0.03%, -0.12%, -0.11% (Y, Cb, Cr) at 98% encoding time and 100% decoding time;  
 RA: 0.02%, -0.09%, -0.09% (Y, Cb, Cr) at 99% encoding time and 100% decoding time;  
 LB: 0.00%, -0.23%, -0.23% (Y, Cb, Cr) at 96% encoding time and 100% decoding time.

The effect in encoder run time is negligible, and the effect is mainly some very small shift of quality from luma to chroma. Main intent of the contribution is to show that 8 states are not really necessary for chroma.

ECM supports both 4 and 8 states, can be selected by encoder but always identical for luma and chroma.

In the current stage of exploration, this is not important.

[JVET-Y0170](https://jvet-experts.org/doc_end_user/current_document.php?id=11377) Crosscheck of JVET-Y0114 (Non-EE2: Dependent quantization with 4 states for chroma components) [T. Lu (Dolby)] [late]

[JVET-Y0124](https://jvet-experts.org/doc_end_user/current_document.php?id=11318) Non-EE2: Intra Block Copy with An Extended Reference Area [J. Xu (ByteDance)]

This contribution presents IBC with an extended reference area. By extending the reference area to cover more CTUs that have been used by other coding modes in the current and above CTU rows, the average BD-rate(Y) numbers are –2.94% for classF and –5.88% for classTGM, with EncT change of 104% for classF and 99% for classTGM; DecT change of 95% for classF and 87% for classTGM in All Intra configurations.

It is claimed that more line buffers are available in ECM which allows extending the reference area. However, only intra template matching uses as many additional lines as this proposal (full CTU row plus 1). Other tools such as TIMD/DIMD only use 5 lines.

Investigate in EE, also in combination with other IBC related proposals.

[JVET-Y0220](https://jvet-experts.org/doc_end_user/current_document.php?id=11429) Cross-check of JVET-Y0160 (EE2-3.13-related: Enlarged HMVP table for IBC) combination with JVET-Y0124 (Non-EE2: Intra Block Copy with An Extended Reference Area) [A. Robert (InterDigital)] [late]

[JVET-Y0125](https://jvet-experts.org/doc_end_user/current_document.php?id=11319) AHG12: Enhanced bi-directional motion compensation [Y.-W. Chen, C.-W. Kuo, N. Yan, W. Chen, X. Xiu, X. Wang (Kwai Inc.)]

In current ECM, the bi-directional motion compensated (MC) predictor is generated by weighted average of two uni-directional MC predictors. In the proposed scheme, the uni-directional MC predictor is regarded as out-of-boundary (OOB) when the location of the reference interpolation sample is located outside the reference picture beyond half sample. For each predictor in a bi-directional MC block, when one uni-directional MC predictor is OOB and the other one is non-OOB, this bi-directional predictor is changed into the non-OOB uni-directional MC predictor because the OOB predictor is less effective.

Compared to the ECM-3.1 anchors, simulation results reportedly show that the proposed enhancements provide average {Y, U, V}BD-rate changes of {-0.22%, -0.15%, -0.15%} and {-0.04%, -0.01%, -0.06%} for RA and LDB configurations, respectively. The corresponding enc/dec run-time are 102%/103% for RA and 101%/100% for LDB. It is noted that SIMD is not used in the current implementation, but the run time could be further optimized after SIMD implementation is applied to the proposed scheme.

It is also noted that, for class D, simulation results reportedly show that the proposed enhancements provide average {Y, U, V}BD-rate changes of {-0.81%, -0.41%, -0.38%} and {-0.15%, -0.19%, -0.59%} for RA and LDB configurations, respectively. The corresponding enc/dec run-time are 102%/104% for RA and 103%/104% for LDB.

Several experts expressed this is interesting and simple. The decision is done on line/column currently, i.e. part of the block may be predicted bidirectional and part unidirectional. Done in the very last step, after DMVR etc.

Cross-checker confirms the result and further reports that they got slightly higher gain with an own implementation (which may be different by some details).

Investigate in EE.

[JVET-Y0219](https://jvet-experts.org/doc_end_user/current_document.php?id=11428) Cross-check of JVET-Y0125 (AHG12: Enhanced bi-directional motion compensation) [H. Huang, Y.-J. Chang, C.-C Chen, M. Karczewicz, V. Seregin, Y. Zhang, Z. Zhang (Qualcomm)] [late]

[JVET-Y0249](https://jvet-experts.org/doc_end_user/current_document.php?id=11458) Crosscheck of JVET-Y0219 (Cross-check of JVET-Y0125 (AHG12: Enhanced bi-directional motion compensation)) [H.-J. Jhu, Y.-W. Chen (Kwai)] [late]

[JVET-Y0128](https://jvet-experts.org/doc_end_user/current_document.php?id=11322) Non-EE2: fixing issues for RPR enabling and non-CTC configuration in ECM [Z. Zhang, H. Huang, C.-C. Chen, V. Seregin, M. Karczewicz (Qualcomm), P. Bordes (InterDigital)]

This contribution proposes slice level enabling condition check for adaptive DMVR and BM AMVP-merge mode. When the enabling condition is not fulfilled, the flag of adaptive DMVR and/or BM AMVP-merge mode is implicitly determined to be 0. The proposed method solves software issues with non-CTC configuration (e.g., RPR enabling). This contribution also proposes fixes for LIC and TM to enable RPR in ECM-3.1.

The proposed methods were implemented on top of ECM3.1 with merge request 47 (MR47). The simulation is conducted with RPR enabled. The test results confirm the software works properly.

It is suggested that LIC is not used and LIC block-level flag is not sent when the ref pic has a different size from the current picture. This allows usage of RPR, avoiding a mismatch..

This is not an encoder-only change and does not affect CTC. It seems to be straightforward fix to allow enabling RPR.

Decision(BF): Adopt JVET-Y0128

[JVET-Y0247](https://jvet-experts.org/doc_end_user/current_document.php?id=11456) Cross-check of JVET-Y0128: Non-EE2: fixing issues for RPR enabling and non-CTC configuration in ECM [K. Andersson (Ericsson)] [late]

[JVET-Y0129](https://jvet-experts.org/doc_end_user/current_document.php?id=11323) Non-EE2: MVD and merge index signaling of AMVP-merge mode [Z. Zhang, H. Huang, C.-C. Chen, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes to not signal MVD when AMVP-merge mode is used for the current block and template matching is enabled. It also proposes to add one pair of AMVP-merge MVPs. The proposed method was implemented on top of ECM-3.1, and it reports the BD-rate -0.08% (Y), -0.14 (U), -0.12% (V) for random access configuration with 100% encoding and 100% decoding run time compared to ECM3.1.

Cross-checker expresses opinion this is straightforward and simple.

Without adding the pair of candidates, gain would be approximately half.

Decision: Adopt JVET-Y0129

[JVET-Y0201](https://jvet-experts.org/doc_end_user/current_document.php?id=11408) Crosscheck of JVET-Y0129 (Non-EE2: MVD and merge index signaling of AMVP-merge mode) [P. Bordes (InterDigital)] [late]

[JVET-Y0130](https://jvet-experts.org/doc_end_user/current_document.php?id=11324) EE2-related: Unification of availability check for intra mode coding [S. Yoo, H. Jang, J. Nam, J. Choi, J. Lim, S. Kim (LGE)]

In this proposal, the unification of usage for neighbor intra modes above the CTU boundary is suggested. Currently, in ECM, the neighbor intra modes can be used in MPM and TIMD process. In VVC, neighboring intra modes above the CTU boundary are not exploited for MPM process to avoid the necessity of storing the intra modes in line buffer memory. However, in ECM, the above mentioned restriction still remains in some parts, but above neighbor intra modes are used as the candidates of the current block for some processes, such as TIMD and in the case of the intra mode propagation from the inter blocks (IPM). Therefore, it is proposed to unify the usage for neighboring intra modes above the CTU boundary. Two unified methods are introduced. In Method 1, the above intra modes in the different CTUs are always restricted as VVC. Method 2 allows to use the intra modes or intra modes from the IPM buffer from the above, above-right, or above-left neighbor blocks, regardless of the CTU boundaries. The experimental results are shown below.

* Method 1 (to restrict intra modes above the CTU boundary)
  + AI: 0.01% / -0.05% / -0.05% for Y/U/V
  + RA: 0.01 % / 0.04% / 0.11% for Y/U/V
* Method 2 (no restriction for intra modes above the CTU boundary)
  + AI: -0.01% / -0.01% / -0.03% for Y/U/V
  + RA: 0.00% / 0.04% / 0.09% for Y/U/V

The method would save one line buffer for storing intra modes (one per 4x4 block, 6 bits).

The benefit is not significant, and this level of fine-tuning is not so important at this stage.

No action.

[JVET-Y0191](https://jvet-experts.org/doc_end_user/current_document.php?id=11398) Crosscheck of JVET-Y0130: “EE2-related: Unification of availability check for intra mode coding” [M. Salehifar (Bytedance)] [late]

[JVET-Y0131](https://jvet-experts.org/doc_end_user/current_document.php?id=11325) EE2-related: Clean-up on DIMD [S. Yoo, H. Jang, J. Nam, J. Choi, J. Lim, S. Kim (LGE)]

In this document, it is proposed for a DIMD block to be blended with planar mode if one intra mode is derived. When DIMD is applied, one or two intra modes are derived from the reconstructed neighbor samples. If two intra modes are derived then those two predictors are combined (i.e., blended) with the planar mode predictor. However, when only one intra mode is derived, the predictor is not combined with planar mode predictor. Hence, it turns out to be the same as a conventional intra prediction mode but the mode is coded in DIMD mode. Therefore, it is proposed to always combine planar mode in DIMD mode. In addition, the default DIMD intra mode is set to DC mode rather than Planar. The experimental results are reportedly shown as following:

* Method 1 (to always combine planar mode in DIMD)
  + AI: 0.00 % / -0.05% / -0.02% for Y/U/V
  + RA: 0.00% / 0.05% / 0.04% for Y/U/V
* Method 2 (Method 1 + DC as default DIMD intra mode)
  + AI: 0.00 % / -0.04% / 0.03% for Y/U/V
  + RA: 0.00% / 0.00% / 0.08% for Y/U/V

No obvious benefit – no action.

[JVET-Y0192](https://jvet-experts.org/doc_end_user/current_document.php?id=11399) Crosscheck of JVET-Y0131 (EE2-related: Clean-up on DIMD) [H.-J. Jhu (Kwai)] [late]

[JVET-Y0135](https://jvet-experts.org/doc_end_user/current_document.php?id=11329) Non-EE2: 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)]

This contribution reports the result of a reordering method for GPM split modes using template matching (TM) cost. The method computes the TM cost of each GPM split mode with the GPM motion information of the current block, and then reorders all GPM split modes in ascending order through their respective TM costs. An index is signaled using Golomb-Rice code to indicate the use of split mode in the list of reordered GPM split modes. Compared with ECM-3.1, it is reported that the contribution achieves BD-rate saving of {0.21% (Y), 0.27% (U), 0.24% (V)} for Random Access and {0.29% (Y), 0.31% (U), 0.25% (V)} for Low-delay B.

1-2% increase in encoding/decoding time. Still in the range of reasonable tradeoff, though not huge coding gain.

Investigate in EE

[JVET-Y0232](https://jvet-experts.org/doc_end_user/current_document.php?id=11441) Crosscheck of JVET-Y0135 (Non-EE2: Template matching based reordering for GPM split modes) [W. Chen (Kwai)] [late]

[JVET-Y0139](https://jvet-experts.org/doc_end_user/current_document.php?id=11333) Non-EE2: On the extended number of active reference pictures and reference picture reordering [H. Huang, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution reports the result of extending the number of active reference pictures in random access configurations. The averaged BD-rate performance is -0.18%, -0.19%, -0.23%. This contribution also proposes a block level reference picture reordering method. Combined with the extended number of active reference pictures, it reports -0.36%, -0.33%, -0.41% BD-rate in random access configurations. Under the common test conditions, -0.15%, -0.11%, -0.12% BD-rate in random access configurations and -0.24%, -0.19%, -0.41% BD-rate in low delay B configurations are reported.

Reordering is done at block level. Signalling is changed. Currently not applied to AMVP merge, but should be straightforward to do.

Encoder run-time is increased, where the numbers given in the contribution (which seem to suggest runtime reduction) are unreliable. Cross-checker reports increase of runtime by approx. 1%.

Additional results show also sloightly more benefit when combined with JVET-Y0126 (changed GOP structure).

Investigate in EE

[JVET-Y0198](https://jvet-experts.org/doc_end_user/current_document.php?id=11405) Crosscheck of JVET-Y0139 (Non-EE2: On the extended number of active reference pictures and reference picture reordering) [R.-L. Liao (Alibaba)] [late]

[JVET-Y0144](https://jvet-experts.org/doc_end_user/current_document.php?id=11338) Non-EE2: DIMD Flag Signalling Clean-up [J. Zhao, S. Kim (LGE)]

In ECM3.1, the position of dimd\_flag is separated from other intra mode related syntax elements, and it unnecessarily complicates the software and signalling logic. This contribution proposes to change the position of dimd\_flag, and place it right before timd\_flag. It is asserted that this change makes the code and signalling logic cleaner, and consistent with timd\_flag signalling.

This change has little impact to coding performance. For class A to E, on average, impact on RD results are -0.01% -0.03%, 0.00% for AI, and 0.01%, 0.06%, 0.10% for RA configurations, no change for LDB.

For class F, the impact is slightly more noticeable, -0.04%, 0.11%, -0.11% for AI, -0.02%, 0.00%, -0.28% for RA and -0.10%, -1.13%, -0.52% for LDB.

Benefit not obvious – no action.

[JVET-Y0179](https://jvet-experts.org/doc_end_user/current_document.php?id=11386) Crosscheck of JVET-Y0144 (Non-EE2: DIMD Flag Signalling Clean-up) [K. Cao (Qualcomm)] [late]

[JVET-Y0148](https://jvet-experts.org/doc_end_user/current_document.php?id=11342) Non-EE2: Spatial-Temporal Adaptive Loop Filter [W. Yin, K. Zhang, Y. Li, H. Liu, L. Zhang (Bytedance)]

In the current ALF design of ECM, the online-trained filters for Luma component only have one fixed filter shape only with the spatial neighbouring samples inside the current picture for filtering. In this contribution, a spatial-temporal adaptive loop filter (ST-ALF) is proposed. With ST-ALF, a sample is filtered with both its spatial and temporal neighbouring samples.

On top of ECM-3.1, simulation results (BD-rate changes, encoding time and decoding time) of ST-ALF are reported as below:

AI: 0.00%, 0.00%, 0.00%, 101%, 100%;

RA: -0.10%, 0.09%, 0.07%, 110%, 102%;

LB: %, %, %, %, %;

Temporal filtering is applied including the current picture plus collocated blocks from reference picture(s). Filter shape is changed.

Relatively small gain is not reasonable tradeoff with complexity. Significant increase in encoding time, and additional access to reference pictures during loop filter operation.

Current status of proposal not attractive enough for inclusion in EE.

[JVET-Y0151](https://jvet-experts.org/doc_end_user/current_document.php?id=11345) Non-EE2: Adaptive re-ordering of merge candidates with refined motion [Y. Wang, K. Zhang, N. Zhang, Z. Deng, L. Zhang (Bytedance)]

This contribution proposes to refine motion information before invoking the adaptive re-ordering of merge candidates (ARMC) process. Specifically, instead of re-ordering the merge candidates using the initial motion information obtained from spatial/temporal/pairwise candidates, template matching (TM) and multi-pass decoder motion vector refinement (DMVR) are applied first to derive the refined motion of each merge candidate. Then the refined motion is used as inputs to the ARMC process. Two methods are tested as below.

Method #1: When multi-pass DMVR is used to derive the refined motion, only the first pass (i.e., PU level) is applied in reordering.

Method #2: Based on Method #1, when template matching is used to derive the refined motion, the template size is set equal to 1.

Compared to ECM-3.1, simulation results of the proposed method are reported as below:

Method #1:

RA: {-0.20%, -0.26%, -0.24%; 103%, 101%};

LDB: {-0.09%, -0.03%, -0.13%; 108%, 100%}.

Method #2:

RA: {-0.15%, -0.21%, -0.16%; 103%, 99%};

LDB: {-0.07%, 0.03%, -0.05%; 107%, 97%}.

Method 2 simplifies method 1 by using smaller size in the first pass.

Tradeoff for LDB not attractive.

Investigate in EE, but possibility to reduce complexity / encoder runtime should be investigated additionally

[JVET-Y0154](https://jvet-experts.org/doc_end_user/current_document.php?id=11348) AHG12: Bilinear Interpolation Filtering for ARMC [W. Chen, X. Xiu, H. Gao, Y.-W. Chen, X. Wang (Kwai)]

In this contribution, the bilinear interpolation filtering is proposed to replace the 12-tap based interpolation filtering used in adaptive reordering of merge candidates with template matching (ARMC) in ECM3 [1]. With the proposal, the 2-tap bilinear filtering is applied in all the cases where ARMC is applied, including regular merge mode, TM merge mode, and affine merge mode. Compared to the ECM-3.1 anchor [3], the simulation results are shown as below:

RA: 0.04% (Y), 0.01% (U), 0.09% (V), 99% (EncT), 97% (DecT)

LB: 0.04% (Y), -0.05% (U), -0.10% (V), 99% (EncT), 96% (DecT)

4-tap and 6-tap filters are also investigated, which have a loss of 0.01% in RA.

Benefit in terms of encoder runtime reduction almost negligible, and at this stge of exploration it is not of high importance to fine-tune tools for implementation friendliness.

No action.

[JVET-Y0231](https://jvet-experts.org/doc_end_user/current_document.php?id=11440) Crosscheck of JVET-Y0154 (AHG12: Bilinear Interpolation Filtering for ARMC) [Yan Zhang (Qualcomm)] [late]

[JVET-Y0157](https://jvet-experts.org/doc_end_user/current_document.php?id=11351) AHG12: Improved probability estimation for CABAC [X. Xiu, Y.-W. Chen, W. Chen, H. Gao, H.-J. Jhu, C.-W. Kuo, X. Wang (Kwai)]

In this contribution, two modifications are proposed to further improve the probability estimation accuracy of the CABAC in ECM-3.1. Firstly, one multi-hypothesis probability estimation with adaptive weight is proposed where the probability of each context is estimated based on the weighted combination of its two probability estimates instead of simple averaging. Secondly, one adaptive context initialization is proposed where the context states in one inter slice can be initialized in either of two ways: 1) being initialized by using one of the existing context tables; 2) being initialized by inheriting the context states of one slice that is previously coded. Compared to the ECM-3.1 anchors, simulation results reportedly show that the proposed modifications provide average BD-rate changes of -0.10%, -0.35% and -0.57% for AI, RA and LD configurations, without encoding/decoding run-time increase.

Inherited context probabilities (from previous picture) gives -0.24% in RA and -0.38 in LB

Multihypothesis context gives about 0.1% (for RA and AI). Adaptive weight (in context initialization, which are fixed for a given slice type) can also be inherited across pictures

The inherited context would not allow parallel parsing across pictures, this could also affect parallel processing within picture (wavefront can only be started when finished for last picture)

Investigate in EE. Multihypothesis should also be tested without the change in adaptation ratio.

[JVET-Y0213](https://jvet-experts.org/doc_end_user/current_document.php?id=11421) Cross-check of JVET-Y0157 (AHG12: Improved probability estimation for CABAC)) [J. Zhao (LGE)] [late]

[JVET-Y0172](https://jvet-experts.org/doc_end_user/current_document.php?id=11379) Non-EE2: Long tap interpolation filtering on chroma components [X. Xie, K. Zhang, L. Zhang, J. Li, M. Wang, S. Wang (Bytedance)] [late]

This contribution proposes to replace the 4-tap interpolation filter on chroma components in motion compensation with a 12-tap interpolation filter. On top of ECM-3.1, experimental results reportedly show {-0.04%, -1.59%, -1.70%} and {-0.11%, -2.10%, -1.78%} BD-rate changes on average for {Y, Cb, Cr}, under RA and LDB configurations, respectively.

Both encoding and decoding time increases by 6%/13% for RA (decoding time less increased in LB)

As chroma is subsampled, it should use shorter interpolation filters.

Has it been investigated if the long filters might cause visible artifacts at colorful sharp edges?

One expert points out that long interpolation filters in chroma could cause problem in subjective quality.

Investigate in EE. In particular, also consider using less extended filters (e.g., 6 taps), and investigate possible quality impact by subjective viewing.

[JVET-Y0181](https://jvet-experts.org/doc_end_user/current_document.php?id=11388) AHG12: CABAC initialization from previous inter slice [V. Seregin, J. Dong, N. Hu, M. Karczewicz (Qualcomm)] [late]

This contribution first reports test results for CABAC initialization for inter slices by using probabilities stored from the center CTU of the previous slice having the same type and slice QP. This contribution also proposes a method to adjusting the window size for each context model, such that the probability estimation adapts better to the local statistics. Experimental results show that this method working together with temporal CABAC reduces BD-rate -0.09% AI, -0.29% RA, and -0.32%\* LDB.

Similar to JVET-Y0157 in terms of inheritance, but position from previous slice/picture where the inheritance comes from is different. This might be preferable for parallelism.

Inheritance is performed from a slice with same slice-level QP. Hypothetically, many tables need to be stored for each QP and slice type.

Initialization from another slice in the same picture is not used. Slices should be independently decodable.

It is pointed out that the LUT for adaptation of window size is relatively large.

Run times are partially unreliable – the method should not have an effect on runtime

Has it been tested with temporal scalability?

Investigate in EE along with JVET-Y0157.

[JVET-Y0233](https://jvet-experts.org/doc_end_user/current_document.php?id=11442) Crosscheck report of JVET-Y0181 (AHG12: CABAC initialization from previous inter slice) [K. Sato (OPPO)] [late]

# High-level syntax (HLS) proposals (12)

## AHG9: SEI message studies and proposals (9)

Contributions in this area were discussed in session 12 at 2330–0145+1 UTC on Friday 14 Jan. 2022, in session 14 at 1600-1730 UTC on Monday 17 Jan. 2022, and in session 18 at 0050-0125 UTC on Wednesday 19 Jan. 2022 (chaired by JRO).

[JVET-Y0044](https://jvet-experts.org/doc_end_user/current_document.php?id=11236) AHG9: Signalling of Green metadata and Video Decoding Interface SEI messages in VVC specification [E. François (InterDigital), Y. He (Qualcomm), C. Herglotz (FAU), Y. Lim (Samsung)]

This contribution proposes an update of the VVC specification for signalling the green metadata and video decoding interface (VDI) SEI messages, as specified in ISO/IEC 23001-11 and ISO/IEC 23090-13, respectively.

Decision: Adopt JVET-Y0044

Output document: Draft 1 of VVC extensions

Plan from last meeting was to have an amendment for these SEI hooks finalized (FDAM) by January 2023, which would mean that CDAM should be started in April 2022, DAM in July 2022. It was confimed by the SC29 secretary that, starting from DAM stage, a specification can be normatively referenced.

It was planned to issue a request for this amendment by WG 5, and issue a WD (as JVET and WG 5 output)

[JVET-Y0050](https://jvet-experts.org/doc_end_user/current_document.php?id=11242) AHG2/AHG9: On the alpha channel information SEI message [Y.-K. Wang (Bytedance)]

See notes in section 4.2

[JVET-Y0053](https://jvet-experts.org/doc_end_user/current_document.php?id=11247) AHG9/AHG13: Film grain blending process for film grain characteristics SEI message [Y. He, M. Coban, M. Karczewicz (Qualcomm)]

See section 6.2

[JVET-Y0073](https://jvet-experts.org/doc_end_user/current_document.php?id=11267) AHG9: Colour component description for post-filter purpose SEI message [T. Chujoh, Y. Yasugi, K. Takada, T. Ikai (Sharp)]

At the previous meeting, the SEI message for sending network parameters of neural network and the SEI message for purposes of post-filter was proposed by JVET-X0112. Multiple companies have presented the same concept at the past several meetings. This contribution suggests maintaining the two-layered SEI structure and proposes that the post-filter purpose SEI message specifies the input or output sample interface for the post-filtering process. For example, it should describe the updating colour components and the output chroma format.

High level interface (Y0073) specifies the input and output chroma sampling formats of the post filter

Neural network interface – Y0074

Neural network complexity – Y0075

[JVET-Y0074](https://jvet-experts.org/doc_end_user/current_document.php?id=11268) AHG9: Data conversion description for NNR post-filter SEI message [Y. Yasugi, T. Chujoh, K. Takada, T. Ikai (Sharp)]

This contribution proposes a post-filter SEI message based on JVET-X0112 to include input and output tensor format information of neural network post-filter and conversion process description. It is argued that the neural network models (input and output shape) do not have association information, i.e., which tensor element corresponds to which component of samples to apply the post-filter. It is also argued that even if NNC carries topology information by either NNR compression, PyTorch, TensorFlow, NNEF, or ONNX specification, the input and output tensor format information is fully dependent on each specific specification and not trivial. It is asserted that thus the format of an input/output tensor shall be specified.

Is the input/output of NNR compression adequately defined? Such an SEI message would need to refer to the corresponding specification.

Should the precision (and possibly precision conversion), e.g. integer bit depth, floating point, be also specified?

Could input and output sizes be different, e.g. to support upsampling?

[JVET-Y0075](https://jvet-experts.org/doc_end_user/current_document.php?id=11269) AHG9: Complexity description for NNR post-filter SEI message [K. Takada, Y. Yasugi, T. Chujoh, T. Ikai (Sharp)]

This contribution proposes a NNR post-filter SEI message based on JVET-X0112. This proposal suggests including syntax element for information of complexity and precision to decide applicability on the decoder side without having to analyze network information specified by the URI or payload in the SEI.

What is the maximum number of parameters? Syntax element nnrpf\_num\_parameters\_idc is 8 bits, and MaxNumParameters = (2048 << nnrpf\_num\_parameters\_idc) – 1. Minimum is 2048.

[JVET-Y0103](https://jvet-experts.org/doc_end_user/current_document.php?id=11297) AHG9: Down-sample phase indication (SEI message) [P. Bordes, P. de Lagrange, E. François (InterDigital)]

Down-sampling pictures before video coding is a tool commonly used for seamless adapting bit-streams to available bandwidth. For example, this feature is supported by MPEG-DASH clients that switches between representations of different resolutions and by VVC that implements per picture resolution adaptation (RPR).

To ensure that the receivers and players appropriately display (up-sample) the decoded video, they need to know the phases that have been used when creating the down-sampled video.

This contribution proposes a new SEI message indicating the phases and other information necessary to correctly interpret the reconstructed video for resampling to the original size. This SEI message can apply to AVC, HEVC and VVC video streams and it may overcome some limitations of previous proposal JVET-X0092 discussed in last JVET meeting and on the reflector.

Comments from the discussion:

* Semantics of subsampling phase unclear, 3 bits should be sufficient for 8 phases
* Semantics of chroma position indicators unclear; are all relevant cases covered? Why is it necessary signalling that for the downsampled case, where it is available in VUI? How do the “collocated” flags relate to common chroma formats? Examples?
* Reference to pps\_id could be improved
* Relationship with VUI to be clarified
* Somehow the corresponding (up- or down) scaling following that SEI should be described, and how the information about the original picture would be used in that process
* Persistence to be clarified – can it be sent per picture?
* Could original picture width/height be an odd number
* Use cases to be made more clear

Further improvements and study of possible application scenarios recommended.

JVET-Y0156 is related.

[JVET-Y0104](https://jvet-experts.org/doc_end_user/current_document.php?id=11298) AHG9: Transparency information SEI for transparent screens [E. Thomas, P. Andrivon, F. Le Léannec, 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, a SEI signaling is presented which is asserted to address the above requirement and enable the transformation of the decoded pictures after decoding for both legacy and new content video content.

Was reviewed in session 18.

Values of mask (0..255) are encoded in the SEI message itself. Simple run-length coding can be used optionally. Mask could be sent for every picture.

It is suggested better using an auxiliary picture, and just signal its properties in an SEI message rather than signaling the mask values in the SEI message. With the proposed method, the bit rate for the mask might become higher than for the video. Such an SEI message should be simple to define, very similar to alpha channels.

How close are transparent displays in terms of introducing services specifically for those?

Generally an interesting domain, typical use case where an SEI message is beneficial. Suggesting an updated version is recommended.

[JVET-Y0107](https://jvet-experts.org/doc_end_user/current_document.php?id=11301) AHG9: Text improvement for the film grain SEI [E. Thomas (Xiaomi)]

It is asserted that the description text and formula of the film grain SEI message may be improved for the reader of the specification. The text is written with the assumption of a 4:4:4 coded sequence while there is no normative constrain of that nature when using this SEI. Hence, using equations with variables such as PicHeightInLumaSamples and c for the different components tends to hide the prior assumption made at the beginning of the clause.

A possible the text modification is proposed to make more explicit the 4:4:4 throughout the equations.

Revision 1: After offline discussion, it was clarified that there was no bug in the specification text but that the readability of the text may be improved.

Decision(ed.): Adopt JVET-Y0107 for V2 of VSEI

Generally it is agreed that editorial changes would help understanding the SEI message.

In the discussion, it is however agreed to keep Pic<Width|Height>InLumaSamples unchanged, as this is commonly used in VVC/HEVC. Further, I-prime should better become I-hat to denote the image that is input to the process. That image should not be called “decoded image” but better “input image” (which may have been converted from the decoded picture).

The above intent to be implemented by editors.

[JVET-Y0115](https://jvet-experts.org/doc_end_user/current_document.php?id=11309) AHG9: On post-filter SEI [M. M. Hannuksela, M. Santamaria, F. Cricri, E. B. Aksu, H. R. Tavakoli (Nokia)]

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 contribution is a follow-up of JVET-X0112.

JVET-X0112 proposed two SEI messages:

1. An NNR post-filter SEI message for the carriage of an MPEG Neural Network Representation (NNR, ISO/IEC 15938-17) bitstream that specifies a neural-network-based post-filter.
2. A post-filter purpose SEI message that currently specifies quality enhancement and super resolution as purposes and allows other purposes to be specified in the future.

The following bullet list copies the comments on JVET-X0112 from the JVET meeting minutes and summarizes how this contribution responds to the comments.

1. Is it really necessary to have two separate messages? The PfP message does not contain much info, and likely both would be always updated together. A single SEI message might be more appropriate.

Response: The post-filter purpose (PfP) SEI message is no longer proposed in this contribution. The purpose indicator is included in the NNR post-filter SEI message.

1. The purpose of PfP is not fully clear: Where would the actual method e.g. of super resolution be defined, and is it at all necessary that the encoder describes it? The resolution after upsampling might finally be display specific, so is it necessary to signal it at all? (it is answered that e.g. the NNR post filter could be used for that purpose)

Response: In this contribution, only a single purpose value (visual quality improvement) is proposed, whereas the super-resolution purpose is no longer proposed in this contribution. A new value for the purpose indicator for super-resolution can be specified in the future. Conditioned by this value, syntax elements related to super resolution can also be added in the future, if needed.

1. Is the NNR specification complete e.g. in describing input/output ordering w.r.t. the video decoder output? Could it e.g. describe a 4:2:0 input and 4:4:4 output?

Response: An input/output ordering indicator is included in the NNR post-filter SEI message proposed in this contribution. A single value of the indicator is specified, and its use requires the 4:2:0 chroma format. The input and output ordering are specified with pseudo code in a manner that suits the post-filter tested in JVET-Y0059 [8]. Other values of the input/output ordering indicator can be specified in the future.

1. How would block-wise processing, block-overlapping, and boundary processing be managed?

Response: The block-wise processing, block overlapping, and boundary processing are specified together with input and output ordering with pseudo code. Sample locations that would be outside the boundaries of the cropped decoded output picture are clipped to be within the boundaries.

1. What could be the maximum amount of payload data in NNR? It is also asked if referring to a relative complex standard is the purpose of a SEI message? For the currently foreseen purposes, it might not be updated frequently. There could also be other systems-related mechanisms.

Response: SEI messages do not have a size limit. In the experiments by the proponent, the NNR-coded data has been up to 2134 bytes.

1. What does URI mean, and what does it refer to?

Response: A cross-reference to URI has been added (IETF Internet Standard 66).

1. Extensibility aspect might require more consideration. In the past, extensibility was not often used, even if it had been foreseen in an SEI message.

Response: This contribution suggests that new values of the purpose indicator and an input/output ordering indicator can be specified in the future.

For the example in JVET-Y0059, the uri would point to the network itself, and the number of 2134 bytes given above would be the sequence-dependent update thereof.

Why 21 bits for nnrpf\_id?

Can line-wise processing be done? Currently not.

Is NNR supporting update of networks? Current NNR allows overwriting weights an biases, and incremental NNR as currently being specified allows differential encoding.

In case when the input and output have same properties in terms of size, color sampling etc., this SEI message (which operates patch-wise) would just take over the picture format parameters from the decoder output. Otherwise, it could be combined with JVET-Y0073 and JVET-Y0074. JVET-Y0075 also has a NNR payload in it, but has more information about complexity.

Would be better to combine everything into one SEI message. Also, the relation with NNR (e.g., for an update made on a network) seems to require more clarification. Ideally, the SEI would just contain an NNR stream, and the interface with the input and output layers of the decoded network.

Further study recommended, merging the different contributions, and also to investigate how this SEI message can be generalized beyond the application of contribution JVET-Y0059.

It is also suggested to consider (in terms of the payload size) the coding efficiency, e.g. for superresolution.

[JVET-Y0156](https://jvet-experts.org/doc_end_user/current_document.php?id=11350) AHG9: SEI message with sample phase indication for consistent rendering [F. Bossen, A. Segall (Sharp)]

An SEI message containing a sample phase indication is proposed. It is asserted to enable more consistent rendering, in particular when rendering at a higher resolution and when switching between multiple resolutions as is common in streaming scenarios where bandwidth may fluctuate.

Same purpose as JVET-Y0103.

Questions:

* Would it be useful to define denominator separately for horizontal and vertical?
* How about chroma? Is it enough to have the chroma position from VUI, and what happens if it is unspecified?
* Would it be necessary to signal a different offset depending on the position of the cropping window?

Further study with common examples would be beneficial to understand what elements are needed in such an SEI message. Contributors of Y0156 and Y0103 are asked to study such cases and report how the elements of their proposed SEI messages can be used in a scaling method to avoid the phase shift problem. It would not be advisable to define two separate SEI messages.

It is reported verbally that updates of the two documents were uploaded which contain such examples. As there is no urgency of defining such an SEI message immediately, it was suggested to have more discussion at the next meeting.

## Film Grain Synthesis (3)

Contributions in this area were discussed in session 19 at 1550–XXXX UTC on Wednesday 19 Jan. 2022 (chaired by JRO).

[JVET-Y0053](https://jvet-experts.org/doc_end_user/current_document.php?id=11247) AHG9/AHG13: Film grain blending process for film grain characteristics SEI message [Y. He, M. Coban, M. Karczewicz (Qualcomm)]

This contribution proposes a film grain blending process for film grain characteristics SEI message. The blending process is identical to the process specified in SMPTE RDD5 except the modifications to reduce line buffers and on-chip memory storage.

The contribution proposes two independent modifications:

1. The grain block size is changed from fixed 8x8-size in RDD5 to 8x1 for picture resolutions less than or equal to 1920x1080, 16x1 for resolution greater than 1920x1080 and up to 3840x2160, and 32x1 for resolutions greater than 4K including 8K. Using Nx1 block size eliminates the need of line buffers to compute block average in the existing methods.
2. Add note to limit the number of film grain patterns per picture to 1 to reduce on-chip memory requirement.

Purpose of second constraint is to reduce the on-chip memory from 40 to 4 kByte.

Some concern is expressed that the proposal may be constraining too much for more generic applicability, RGB, other than 4:2:0 color sampling, etc. Evaluation based on 1080p might not be sufficient to identify it is sufficient.

The Cross-checker is also sceptical, in particular about the modification 2. It is mentioned that the line-wise processing of modification 1 is only related to blending, the patterns can still be 2D.

Modification 1 is asserted to be uncritical in terms of subjective quality.

Should an SEI message include such constraints? Likely not. However, if such a restriction is deemed applicable and useful, the TR could express that, but then clearly mention where possible limits are.

It was suggested to discuss and study this further in the context of the AHG. It is however not practical to perform extensive tests about possible implications on subjective quality. It should, for example, be discussed if the report could point out ways to simplify the synthesis with regard to hardware complexity. If experts have an opinion that such simplifications might not be appropriate for certain type of content, this should also be expressed.

[JVET-Y0165](https://jvet-experts.org/doc_end_user/current_document.php?id=11372) Crosscheck of JVET-Y0053 (AHG9/AHG13: Film grain blending process for film grain characteristics SEI message) [M. Radosavljević (InterDigital)] [late]

[JVET-Y0158](https://jvet-experts.org/doc_end_user/current_document.php?id=11352) AHG13: Draft Film Grain Technical Report Text [W. Husak, A. Tourapis, W. Wan, M. Radosavljević, D. Grois] [late]

This document contains the draft text for the Technical Report on Film Grain Technologies for Video Compression Applications. This is a working title and is subject to change.

The proposed text was presented. It contains description of different methods, including frequency-based and autoregressive, and the associated analysis/synthesis processes.

The purpose of applying this in the context of video coding is described at high level in the introduction. Key application that is described is adding film grain to hide compression artifacts. QP dependency on that?

Better title is needed. Avoid “Metadata”. “recommendations” and “guidance” should be avoided in a TR.

Way of referencing AV1 would need to be carefully expressed. Should not be termed as “major XXX”. Main focus should be about expressing the differences and commonalities in the technology, e.g. that the autoregressive model in the SEI is different from the one of AV1.

There is no implementation available for the autoregressive model in the SEI message, even though by the time of its proposal (in JVT) it might have been presented.

For example, Documents JVT-I013 and JVT-H022 could be used as references how the autoregressive model in the SEI message could be used, and then, with reference to the AV1 model, it could be described that the models are conceptually similar but not compatible, even though to some extent it might be possible to generate similar noise characteristics when mapping one model into another.

It is also mentioned that an external reference to the technology of AV1 is currently evolving (in an external organization), developing a kind of SEI message for it that could be references better than AV1. Timeline for the TR might also be made dependent on the availability of such a spec.

Two options:

* Not having an output, wait until next meeting
* Produce a WD from the current meeting, but only include the description of the AVC/HEVC/VSEI based SEI message (or usage in systems), and associated analysis synthesis, reference to RDD-5, potentially old JVT docs, etc., leave out reference AV1 for now. Could have editing notes that it is intended to add reference to other existing technology.

It is agreed to move forward with option 2. Editing period could be 5 weeks.

“Film grain technology for video applications”

“Grain-noise technology for video applications”

## Non-SEI HLS aspects (0)

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

# Plenary meetings, joint meetings, BoG reports, and liaison communications

## JVET plenaries

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
* Whether to hold a hybrid meeting in April 2022
* Review docs from section 4.x

Fri. 21 Jan. 1300-XXXX:

* See notes under sections 8,9,10, and 11.

## Information sharing meetings

In addition to the joint meetings listed below, information sharing sessions with other WGs of the MPEG community were held on Monday 17 Jan. 0500–XXXX, Wednesday 19 Jan. 0500–0600, and Friday 21 Jan. 2100–2300. The status of the work in the MPEG WGs was reviewed at these information sharing sessions.

## Joint meeting with … XXXX–XXXX on XXday X January

The following topics were discussed in this joint session. See also the notes recorded on these topics in other sections of this document.

## BoGs (3)

The following break-out groups were established at this meeting to conduct discussion and develop recommendations on specific topics.

## Liaison communications

The JVET did receive the following liaison related communication at its current meeting.

[m58614](https://dms.mpeg.expert/doc_end_user/current_document.php?id=81344&id_meeting=189) Summary of voting on Request, Liaison Representative (5th WG 5 mtg) [SC 29 Secretariat]

TBP

Follow-up on the request to appoint Mathias Wien as WG 5 Liaison Representative to ITU-T SG 12.

Update of Liaison representatives doc?

[m59114](https://dms.mpeg.expert/doc_end_user/current_document.php?id=81844&id_meeting=189) Response liaison letter from SC 29/WG 01 to WG 5 (JVET) on JPEG AI [WG 1 via SC 29 Secretariat]

TBP

# Project planning

## Software timeline

ECM4 software (including all adoptions) was planned to be available 3 weeks after the meeting.

VTM16 software was planned to be available on 2022-02-18. (Note that updates 15.1/15.2 are planned to be released shortly after the meeting)

HM16.25 software was planned to be available on 2022-02-18.

## Core experiment and exploration experiment planning

An EE on neural network-based video coding was established, as recorded in output document JVET-X2023.

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-X2024.

Initial versions of these documents were presented and approved in the session 23 on Thursday 14 October.

## 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, we would like 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)).

Review of AHG plans was conducted during the closing plenary on Friday 21 January 2022.

|  |  |  |
| --- | --- | --- |
| **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-X2005 and JVET-X2006). * 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-X1004 errata output collection. * Produce and finalize JVET-X2002 VVC Test Model 15 (VTM 15) 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, J. Chen, 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 in relation to video test materials, including new test materials. * Identify and recommend appropriate test materials for testing the VVC standard and potential future extensions, as well as exploration activities. * Identify 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. * Prepare availability of viewing equipment and facilities arrangements for future meetings. | V. Baroncini, T. Suzuki, M. Wien (co-chairs), E. François, S. Liu, A. Norkin, A. Segall, P. Topiwala, S. Wenger, Y. Ye (vice-chairs) | N |
| **Conformance testing (AHG5)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the JVET-X2026 draft conformance testing for operation rage extensions and investigate the need for improvements. * Study the JVET-X2008 draft conformance testing specification and investigate the need for extensions. * 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-3.0 software version and the reference configuration encodings according to JVET-X2017 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 key application scenarios (e.g. cloud gaming) and their requirements for low latency and constrained complexity, taking into account aspects of real-time encoding and decoding. * Evaluate and propose new CTC for low latency and constrained complexity application scenarios. * Conduct tests with ECM and VTM to determine the impact of discussed configurations on coding efficiency and run time. * 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. | 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 lossless coding characteristics of VVC. * Identify technologies for future extension of VVC to support such application usage. * Study the JVET-U2018 testing conditions for high bit depth, high bit rate, and high frame rate coding, and suggest improvements as applicable. * Contribute to the development of software and conformance testing for operation range extensions in coordination with AHG3 and AHG5. * Study the draft text JVET-X2005 and suggest improvements in coordination with AHG 2. * Identify suitable test material for testing of high bit depth, high bit rate, and high frame rate coding in coordination with AHG4. * Study VVC entropy decoding throughput and latency in the cases of high bit depth, high bit rate, and high frame rate coding. | A. Browne, T. Ikai (co-chairs), D. Rusanovskyy, M. Sarwer, 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. * Study the draft text JVET-X2006 and suggest improvements in coordination with AHG 2. * Study signalling of essential resampling phase indication and prepare draft text for such signalling. * 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, Ye-kui Wang (co-chairs), C. Fogg, P. de Lagrange, J. Samuelsson, G. J. Sullivan, A. Tourapis, Y.-K. Wang, 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 software configuration settings optimized for subjective quality, and coordinate such efforts with AHG3. | P. de Lagrange, 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 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. * Finalize, conduct and discuss the EE on neural network-based video coding. * Refine the test conditions for NN-based video coding. Generate and distribute anchor encoding, and develop supporting software as needed. * Investigate technical aspects specific to NN-based video coding, such as encoding and decoding complexity of neural networks, design network representation, operation, tensor, on-the-fly network adaption (e.g. updating during encoding) etc; * Study the impact of training (including the impact of loss function) on the performance of candidate technology, and identify suitable materials for training. * Analyse complexity characteristics, perform complexity analysis, and develop complexity reductions of candidate technology. * Refine testing methods for assessment of the effectiveness and complexity of considered technology. * Study the impact of parameter quantization and fixed-point computations in NN-based video coding. * Study and collect information related to near-term and longer-term architectures for neural network-based video coding. * Study and maintain the SADL (Small Adhoc Deep Learning Library). Identify gaps in library support and develop improvements as needed. * Review the outcome of the expert viewing conducted at the meeting, refine the methodology, and prepare viewing for the next meeting. * Coordinate with AHG12 to study the interaction with ECM coding tools. * Contribute to a workshop on NN based technologies with WG 1 and WG 4. * Coordinate with other relevant groups, including SC29/AG5 on visual quality assessment. | E. Alshina, S. Liu, A. Segall (co‑chairs), J. Chen, F. Galpin, J. Pfaff, S. S. Wang, Z. Wang, M. Wien, P. Wu, J. Xu (vice‑chairs) | Y (2 weeks notice) |
| **Enhanced compression beyond VVC capability (AHG12)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Solicit and study non-neural-network video coding tools with enhanced compression capabilities beyond VVC. * Discuss and propose refinements to the ECM3 algorithm description JVET-X2025. * Study the performance and complexity tradeoff of these video coding tools. * Coordinate with AHG6 on ECM software development. * Refine test conditions in JVET-X2017, generate anchors, identify new test sequences to be added, especially high resolution ones in 8K, in coordination with AHG4. * Analyse the results of exploration experiments described in JVET-X2024 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 existing FGC SEI messages in VSEI, HEVC, and AVC. * Study the benefits and characteristics of film grain technologies, including autoregressive and frequency-filtering technologies. * Study encoder technologies for determining values for FGC SEI message syntax elements. * Prepare draft text for a technical report on the use of the FGC SEI message providing examples of different technologies including examples of encoder-side and post-decode processing while considering existing implementations. * Identify potential need for additional film grain technology SEI messages. * Coordinate development of film grain technology software and configuration files. * Study and evaluate available test content (both SDR and HDR) for use in film grain technology development in coordination with AHG4. * Coordinate with AHG3 for software support of the FGC SEI message. | W. Husak, M. Radosavljević, W. Wan (co-chairs), D. Grois, A. Tourapis (vice-chairs) | N |

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 however 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 118](https://dms.mpeg.expert/doc_end_user/current_document.php?id=81008&id_meeting=188)) in order to make it easy to reference.

# Output documents

The following documents were agreed to be produced or endorsed as outputs of the meeting. Names recorded below indicate the editors responsible for the document production. Where applicable, dates of planned finalization and corresponding parent-body document numbers are also noted.

It was reminded that in cases where the JVET document is also made available as a WG 5 output document, a separate version under the WG 5 document header should be generated. This version should be sent to GJS and JRO for upload.

The list of JVET ad hoc groups was also issued as a WG 5 output document [WG 5 N 118](https://dms.mpeg.expert/doc_end_user/current_document.php?id=81008&id_meeting=188), as noted in section 9.

[JVET-Y1000](https://jvet-experts.org/doc_end_user/current_document.php?id=11222) Meeting Report of the 25th JVET Meeting [J.-R. Ohm] [WG 5 N 96] (2022-02-18)

Initial versions of the meeting notes (d0 … d9) were made available on a daily basis during the meeting.

Remains valid – not updated: [JCTVC-H1001](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=5095) HEVC software guidelines [K. Sühring, D. Flynn, F. Bossen (software coordinators)]

Remains valid – not updated: [JVET-Y1002](https://jvet-experts.org/doc_end_user/current_document.php?id=10846) 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 106] (2022-04-15)

Encoder optimization from JVET-Y0077

Remains valid – not updated: [JVET-T1003](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10535) Revised coding-independent code points for video signal type identification (Draft 2) [G. J. Sullivan, T. Suzuki, A. Tourapis]

[JVET-Y1004](https://jvet-experts.org/doc_end_user/current_document.php?id=11223) 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] (2022-04-15, near next meeting)

[JVET-Y1005](https://jvet-experts.org/doc_end_user/current_document.php?id=11224) New levels for HEVC (Draft 2) [T. Suzuki, A. Tourapis, Y.-K. Wang] [WG 5 CDAM N 102] (2022-02-11)

Also includes corrections related to HEVC from previous JVET-X1004.

A request for a new amendment ([WG 5 N 101](https://dms.mpeg.expert/doc_end_user/current_document.php?id=81002&id_meeting=188)) was reviewed XXday XX January in session XX.

Remains valid – not updated [JVET-T1006](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10538) Annotated regions and shutter interval information SEI messages for AVC (Draft 2) [J. Boyce, S. McCarthy, Y.-K. Wang]

This is included in a 10th ISO/IEC edition issued as WG 5 N 98 (due 2022-04-15). Twin text already published in ITU-T.

DoCR of DIS (WG5 N 97) was reviewed Thursday 20 January in session 24.

Remains valid – not updated: [JCTVC-V1007](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=10312) SHVC Test Model 11 (SHM 11) Introduction and Encoder Description [G. Barroux, J. Boyce, J. Chen, M. M. Hannuksela, Y. Ye] [WG 11 N 15778]

Remains valid – not updated: [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: [JCTVC-Z1020](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=10692) Common Test Conditions for HDR/WCG Video Coding Experiments [E. François, J. Sole, J. Ström, P. Yin]

[JVET-Y1100](https://jvet-experts.org/doc_end_user/current_document.php?id=10675) Common Test Conditions for HM Video Coding Experiments [K. Sühring, K. Sharman] (2022-02-25)

After merging CTC for HM and VTM into JVET-Y-2010, this needs to be modified to contain 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-Y2002](https://jvet-experts.org/doc_end_user/current_document.php?id=11225) Algorithm description for Versatile Video Coding and Test Model 16 (VTM 16) [A. Browne, J. Chen, Y. Ye, S. Kim] [WG 5 N 92] (2022-04-15, near next meeting)

New description elements: Encoder optimization from JVET-Y0077, JVET-Y0085, JVET-Y0105, JVET-Y0126

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-Y2005](https://jvet-experts.org/doc_end_user/current_document.php?id=11226) VVC operation range extensions (Draft 6) [F. Bossen, B. Bross, T. Ikai, D. Rusanovskyy, G. J. Sullivan, Y.-K. Wang] (2022-03-18)

Was integrated into a FDIS of new edition issued as WG 5 N 105 (due 2022-03-18), and submitted for ITU-T consent (initial version needed 2022-01-24).

DoCR on CDAM (WG 5 N 104) was reviewed Thursday 20 January in session 24.

[JVET-Y2006](https://jvet-experts.org/doc_end_user/current_document.php?id=11227) Additional SEI messages for VSEI (Draft 6) [J. Boyce, G. J. Sullivan, Y.-K. Wang] (2022-03-18)

Was integrated into a DIS of new edition issued as WG 5 N 100 (due 2022-03-18), and submitted for ITU-T consent (initial version needed 2022-01-24).

DoCR on CDAM (WG 5 N 99) was reviewed Thursday 20 Januay in session 24.

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] (2021-11-30)

[JVET-Y2009](https://jvet-experts.org/doc_end_user/current_document.php?id=10680) Reference software for versatile video coding (Draft 3) [F. Bossen, K. Sühring, X. Li] (2022-03-18)

Software related to v2 of VVC and VSEI was integrated.

Was issued as ISO/IEC FDIS 23090-16 as [WG 5 N 112](https://dms.mpeg.expert/doc_end_user/current_document.php?id=81001&id_meeting=188), and submitted for ITU-T consent.

A DoCR ([WG 5 N 111](https://dms.mpeg.expert/doc_end_user/current_document.php?id=81000&id_meeting=188)) of the NB comments received in [m57767](https://dms.mpeg.expert/doc_end_user/current_document.php?id=80226&id_meeting=188) from the ISO/IEC JTC 1/SC 29 DIS ballot was reviewed Thursday 20 January in session 24.

[JVET-Y2010](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10545) VTM common test conditions and software reference configurations for SDR video [F. Bossen, X. Li, V. Seregin, K. Sharman, K. Sühring] (2022-02-25)

New merged SDR conditions for HM and VTM, as per JVET-Y0112.

[JVET-Y2011](https://jvet-experts.org/doc_end_user/current_document.php?id=10851) VTM common test conditions and evaluation procedures for HDR/WCG video [A. Segall, E. François, W. Husak, S. Iwamura, D. Rusanovskyy] (2022-02-04)

Update aligned with ECM CTC.

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]

Remains valid – not updated: [JVET-X2016](https://jvet-experts.org/doc_end_user/current_document.php?id=11229) Common Test Conditions and evaluation procedures for neural network-based video coding technology [E. Alshina, R.-L. Liao, S. Liu, A. Segall]

[JVET-Y2017](https://jvet-experts.org/doc_end_user/current_document.php?id=11230) Common Test Conditions and evaluation procedures for enhanced compression tool testing [M. Karczewicz and Y. Ye] (2022-02-25)

Change to CatRobot, and add HDR as optional, and per-class configurations to be added.

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]

[JVET-Y2019](https://jvet-experts.org/doc_end_user/current_document.php?id=11229) New level and systems-related supplemental enhancement information for VVC (Draft 1) [E. François, A. Tourapis, Y.-K. Wang] [WG 5 WD N 108] (2022-03-18)

Hooks for green metadata and VDI; unconstrained level

A request for a new amendment ([WG 5 N 107](https://dms.mpeg.expert/doc_end_user/current_document.php?id=81002&id_meeting=188)) was reviewed XXday XX January in session XX.

[JVET-Y2020](https://jvet-experts.org/doc_end_user/current_document.php?id=11229) Film grain technology for video applications (Draft 1) [D. Grois, Y. He, W. Husak, M. Radosavljević, A. Tourapis, W. Wan] (2022-02-25)

No output: JVET-X2021, JVET-X2022

Numbers retained for future purposes of planning possible additional verification testing and CE.

[JVET-Y2023](https://jvet-experts.org/doc_end_user/current_document.php?id=11221) 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 113] (2022-02-04)

An initial draft of this document was reviewed and approved during session XX on XXday XX January.

Three categories of experiments were planned: Enhancement filters (loop and post), super resolution, and intra prediction. Except for 1.1. and 1.3, proposals from last meeting, more extensive cross-check to be conducted. Tests on top of ECM are also included.

[JVET-Y2024](https://jvet-experts.org/doc_end_user/current_document.php?id=11220) Exploration Experiment on Enhanced Compression beyond VVC capability (EE2) [V. Seregin, J. Chen, L. Li, K. Naser, J. Ström, M. Winken, X. Xiu, K. Zhang] [WG 5 N 114] (2022-02-18)

An initial draft was reviewed and approved during session 23 on Thursday 14 Oct. It was suggested to add a rule that encoder modifications that could also be applied to the anchors should also be implemented in a modified version of the anchor, and results reported.

[JVET-Y2025](https://jvet-experts.org/doc_end_user/current_document.php?id=11231) Algorithm description of Enhanced Compression Model 4 (ECM 4) [M. Coban, F. Le Léannec, K. Naser, J. Ström] [WG 5 N 115] (2022-03-04)

New elements:

* Adopt JVET-Y0116 version 2.1a (extension of MRL list to 5 lines 1,3,5,7,12)
* Adopt JVET-Y0065 Test 3.1c, not in LP CTC
* Adopt JVET-Y0134 Test 3.6a
* Adopt JVET-Y0067 Test 3.9d
* Adopt JVET-Y0058 Test 3.13
* Adopt JVET-Y0142 Test 4.4a (adaptive intra MTS with fixed threshold
* Adopt JVET-Y0106 (CCSAO improvement)
* Adopt JVET-Y0141 test 3
* Adopt JVET-Y0159 method 1
* Adopt JVET-Y0089 (DMVR with BCW)
* Adopt JVET-Y0128 (bug fix)
* Adopt JVET-Y0129 (MVD signalling)

It is noted that the list above may not be complete; if some adoption is missing that is recorded somewhere else in the meeting notes it shall also be considered included.

[JVET-Y2026](https://jvet-experts.org/doc_end_user/current_document.php?id=11232) Conformance testing for VVC operation range extensions (draft 3) [D. Rusanovskyy, T. Hashimoto, H.-J. Jhu, I. Moccagatta, M. G. Sarwer, Y. Yu] (2022-03-18)

The basis is document JVET-Y0127 which had been reviewed and approved during the meeting.

This was issued as ISO/IEC 23090-15 DAM 1 as [WG 5 N 110](https://dms.mpeg.expert/doc_end_user/current_document.php?id=81003&id_meeting=188) (with an editing period as noted).

A DoCR on CDAM ([WG 5 N 109](https://dms.mpeg.expert/doc_end_user/current_document.php?id=81002&id_meeting=188)) was reviewed Thursday 20 January in session 24.

# 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. 20 – Fri. 22 and Mon. 25 – Fri. 29 April 2022, 26th meeting under ISO/IEC JTC 1/‌SC 29 auspices, to be held as teleconference meeting.
* 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.
* During October 2022, 28th meeting under ITU-T SG16 auspices, date and location t.b.d.
* 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 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 in Geneva, CH, date 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.

The agreed document deadline for the 26th JVET meeting was planned to be Wednesday 13 April 2022.

Huawei was thanked for providing financial support for the VVC verification tests.

Mathias Wien was thanked for planning, organizing and conducting the remote expert viewings related to the exploration experiment on neural network-based video compression, and to deblocking filter improvements.

Apple and InterDigital were thanked for offering new test materials that can be used for developing and testing video technology standards.

Warmestv thanks were expressed to Christian Tulvan for his continuous support and personal engagement in maintaining the site jvet-experts.org.

The 25th JVET meeting was closed at approximately XXXX hours UTC on Saturday 22 January 2022.

# Annex A to JVET report: List of documents

# Annex B to JVET report: List of meeting participants

The participants of the twenty-fifth 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:



# Annex C to JVET report: Recommendations of the 6th 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 XX**](https://sd.iso.org/documents/ui/#!/browse/iso/iso-iec-jtc-1/iso-iec-jtc-1-sc-29/iso-iec-jtc-1-sc-29-wg-5/library/4/Recommendations%20of%20the%204th%20WG%205%20meeting)