|  |  |
| --- | --- |
| **Joint Video Experts Team (JVET)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29**  24th Meeting, by teleconference, 6–15 October 2021 | Document: JVET-X\_Notes\_d7 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Meeting Report of the 24th Meeting of the Joint Video Experts Team (JVET), by teleconference, 6–15 October 2021** | | |
| *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-fourth meeting during 6–15 October 2021 as an online-only meeting. It had previously been planned to be in Antalya, TR, 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 fifth meeting as WG 5. The JVET meeting was held under the chairmanship of Dr Jens-Rainer Ohm (RWTH Aachen/Germany). For rapid access to particular topics in this report, a subject categorization is found (with hyperlinks) in section 2.14 of this document. It is further noted that work items which had originally been conducted by the Joint Collaborative Team on Video Coding (JCT-VC) were continued in JVET as a single joint team, and explorations towards possible future need of standardization in the area of video coding are also conducted by JVET, as negotiated by the parent bodies.

The JVET meeting began at approximately 0500 hours UTC on Wednesday 6 October 2021. Meeting sessions were held on all days except the weekend days of Saturday and Sunday 9 and 10 October 2021, until the meeting was closed at approximately XXXX hours UTC on Friday 15 October 2021. Approximately 1 people attended the JVET meeting, and approximately XXX input documents (not counting crosschecks), 12 AHG reports, 3 CE/EE summary reports, and X BoG reports were discussed. The meeting took place in coordination with a meeting of various 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, under whose auspices this JVET meeting was held. The subject matter of the JVET meeting activities consisted of work on further development and maintenance of the twin-text video coding technology standards *Advanced Video Coding* (AVC), *High Efficiency Video Coding* (HEVC), *Versatile Video Coding* (VVC)*, Coding-independent Code Points (Video)* (CICP), and *Versatile Supplemental Enhancement Information Messages for Coded Video Bitstreams* (VSEI), as well as related technical reports, reference software and conformance testing packages. As a primary goal, the JVET meeting reviewed the work that was performed in the interim period since the twenty-third JVET meeting in producing the following documents:

* JVET-W1000 Meeting report
* JVET-W1004 Errata report items for VVC, HEVC, AVC, Video CICP, and CP usage TR
* JVET-W2002 Algorithm description for Versatile Video Coding and Test Model 14 (VTM 14)
* JVET-W2005 VVC operation range extensions (Draft 4)
* JVET-W2006 Additional SEI messages for VSEI (Draft 4)
* JVET-W2016 Common Test Conditions and evaluation procedures for neural network-based video coding technology
* JVET-W2017 Common Test Conditions and evaluation procedures for enhanced compression tool testing
* JVET-W2019 White paper on VVC
* JVET-W2020 VVC verification test report for HDR video content
* JVET-W2022 Core Experiment on film grain synthesis
* JVET-W2023 Exploration Experiment on neural network-based video coding (EE1)
* JVET-W2024 Exploration Experiment on enhanced compression beyond VVC capability (EE2)
* JVET-W2025 Algorithm description of Enhanced Compression Model 2 (ECM 2)
* JVET-W2026 Conformance testing for VVC operation range extensions (Draft 1)

Further important goals were reviewing the results of the CE on Film Grain Synthesis, 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.

The JVET produced XX output documents from the current meeting (update):

* JVET-W1004 Errata report items for VVC, HEVC, AVC, Video CICP, and CP usage TR
* JVET-W2002 Algorithm description for Versatile Video Coding and Test Model 14 (VTM 14)
* JVET-W2005 VVC operation range extensions (Draft 4)
* JVET-W2006 Additional SEI messages for VSEI (Draft 4)
* JVET-W2016 Common Test Conditions and evaluation procedures for neural network-based video coding technology
* JVET-W2017 Common Test Conditions and evaluation procedures for enhanced compression tool testing
* JVET-W2019 White paper on VVC
* JVET-W2020 VVC verification test report for HDR video content
* JVET-W2022 Core Experiment on film grain synthesis
* JVET-W2023 Exploration Experiment on neural network-based video coding (EE1)
* JVET-W2024 Exploration Experiment on enhanced compression beyond VVC capability (EE2)
* JVET-W2025 Algorithm description of Enhanced Compression Model 2 (ECM 2)
* JVET-W2026 Conformance testing for VVC operation range extensions (Draft 1)

For the organization and planning of its future work, the JVET established XX “ad hoc groups” (AHGs) to progress the work on particular subject areas. At this meeting, X Core Experiment (CE) and X Exploration Experiments (EE) were defined. The next eight JVET meetings were planned for Tue. 11 – Fri. 14 January and Mon. 17 – Wed. 19 January 2022 as a virtual meeting under ITU-T SG16 auspices; during Fri. 22 – Fri. 29 April 2022 under ISO/IEC JTC 1/‌SC 29 auspices in Alpbach, AT; during Fri. 15 – Fri. 22 July 2022 under ISO/IEC JTC 1/‌SC 29 auspices in Cologne, DE; during October 2022 under ITU-T SG16 auspices in Geneva, CH; during January 2023 under ISO/IEC JTC 1/‌SC 29 auspices, location t.b.d.; during April 2023 under ISO/IEC JTC 1/‌SC 29 auspices, location t.b.d.; during July 2023 under ITU-T SG16 auspices in Geneva, CH; and during October 2023 under ISO/IEC JTC 1/‌SC 29 auspices, 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-fourth meeting during 6–15 October 2021 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 fifth 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/2021_10_X_Virtual/>.

## Primary goals

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

* JVET-W1000 Meeting report
* JVET-W1004 Errata report items for VVC, HEVC, AVC, Video CICP, and CP usage TR
* JVET-W2002 Algorithm description for Versatile Video Coding and Test Model 14 (VTM 14)
* JVET-W2005 VVC operation range extensions (Draft 4)
* JVET-W2006 Additional SEI messages for VSEI (Draft 4)
* JVET-W2016 Common Test Conditions and evaluation procedures for neural network-based video coding technology
* JVET-W2017 Common Test Conditions and evaluation procedures for enhanced compression tool testing
* JVET-W2019 White paper on VVC
* JVET-W2020 VVC verification test report for HDR video content
* JVET-W2022 Core Experiment on film grain synthesis
* JVET-W2023 Exploration Experiment on neural network-based video coding (EE1)
* JVET-W2024 Exploration Experiment on enhanced compression beyond VVC capability (EE2)
* JVET-W2025 Algorithm description of Enhanced Compression Model 2 (ECM 2)
* JVET-W2026 Conformance testing for VVC operation range extensions (Draft 1)

Further important goals were reviewing the results of the CE on Film Grain Synthesis, 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 plan 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.

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 Thursday, 30 September 2021. Any documents uploaded after 2359 hours Paris/Geneva time on Thursday 30 September 2021 were considered “officially late”. The usual grace period of 12 hours (to accommodate those living in different time zones of the world) was not applied, as the deadline was already one day later than originally intended, due to a wrong expression of date in the previous meeting report. 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-X0146 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-X0XXX (a proposal on …), uploaded 10-XX.
* …

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

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

* JVET-X0XXX (a document on …), uploaded 10-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 contribution registrations were noted that were later cancelled, withdrawn, never provided, were cross-checks of a withdrawn contribution, or were registered in error: JVET-X0042, JVET-X0057, JVET-X0061, JVET-X0062.

“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-W1000, the Errata report items for VVC, VSEI, HEVC, AVC, Video CICP, and CP usage TR JVET-W1004, the Algorithm description for Versatile Video Coding and Test Model 14 (VTM 14) JVET-W2002, the Operation range extensions for VVC (Draft 4) JVET-W2005, the Additional SEI messages for VSEI (Draft 4) JVET-W2006, the Common Test Conditions and evaluation procedures for neural network-based video coding technology JVET-W2016, the Common Test Conditions and evaluation procedures for enhanced compression tool testing JVET-W2017, the White paper on VVC JVET-W2019, the VVC verification test report for HDR video content JVET-W2020, the Description of the CE on Film Grain Synthesis JVET-W2022, the Description of the EE on Neural Network-based Video Coding JVET-W2023, the Description of the EE on Enhanced Compression beyond VVC capability JVET-W2024, the Algorithm description of Enhanced Compression Model 2 (ECM 2) JVET-W2025, and the Conformance testing for VVC operation range extensions (Draft 1) JVET-W2026, had been completed and were approved. The software implementations of HM (version 16.24), VTM (version 14.0), ECM (version 2.0), and HDRTools (version 0.23) were also approved.

Only minor editorial issues were found in the meeting report JVET-W1000 – 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 core experiment on film grain synthesis
* 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 considerationThe plans for the times of meeting sessions were established as follows, in UTC (2 hours behind the time in Geneva, Paris; 7 hours ahead of the time in Los Angeles, etc.). No session should last longer than 2 hrs.

* 0500–0700 1st “morning” session [break after 2 hours]
* 0720–0920 2nd “morning” session
* [“midday” break – nearly 4 hours]
* 1300–1500 1st “afternoon” session [break after 2 hours]
* 1520–1720 2nd “afternoon” session

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 21st JVET meeting (which had been the fourth 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 1340 (as of 5 October 2021). All future discussions (including those on HEVC, VVC, CICP, etc.) shall be conducted on the JVET reflector rather than the JCT-VC reflector, while the old reflectors (including JVT, JCT-VC, and JCT-3V) are retained for archiving purposes.

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

## Terminology

* **ACT**: Adaptive colour transform
* **AFF**: Adaptive frame-field
* **AI**: All-intra
* **AIF**: Adaptive interpolation filtering
* **ALF**: Adaptive loop filter
* **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
* **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 6 October at 0500 UTC were as follows.

* Timing and organization of online meetings, calendar
* Standards approval and publication status
  + Working practices using objective metrics report
    - HSTP-VID-WPOM V1: 2020-07-03, published 2020-11
    - ISO/IEC TR 23002-8 (Ed. 1) published 2021-05-20
  + AVC
    - H.264v14 Consented at 22nd meeting (with annotated regions, shutter interval, and miscellaneous corrections), Last Call closed during the 23rd meeting, not published yet
    - ISO/IEC 14496-10:2020 (Ed. 9) FDIS closed 2020-11-27, published 2020-12-15 – corresponding aspects are partly in the in-progress CDAM (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), Last Call closed during the 23rd meeting, not published yet
    - 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
    - H.273v2 (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) Last Call closed during the 23rd meeting, not published yet
  + 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
* Draft standards progression status
  + VVC conformance – under DIS 23090-15 ballot closed 2021-09-29, FDIS from current meetng
    - [m57766](https://dms.mpeg.expert/doc_end_user/documents/136_OnLine/wg11/m57766-v1-m57766.zip) Summary of voting on ISO/IEC DIS 23090-15
  + VVC reference SW – under DIS 23090-16 ballot closed 2021-09-30, FDIS from current meeting
    - [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
  + AVC additional SEI – integrated into DIS of new edition at 23rd meeting, ballot closes XX-XX, FDIS in January
  + VSEI extensions – integrated into DIS of new edition at 23rd meeting, ballot closes XX-XX, FDIS in January, some inputs at this meeting, possibly new version JVET-X2006
  + VVC operation range extensions – integrated into DIS of new edition at 23rd meeting, ballot closes XX-XX, FDIS in January, some inputs at this meeting, possibly new version JVET-X2005
  + 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 14496-10:2020/Amd 1 – AVC to be done when finishing new edition
    - ISO/IEC 23008-2:2020/Amd 1 – HEVC done last meeting
    - ISO/IEC DIS 23091-2, 2nd edition – CICP done last meeting
    - ISO/IEC 23002-7:2021/Amd 1 – VSEI to be done when finishing new edition
    - ISO/IEC 23090-3:2021/Amd 1 – VVC to be done when finishing new edition
* The meeting logistics, agenda, working practices, policies, and document allocation were reviewed.
  + The meeting was conducted using Zoom.
  + Having text and software available is crucial (and not just arriving at the end of the meeting).
  + There were no objections voiced in the opening plenary to the consideration of late contributions.
* The results of the previous meeting and the meeting report JVET-W1000 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 section 2.6, it was stated that meeting hours of the 23rd meeting were identical to those of the 19th meeting, but actually they were matching with those of the 20th meeting
  + In the report JVET-W0005 of AHG5 (conformance), it was stated that FDIS 23090-15 would be an output from the 23rd meeting, but actually the DIS was still under ballot by that time.
  + In section 11 and Annex C 6.3.1, the document deadline for the current meeting was defined as “Wednesday 30 September”, but actually 2021-09-30 was a Thursday, Even though the intention might have been to have Wednesday 29 as the deadline, a deadline of Thursday 30 was later defined in the meeting announcement.
  + In Annex C 4, recommendations related to VVC are under wrong headers for HEVC (already noted in JVET-W1000).
* 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
  + FDIS on v1 of VVC conformance and software
  + Errata
  + Conformance and software for version2 of VVC & VSEI
  + Extensions of VVC (High bit rate / high bit depth)
  + Additional SEI messages for VSEI
  + Core Experiment on Film Grain Synthesis
  + 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.
* Liaison communication
* Number of documents slightly higher than last meeting (90->110)
* 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.

## Scheduling of discussions

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

* 0500–0700 1st “morning” session [break after 2 hours]
* 0720–0920 2nd “morning” session
* [“midday” break – nearly 4 hours]
* 1300–1500 1st “afternoon” session [break after 2 hours]
* 1520–1720+1 2nd “afternoon” session

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

* Wed. 06 Oct., 1st day
  + Session 1:
    - 0500–0540 Opening remarks, review of practices, agenda, IPR reminder
    - 0540–0700 Reports of AHGs 1–5
  + Session 2:
    - 0720–0810 Reports of AHGs 6–12
    - 0820-0920 Review of CE on film grain synthesis
  + Session 3:
    - 1300–1500 Review of EE1 and EE2
  + Session 4:
    - 1520–1720 Further review of EE2
* Thu. 7 Oct., 2nd day
  + Session 5:
    - 0500–0700 Review of v2 related items: 4.2, 4.12, 5.1
    - BoG on NN EE1 (A. Segall) meets in parallel
  + Session 6:
    - 0720–0920 Review of v2 related items: 4.2, 4.12, 5.1
    - BoG on NN EE1 (A. Segall) may continue to meet
  + Session 7:
    - 1300–1500 Further review of EE1 and EE1 related
  + Session 8:
    - 1520–1720 Further review of EE2 and EE2 related
* Fri. 8 Oct., 3rd day
  + Session 9:
    - 0500–0700 Further review of v2 related items: 4.2, 4.12, 5.1
    - BoG on NN EE1 (A. Segall) meets in parallel from 0600
  + Session 10:
    - 0720–0920 Further review of v2 related items: 4.2, 4.12, 5.1, SEI messages (6.1)
    - BoG on NN EE1 (A. Segall) meets in parallel
  + Session 11:
    - 1300–1500 Further review of EE1 and EE1 related
  + Session 12:
    - 1520–1720 Further review of EE2 related
* Mon. 11 Oct., 4th day
  + 0500–0730 MPEG information sharing session
  + Session 13:
    - 0745–0920 Coordination, planning, topics from section 4
  + Session 14:
    - 1300–1500 Remaining docs 5.2.3/5.2.4
  + Session 15:
    - 1520–1615 Remaining docs 5.3.3
  + Joint meeting with AG5/WG4/WG7:
    - 1620–1720 Guidelines for verification testing and remote experts viewing
* Tue. 12 Oct., 5th day
  + 0500-0600 VCEG
  + 0500-0700 BoG on v1/v2 conformance (I. Moccagatta/D. Rusanovskyy)
  + Session 16:
    - 0600–0700 Remaining docs 5.3.4
  + Session 17:
    - 0720–0920 SEI messages, VVC/VSEI v2 revisits
  + 1300-1400 EE1 experts viewing
  + Session 18:
    - 1300–1500 Remaining docs 5.3.4/5.2.4
* Wed. 13 Oct., 6th day
  + 0500–0600 MPEG information sharing session
  + Session 19:
    - 0620–0920 Remaining docs 5.3.4
  + Session 20:
    - 1300–1500 4.11, 4.4, 4.6
  + Session 21:
    - 1520–1720 NN REV results, remaining docs 4.6, BoG reports, revisit on CE, further planning
* Thu. 14 Oct., 7th day
  + Session 22:
    - 0500–0700 Revisits, remaining business
  + Session 23:
    - 1300–1500 Review of DoCs, EE descriptions and other output documents
* Fri. 15 Oct., 8th day
  + Plenary:
    - 0500–0700 Remaining business, Approval of output docs, AHGs, recommendations
  + Plenary:
    - 0720–0920 Remaining business, Approval of output docs, AHGs, recommendations
  + 2100–2300 MPEG information sharing session
  + XXXX–XXXX WG 5 Closing plenary: Approval of 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 (section 3)
* Project development (section 4)
  + Deployment and advertisement of standards (2)
  + Text development and errata reporting (6)
  + Test conditions (1)
  + Verification testing (1)
  + Test Material (0)
  + Quality assessment (4)
  + Conformance test development (2)
  + Software development (0)
  + Implementation studies (2)
  + Complexity analysis (0)
  + Encoding algorithm optimization (3)
  + Profile/tier/level specification (7)
  + 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 (5) (section 5.1)
  + AHG11 and EE1: Neural network-based technology (29) (section 5.2)
  + AHG12 and EE2: Enhanced compression beyond VVC capability (52) (section 5.3)
* High-level syntax (HLS) proposals (section 6) with subtopics
  + AHG9: SEI message studies and proposals (4) (section 6.1)
  + CE on film grain synthesis (2) (section 6.2)
  + Non-SEI HLS aspects (0) (section 6.3)
* Joint meetings, plenary discussions, BoG and viewing reports (3), 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 (12)

These reports were discussed Wednesday 6 Oct. 2021 during 0540–0700 and 0720–0810 UTC (chaired by JRO).

[JVET-X0001](https://jvet-experts.org/doc_end_user/current_document.php?id=11146) JVET AHG report: Project management (AHG1) [J.-R. Ohm, G. J. Sullivan]

At the 23rd meeting of the ITU-T/ISO/IEC Joint Video Experts Team (JVET), an ad hoc group on Project Management was established with the following mandates:

* Coordinate overall JVET interim efforts.
* Supervise AHG studies.
* Report on project status to JVET reflector.
* Provide a report to next meeting on project coordination status.

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

The number of subscribers (by the beginning of the current meeting) was 1340. Furthermore, the JCT-VC email list (which was still kept open) had 1294 subscribers.

**Goals and activity**

The work of the JVET overall had proceeded well in the interim period with approximately same 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_07_W_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-W1000) [Posted 2021-08-14]
* Errata report items for VVC, HEVC, AVC, Video CICP, and CP usage TR (JVET-W1004) [Posted 2021-09-28]
* Algorithm description for Versatile Video Coding and Test Model 14 (VTM 14) (JVET-W2002) [Posted 2021-09-25]
* VVC operation range extensions (Draft 4 (JVET-W2005) [Posted 2021-08-24, last update 2021-09-29]
* Additional SEI messages for VSEI (Draft 4) (JVET-W2006) [Posted 2021-05-07]
* Common Test Conditions and evaluation procedures for neural network-based video coding technology (JVET-W2016) [Posted 2021-05-26, last update 2021-08-16]
* Common Test Conditions and evaluation procedures for enhanced compression tool testing (JVET-W2017) [Posted 2021-07-31]
* White paper on Versatile Video Coding (JVET-W2019) [Posted 2021-10-04]
* VVC verification test report for high dynamic range video content (JVET-W2020) [Posted 2021-08-27]
* CE on Film Grain Synthesis (JVET-W2022) [Posted 2021-07-16, last update 2021-07-30]
* EE on Neural Network-based Video Coding (EE1) (JVET-W2023) [Posted 2021-07-16, last update 2021-08-02]
* EE on Enhanced Compression beyond VVC capability (EE2) (JVET-W2024) [Posted 2021-07-15, last update 2021-09-23]
* Algorithm description of Enhanced Compression Model 2 (ECM 2) (JVET-W2025) [Posted 2021-09-02]
* Conformance testing for VVC operation range extensions (draft 1) (JVET-W2026) [Posted 2021-07-16, last update 2021-09-01]

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

* Disposition of comments on ISO/IEC 14496-10:2020 CDAM 1 Additional SEI messages (WG 5 N 65)
* Text of ISO/IEC DIS 14496-10:202X Advanced Video Coding (10th edition) (WG 5 N 66)
* Disposition of comments on ISO/IEC 23002-7:2021 CDAM 1 Additional SEI messages (WG 5 N 67)
* Text of ISO/IEC DIS 23002-7:202X Versatile supplemental enhancement information messages for coded video bitstreams (2nd edition) (WG 5 N 68)
* Disposition of comments on ISO/IEC 23090-3:2021 CDAM 1 Operation range extensions (WG 5 N 69)
* Text of ISO/IEC DIS 23090-3:202X Versatile Video Coding (2nd edition) (WG 5 N 70)

The twelve *ad hoc* groups had made progress, and reports from those activities had been submitted. Furthermore, one core experiment (CE) on film grain synthesis, and 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 verification 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 110 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 24th meeting had been made publicly available on the ITU-hosted ftp site <http://wftp3.itu.int/av-arch/jvet-site/2021_10_X_Virtual/>.

The AHG recommends adding a mandate on supervising and organizing the processing and delivery of output documents.

[JVET-X0002](https://jvet-experts.org/doc_end_user/current_document.php?id=11147) 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]

This document reports on the work of the JVET ad hoc group on Draft text and test model algorithm description editing (AHG2) between the 23rd meeting by teleconference, (7 – 16 July 2021) and the 24th meeting by teleconference, (6 – 15 October 2021).

1. **Mandates**

At the 23rd meeting of the ITU-T/ISO/IEC Joint Video Experts Team (JVET), an ad hoc group on Draft text and test model algorithm description editing (AHG2) was established with the following mandates:

* Produce and finalize draft text outputs of the meeting (JVET-W2005 and JVET-W2006).
* 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-W1004 errata output collection.
* Produce and finalize JVET-W2002 VVC Test Model 14 (VTM 14) 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

1. **Ad hoc group activity**

**Output documents produced**

**JVET-W2005 VVC operation range extensions (Draft 4) and the corresponding MPEG output document MDS20614\_WG05\_N00070 Text of ISO/IEC DIS 23090-3:202X Versatile Video Coding (2nd edition)**

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 the SEI manifest and SEI prefix indication SEI messages, including SEI payload type values and other interfaces for SEI messages added to the VSEI specification, as well as some technical corrections.

Draft 4 incorporated items:

* Method of entropy coding for high bit depth (JVET-W0046).
* High bit depth profiles (JVET-W0136v5).
* HLS cleanup proposal 3 in JVET-W0070.
* Proposal 2 case 1 of JVET-W0109.
* Constrains of option 3 of JVET-W0178.
* Editorial changes from WG5 N0070.

**JVET-W2006 Additional SEI messages for VSEI (Draft 4) and the corresponding MPEG output document MDS20610\_WG05\_N00068 Text of ISO/IEC DIS 23002-7:202X Versatile supplemental enhancement information messages for coded video bitstreams (2nd edition)**

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 4 incorporated items:

* JVET-W0077 AHG9: Comments on multiview-related SEI messages in VSEI
* JVET-W0078 AHG9: Multiview view position SEI message
* JVET-W0080 AHG9: Some errata and clarification items for Additional SEI messages for VSEI
* JVET-W0083 AHG9: Bug fixes for some SEI messages in the VSEI amendment
* DoCR items not covered by any of the above

**JVET-W1004 Errata report items for VVC, 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-W meeting:

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

**JVET-W2002 Algorithm description for Versatile Video Coding and Test Model 14 (VTM 14)**

The JVET established the VVC Test Model 14 (VTM14) algorithm description and encoding method at its 23rd meeting (7 – 16 July. 2021, teleconference). This document serves as a source of general tutorial information on the VVC design and also provides an encoder-side description of VTM14. 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 14 (VTM14) algorithm description and encoding method

* Incorporated JVET-V0047: CE-3.1 and CE-3.2: Transform coefficients range extension for high bit-depth coding
* Incorporated JVET-V0054: CE-2.1: Slice based Rice parameter selection for transform skip residual coding
* Incorporated JVET-V0106: CE-related: On history-enhanced method of Rice parameter derivation for regular residual coding (RRC) at high bit depths
* Incorporated JVET-W0046: CE-1.1: coding of last significant coefficient position for high bit depth and high bit rate extensions
* Incorporated JVET-W0136: Suggested initial profile text for VVC operation range extension

1. **Related input contributions**

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

* JVET-X0050: AHG2: On editing notes for the 2nd edition draft text of VVC.
* JVET-X0059: AHG2/AHG9: Comments on the 2nd edition draft text for VSEI
* JVET-X0073: AHG2: On specifying the range extensions profiles
* JVET-X0075: AHG8: Level refinement for VVC operation range extension profiles
* JVET-X0079: Proposals on maximum bit rate for HEVC and VVC
* JVET-X0093: AHG2/AHG8: Comments on VVC operation range extensions
* JVET-X0096: AHG2/AHG9: On Multiview View Position (MVP) SEI message

1. **Remaining bug tickets**

* [#1504](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1504) is discussed in JVET-X0075.
* [#1513](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1513) point to a mismatch between VTM software and draft text. It needs to be clarified that the VTM behaviour reflects the original intent.

1. **Other AHG2 related inputs**

From: Ajay Luthra <ajay@picsellabs.com>

Sent: Wednesday, August 18, 2021 22:06

Subject: Formatting error in VVC v2 DIS

1. First two tables in the Annex A are labeled and referred to as A.1 and A.2 (which is right). However, the rest of the Tables in Annex A are labeled as Table 3, 4 ...11, instead of A.3, A.4 ... A.11 and they are also referred to as such in many places. E.g. in Annex A it says "All other combinations of the syntax elements in Table 4 with general\_profile\_idc equal to 2, 10, 66, 34, 42, 98, 36, 44, or 100 are reserved .... This table should be A.4 and needs to be referred as Tables A.4 in this Annex as well and at other places (like Annex D)

2. This problem continues in Annex D (Table 12 should have been Table D.1)

3. There is Table 143 in Annex D - it is quite a jump in the number and (I think) should have been D.2

4. Minor error: In the Content list (Table of Contents) some titles show up as Bold: 8.1.1, 8.1.2 and B.2 through B.4 (on Pages iv and v)

1. **Recommendations**

The AHG recommends to:

* Approve JVET-W1004, JVET-W2002, JVET-W2005, and JVET-W2006 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-X0003](https://jvet-experts.org/doc_end_user/current_document.php?id=11148) JVET AHG report: Test model software development (AHG3) [F. Bossen, X. Li, K. Sühring, Y. He, K. Sharman, V. Seregin, A. Tourapis]

**Abstract**

This report summarizes the activities of the AhG3 on Test model software development that has taken place between the 23rd and 24th JVET meetings.

1. **Introduction**

The mandates given to the AHG are:

* Coordinate development of test model (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.

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

* [VTM 14.0](https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/-/releases/VTM-14.0) (Aug. 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 published by ITU-T and ISO/IEC. It is planned to create repositories with the latest versions available in ITU-T H.264.2 (02/2016). All development history is lost.

1. **Software development**

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

The server is located at:

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

The registration and development workflow are documented at:

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

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

1. **VTM related activities**

The VTM software can be found at

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

The software development continued on the GitLab server. VTM versions 13.2 and 13.3 were tagged on Jul. 19 and Jul. 20, and VTM version 14.0 was tagged on Aug. 26. VTM 14.1 is expected during the 24th JVET meeting.

VTM 13.1 was tagged on Jul. 19, 2021. Changes include:

* JVET-V0061 display orientation SEI message
* JVET-U0084: EDRAP SEI message
* JVET-U0082 SDI SEI and three other SEIs: MAI, ACI, and DRI.
* JVET-V0108: Colour transform information SEI
* JVET-V0111: Decoding Unit Information
* JVET-Q0443: Implement MinCR check for subpictures using Subpicture SEI message
* Fix for Ticket[#1489](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1489) - Cleanup of variable names for m\_iSourceWidth and m\_iSourceHeight
* Fix TMVPMode=2: don't search affine candidates when there are none
* Fix TestSAODisableAtPictureLevel=1: reset picture when SAO disabled for picture
* Fix EncDbOpt=1: make sure reference luma block has same size as chroma block
* Fix [#1492](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1492): Values in cfg\_sliFractions and cfg\_sliNonSubpicLayersFractions should have a range of 0 to 255

VTM 13.2 was tagged on Jul. 20, 2021. Changes include:

* Fix [#1490](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1490): Some pictures are not output when bitstream has mixed NALUs with CRA subpicture
* Remove macros from previous cycle

VTM 14.0 was tagged Aug. 26, 2021. Changes include:

* JVET-W0134 Uniform metrics log
* JVET-W0070 & W0121: Condition the signaling of sps\_ts\_residual\_coding\_rice\_present\_in\_sh\_flag on...
* JVET-W0043: Alignment of smooth block QP control with adaptive QP in VTM
* JVET-W0129: change default encoder cfg: enable using true original samples for ALF
* JVET-W0178: Bitsream constraints on RExt tools for low bit-depth
* JVET-W0046: reverse last significant coefficient position
* JVET-W2005: Operation range extension profiles based on JVET-W0136 and meeting decisions
* Fix high bit depth support after !2081 (JVET-V0108)
* Fix segmentation fault when calling xCreateUnavailablePicture() on RPL1
* fix #1496: add missing condition on cu height
* Fix memory leak after !2081
* Fix #1497: Encoder crash for RGB content when LFNST and TransformSkip disabled
* Fix #1498: Fix decoder mismatch for RGB content when LFNST is disabled
* Fix #1499: Fix a bitsream constraint on RExt tools for low bit-depth (JVET-W0178)
* Fix #1503: Added check for transform skip being disabled if TSRCRicePresent is enabled
* Fix #1502: Low bit-depth encoding fails due to the check in riceStatReset
* Fix #1495: scalable encoding fails when GDR\_ENABLED=1
* Fix parameters sets used for unavailable picture creation
* Fix #1495 and #1501: GDR-related issues
* Fix for #1505: access to m\_pProfile, when no profile is set (Profile=none)
* Change hbd.cfg etc file to set appropriate Profile for HBD CTC
* Revert chroma scaling changes introduced in !2095

VTM 14.1 is expected to be tagged during the 24th JVET meeting. 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
* 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
* (open merge requests)

***CTC Performance***

The following tables show **VTM 14.0** performance over **HM 16.24**:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **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 14.0** performance compared to **VTM 13.2**:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **All Intra Main10** |  |  |
|  |  |  | **Over VTM 13.2** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | 0,00% | 0,00% | 0,00% | 97% | 101% |
| Class A2 | 0,00% | 0,00% | 0,00% | 96% | 101% |
| Class B | 0,00% | 0,00% | 0,00% | 96% | 99% |
| Class C | 0,00% | 0,00% | 0,00% | 96% | 103% |
| Class E | 0,00% | 0,00% | 0,00% | 96% | 104% |
| **Overall** | 0,00% | 0,00% | 0,00% | 96% | 101% |
| Class D | 0,00% | 0,00% | 0,00% | 96% | 104% |
| Class F | 0,00% | 0,00% | 0,00% | 99% | 102% |
|  |  |  |  |  |  |
|  |  |  | **Random access Main10** |  |  |
|  |  |  | **Over VTM 13.2** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | -0,23% | -0,07% | 0,10% | 101% | 100% |
| Class A2 | -0,07% | 0,05% | 0,16% | 101% | 101% |
| Class B | -0,16% | 0,14% | 0,03% | 100% | 103% |
| Class C | -0,51% | -0,28% | -0,52% | 100% | 103% |
| Class E |  |  |  |  |  |
| **Overall** | -0,25% | -0,03% | -0,08% | 100% | 102% |
| Class D | -0,86% | -0,61% | -0,68% | 100% | 103% |
| Class F | 0,00% | 0,00% | 0,00% | 100% | 103% |
|  |  |  |  |  |  |
|  |  |  | **Low delay B Main10** |  |  |
|  |  |  | **Over VTM 13.2** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | 0,00% | 0,00% | 0,00% | 101% | 101% |
| Class C | 0,00% | 0,00% | 0,00% | 101% | 107% |
| Class E | 0,00% | 0,00% | 0,00% | 100% | 105% |
| **Overall** | 0,00% | 0,00% | 0,00% | 101% | 104% |
| Class D | 0,00% | 0,00% | 0,00% | 99% | 105% |
| Class F | 0,00% | 0,00% | 0,00% | 100% | 102% |
|  |  |  |  |  |  |
|  |  |  | **Low delay P Main10** |  |  |
|  |  |  | **Over VTM 13.2** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | 0,00% | 0,00% | 0,00% | 100% | 100% |
| Class C | 0,00% | 0,00% | 0,00% | 100% | 106% |
| Class E | 0,00% | 0,00% | 0,00% | 100% | 110% |
| **Overall** | 0,00% | 0,00% | 0,00% | 100% | 104% |
| Class D | 0,00% | 0,00% | 0,00% | 101% | 106% |
| Class F | 0,00% | 0,00% | 0,00% | 100% | 102% |

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

***Issues in VTM 14.x affecting conformance***

The following issues in VTM master branch (Oct. 6, 2021) 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, as issue #1490 was resolved (handling of CRA subpicture with mixed NAL units).

***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-S0117 (JVET-S0141 item 11)
* JVET-S0157 item 2 (JVET-S0141 item 13)
* JVET-S0157 item 4 (JVET-S0141 item 14)
* JVET-S0175 aspect 3 (JVET-S0141 item 16)
* JVET-S0175 aspect 1, 2 (JVET-S0141 item 17)
* JVET-S0175 aspects 4 and 5 (JVET-S0141 item 18)
* JVET-S0175 aspect 6 (JVET-S0141 item 19)
* JVET-S0198/ JVET-S0223 (JVET-S0141 item 24)
* JVET-S0173 aspect 2 (JVET-S0141 item 40.b)
* JVET-S0173 item 1 (JVET-S0141 item 51)
* JVET-S0173 item 3 (JVET-S0141 item 52)
* JVET-S0173 item 5 (JVET-S0141 item 53)
* JVET-S0173 item 6 (JVET-S0141 item 54)
* JVET-S0173 item 4 (JVET-S0141 item 56)
* JVET-S0176 item 4 (JVET-S0141 item 60)
* JVET-S0154 aspect 5 (JVET-S0141 item 68)
* JVET-S0154 aspect 6 (JVET-S0141 item 69)
* JVET-S0154 aspect 8 (JVET-S0141 item 71)
* JVET-S0095 aspect 5 (JVET-S0145 item 5)
* JVET-S0095 aspect 6 (JVET-S0145 item 6)
* JVET-S0100 aspect 1, depends on JVET-R0193 (JVET-S0147 item 2)
* FINB ballot comments
* Make high tier support up to 960.

1. **HM related activities**

HM 16.24 is was tagged on Oct. 5, 2021. Changes include:

* JVET-V0056: Changes to MCTF
* JVET-V0078: QP control for very smooth blocks
* JCTVC-AM0023: Illustration of the film grain characteristics SEI message in HEVC
* JCTVC-AM0024: Illustration of the shutter interval info SEI message in HEVC Draft
* Update Conformance Window code (backport from VTM)
* Update of copyright year

The following actions have yet to be included:

* JVET-T0050: Add ability to detect static objects to encoder

Merge requests are available, but have pending discussions.

The following tables show **HM 16.24** performance compared to **HM 16.23**:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **All Intra Main10** |  |  |
|  |  |  | **Over HM-16.23** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | 0,00% | 0,00% | 0,00% | 100% | 102% |
| Class A2 | 0,00% | 0,00% | 0,00% | 100% | 102% |
| Class B | 0,00% | 0,00% | 0,00% | 99% | 98% |
| Class C | 0,00% | 0,00% | 0,00% | 99% | 101% |
| Class E | -0,01% | 0,00% | 0,00% | 99% | 98% |
| **Overall** | 0,00% | 0,00% | 0,00% | 99% | 100% |
| Class D | -0,01% | -0,01% | -0,01% | 104% | 112% |
| Class F | 0,00% | 0,00% | 0,00% | 99% | 98% |
|  |  |  |  |  |  |
|  |  |  | **Random access Main10** |  |  |
|  |  |  | **Over HM-16.23** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | -1,42% | -1,46% | -1,84% | 100% | 99% |
| Class A2 | -1,14% | -1,10% | -1,24% | 101% | 100% |
| Class B | -1,57% | -1,82% | -1,68% | 101% | 103% |
| Class C | -0,98% | -1,07% | -0,86% | 101% | 97% |
| Class E |  |  |  |  |  |
| **Overall** | -1,30% | -1,40% | -1,41% | 101% | 100% |
| Class D | -0,98% | -0,77% | -0,62% | 100% | 93% |
| Class F | 0,00% | 0,00% | 0,00% | 99% | 95% |
|  |  |  |  |  |  |
|  |  |  | **Low delay B Main10** |  |  |
|  |  |  | **Over HM-16.23** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | 0,00% | 0,00% | 0,00% | 99% | 101% |
| Class C | 0,00% | 0,00% | 0,00% | 98% | 102% |
| Class E | 0,00% | 0,00% | 0,00% | 97% | 97% |
| **Overall** | 0,00% | 0,00% | 0,00% | 98% | 100% |
| Class D | 0,00% | 0,00% | 0,00% | 99% | 104% |
| Class F | 0,00% | 0,00% | 0,00% | 97% | 98% |
|  |  |  |  |  |  |
|  |  |  | **Low delay P Main10** |  |  |
|  |  |  | **Over HM-16.23** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | 0,00% | 0,00% | 0,00% | 99% | 99% |
| Class C | 0,00% | 0,00% | 0,00% | 98% | 99% |
| Class E | 0,00% | 0,00% | 0,00% | 97% | 94% |
| **Overall** | 0,00% | 0,00% | 0,00% | 98% | 98% |
| Class D | 0,00% | 0,00% | 0,00% | 99% | 103% |
| Class F | 0,00% | 0,00% | 0,00% | 98% | 97% |

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

Help to address these tickets would be appreciated.

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

1. **360Lib related activities**

There is no update for 360Lib software.

The following table is for the projection formats comparison using VTM-14.0 according to 360-degree video CTC (JVET-U2012) compared to that using VTM-13.0 (VTM-13.0 as anchor).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **PERP: VTM-14.0 over VTM-13.0** | | | | | |
|  | **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-14.0 (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-14.0 and HM-16.22 (HM as anchor), respectively.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **VTM-14.0 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-13.0 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% |

1. **SCM related activities**

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

1. **SHM related activities**

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

1. **HTM related activities**

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

1. **HDRTools related activities**

HDRTools version 0.23 was tagged on October 5th, 2021. Changes include:

* Add fix to support integer N->M bitdepth conversion.
* Also add support for log conversion of MSSSIM (JVET module))
* Add start frame number CLI option
* clean up old project files
* add a few guidelines on how to use libpng with Xcode

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

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

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

1. **Bug tracking**

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

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

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

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

Bug tracking for HDRTools is located at:

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

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

1. **Software repositories**

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

The subversion repositories for SHM and HTM were converted to git and are now stored on the GitLab server to unify access and development process.

1. **CTC alignment**

The following differences were found in CTC alignment between HEVC and VVC:

- For HM two test configurations are described: one for 8-bit coding bit depth for Main profile and a second one for 10-bit coding bit depth for Main 10. VTM only specifies a 10-bit test case. These should be aligned, so that the same templates can be used.

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.

There was no activity during this meeting cycle on creating merged CTC documents.

1. **Recommendations**

The AHG recommends to:

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

[JVET-X0004](https://jvet-experts.org/doc_end_user/current_document.php?id=11149) 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]

This report summarizes the activities of the AhG4 on Test material and visual assessment that has taken place between the 23rd and 24th JVET meetings.

The test materials available for JVET are summarized in the Annex.

1. **Introduction**

The mandates of this AHG were:

* Produce the verification test report JVET-W2020, and consider plans for additional verification testing of VVC capability.
* 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.

1. **Activities**

***Verification test***

Report JVET-W2020 has been produced as output of the last meeting.

Offline preparation work towards the development of verification tests for VVC scalable coding has been performed by InterDigital.

***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).

1. **Related contributions**

|  |  |  |  |
| --- | --- | --- | --- |
| [JVET-X0051](https://jvet-experts.org/doc_end_user/current_document.php?id=11044) | m57844 | AG5-Related: Full Reference Video Quality Analysis |  |

1. **Recommendations**

The AHG recommends:

* 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 continue to collect new test sequences available for JVET with licensing statement.

Joint meeting with AG5 to be planned (next week)

Expert viewing on NN (if yes, should be prepared as early as possible)

[JVET-X0005](https://jvet-experts.org/doc_end_user/current_document.php?id=11150) JVET AHG report: Conformance testing (AHG5) [J. Boyce, W. Wan, E. Alshina, F. Bossen, I. Moccagatta, K. Kawamura, D. Rusanovskyy, K. Sühring, X. Xu, T. Zhou]

This document summarizes the activity of AHG5: “Conformance testing” between the 23rd Meeting (teleconference, 7–16 July 2021) and the 24th Meeting (teleconference, 6–15 October 2021).

1. **Mandates**

At the 23rd JVET meeting, the AHG on Conformance testing was established with the following mandates:

* Study the JVET-U2008 draft conformance testing specification and investigate the need for extensions.
* Study the JVET-W2026 draft conformance testing for operation rage extensions and investigate the need for improvements.
* Study the requirements of VVC, HEVC, and AVC conformance testing to ensure interoperability.
* Maintain and update the conformance bitstream database.
* Study additional testing methodologies to fulfil the needs for VVC conformance testing.

1. **Timeline**

The progress on the Conformance testing specification is consistent with the preliminary timeline agreed at the 16th JVET meeting, as follows:

* 17th meeting Jan. 2020: Preliminary guidelines for bitstream preparation (e.g., naming conventions),  
  improved list of conformance bitstreams
* 18th meeting Apr. 2020: Final guidelines for bitstream preparation and improved list of conformance  
  bitstreams with identified responsible experts, initial bitstreams provided
* 19th meeting July 2020: Confirmed list of bitstreams to be included in VVC, collection of bitstream  
  candidates for CD ballot at next meeting
* 20th meeting Oct. 2020: CD of conformance specification
* 21st meeting Jan. 2021: Final bitstreams provided, DIS ballot in ISO/IEC
* 22nd meeting April 2021: No action pending DIS ballot
* 24th meeting Oct 2021: Final conformance specification

The timeline of Conformance testing for operation rage extensions specification has not been discussed.

1. **Status on bitstream submission**

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

* conformance bitstreams for VVC:
  + 104 bitstream categories have been identified
  + At least one bitstream has been submitted in each identified category
  + 283 total bitstreams have been provided, checked, and made available
* conformance bitstreams for VVC operation range extensions:
  + 51 bitstream categories have been identified
  + Volunteers have been identified to generate bitstreams in 47 categories
  + Volunteers have been identified to cross-check bitstreams in 37 categories
  + 14 bitstream descriptions have been provided
  + 76 bitstreams of 12 identified categories have been uploaded

1. **Activities and Discussion**

The AHG activities are on schedule with the preliminary timeline shown in section 2. The final conformance specification is expected to be output from this meeting.

VVC activities:

* At the 23rd meeting July 2021 it was reported that a problem was identified with the MNUT\_A\_Nokia\_3 conformance bitstream which requires a change to the VTM13.0 addressed in the merge request in ticket [#1490](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1490). The merge request has been merged, and the conformance bitstream has been re-generated (MNUT\_A\_Nokia\_4).
* MNUT\_B\_Nokia\_2 has been re-generated (MNUT\_B\_Nokia\_3) to slightly improve coverage.
* IBC\_E\_Tencent\_1 has been generated to cover all possible block sizes for IBC in 4:2:0 color format.
* Subfolder has been removed from some .zip, older version empty .txt file has been removed from two packages.

There are not currently any known issues with the other provided VVC conformance bitstream packages. A VTM14 directory of bitstreams was made available. All provided bitstreams can be decoded using VTM14.

VVC operation range extensions activities:

* Volunteers to generate and cross-check the bitstreams have been solicited and identified.
* Folders have been generated in the ftp and http sites to upload bitstreams and to distribute them to cross-checkers.

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

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

1. **Contributions**

JVET-X0161 AHG5: Editors update on VVC conformance testing [J. Boyce, E. Alshina, F. Bossen, K. Kawamura, I. Moccagatta, W. Wan]

Another input (update of v2 conformance) may be submitted.

1. **Ftp site information**

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

* VVC:

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

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

* VVC operation range extensions:

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

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

The ftp site for uploading bitstream file is as follows.

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

(user id: avguest, passwd: Avguest201007)

If using FileZilla, the following configuration is suggested:



1. **Recommendations**

The AHG recommends the following:

* Finalize the conformance specification for VVC as an output document from this meeting
* Solicit and identify volunteers to provide and cross-check bitstreams for the remaining 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 1
* Discuss the conformance VVC operation range extensions timeline

It is discussed whether it would be useful to delay the approval of reference SW FDIS until next meeting, and include v2 features right away. If conformance has a normative reference to SW, it should be delayed as well. This also would make sense to align approval in ITU-T and ISO/IEC (both in January).

On the other hand, it is pointed out that often SW bugs are found in context of conformance development, and conformance for v2 would not be ready by January.

It is pointed out that v2 conformance could also include new streams for VVC v1 profiles.

Conformance for both v1 and v2 should progress as quick as possible, considering that volunteers might be leaving. However, for v1 everything is basically ready.

Plan for timing:

* FDIS v1 conformance 2022/01 (plus ITU consent)
* FDIS v1+v2 software 2022/01 (plus ITU consent)
* CDAM v2 conformance this meeting (if possible), FDAM 2022/07, ITU consent in October
* If necessary (bug fixing for v2), new software in October

[JVET-X0006](https://jvet-experts.org/doc_end_user/current_document.php?id=11151) JVET AHG report: ECM software development (AHG6) [V. Seregin, J. Chen, F. Le Léannec, K. Zhang]

This report summarizes the activities of the AHG6 on ECM software development that has taken place between the 23rd and 24th JVET meetings.

1. **Introduction**

The mandates for the AHG are as follows.

* Coordinate development of the ECM software and associated configuration files.
* Produce documentation of software usage for distribution with the software.
* Prepare and deliver ECM-2.0 software version and the reference configuration encodings according to JVET-W2017 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.

1. **Software development**

ECM software was decoupled from VTM and a new repository was created 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-2.0:

JVET-W0066: Cross-component SAO (CCSAO)

JVET-W0069: TMP boundary handling

JVET-W0090: Adaptive reordering of merge candidates with template matching (ARMC)

JVET-W0097: GPM with MMVD and template matching GPM

JVET-W0103: Extended intra MTS

JVET-W0119: LFNST extension with large kernel (test 4.6)

JVET-W0123: TIMD with fusion

Software improvements and bug fixes:

Align config settings with VTM related to <https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1439>

Fix OBMC memory

Fix a crash with disabled TM\_AMVP macro

Fix ECM encoder crash on small sequences

Add initialization for TU 256

Bug fix for bilateral filter when LMCS is disabled.

ECM-2.0 was tagged on August 2, 2021.

***CTC Performance***

In this section, ECM-2.0 test results following ECM CTC configuration descried in JVET-W2017 are summarized.

The following tables show ECM-2.0 performance over ECM-1.0 anchor.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | |
|  | **Over ECM-1.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -1.04% | -1.43% | -2.23% | 140% | 109% |
| Class A2 | -1.01% | -2.27% | -0.94% | 139% | 113% |
| Class B | -0.97% | -3.01% | -3.99% | 137% | 116% |
| Class C | -0.72% | -1.49% | -1.62% | 133% | 123% |
| Class E | -1.10% | -2.01% | -2.79% | 131% | 119% |
| **Overall** | -0.96% | -2.12% | -2.46% | 136% | 116% |
| Class D | -0.88% | -1.53% | -1.07% | 132% | 128% |
| Class F | -0.55% | -1.97% | -2.33% | 114% | 110% |
| Class TGM | -1.08% | -1.61% | -1.63% | 110% | 103% |
|  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | |
|  | **Over ECM-1.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -1.45% | -3.15% | -5.45% | 123% | 111% |
| Class A2 | -1.07% | -3.59% | -2.70% | 120% | 107% |
| Class B | -1.07% | -5.84% | -6.26% | 131% | 114% |
| Class C | -1.26% | -3.64% | -3.56% | 132% | 117% |
| Class E |  |  |  |  |  |
| **Overall** | -1.20% | -4.26% | -4.66% | 127% | 113% |
| Class D | -1.08% | -3.33% | -2.88% | 128% | 118% |
| Class F | -1.22% | -3.94% | -3.77% | 126% | 126% |
| Class TGM | -2.02% | -3.79% | -4.08% | 133% | 121% |
|  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | |
|  | **Over ECM-1.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -0.86% | -12.85% | -12.12% | 127% | 114% |
| Class C | -0.87% | -7.38% | -6.20% | 130% | 113% |
| Class E | -1.46% | -4.02% | -6.78% | 133% | 123% |
| **Overall** | -1.01% | -8.82% | -8.81% | 130% | 116% |
| Class D | -0.66% | -6.60% | -5.70% | 125% | 113% |
| Class F | -0.84% | -7.38% | -7.04% | 125% | 134% |
| Class TGM | -3.45% | -8.88% | -8.91% | 136% | 131% |

Next tables show ECM-2.0 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 | -6.76% | -10.85% | -12.55% | 306% | 235% |
| Class A2 | -6.43% | -9.83% | -6.78% | 294% | 226% |
| Class B | -5.92% | -9.95% | -11.25% | 337% | 248% |
| Class C | -6.73% | -8.79% | -9.19% | 329% | 243% |
| Class E | -7.23% | -9.70% | -9.20% | 329% | 286% |
| **Overall** | -6.54% | -9.78% | -9.92% | 321% | 247% |
| Class D | -5.70% | -7.02% | -6.59% | 332% | 256% |
| Class F | -10.50% | -13.32% | -14.04% | 244% | 285% |
| Class TGM | -15.50% | -17.44% | -17.29% | 233% | 290% |
|  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | |
|  | **Over VTM-11.0 + V0056** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -13.50% | -15.91% | -20.31% | 342% | 504% |
| Class A2 | -14.37% | -17.39% | -16.47% | 321% | 584% |
| Class B | -12.47% | -17.52% | -17.43% | 355% | 548% |
| Class C | -14.37% | -16.46% | -16.52% | 351% | 488% |
| Class E |  |  |  |  |  |
| **Overall** | -13.56% | -16.89% | -17.57% | 345% | 529% |
| Class D | -15.35% | -16.36% | -15.88% | 358% | 530% |
| Class F | -13.20% | -16.71% | -16.88% | 319% | 438% |
| Class TGM | -14.41% | -17.97% | -18.41% | 325% | 307% |
|  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | |
|  | **Over VTM-11.0 + V0056** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -10.33% | -21.17% | -20.57% | 301% | 326% |
| Class C | -11.78% | -16.02% | -16.53% | 326% | 286% |
| Class E | -10.56% | -13.79% | -14.32% | 266% | 296% |
| **Overall** | -10.87% | -17.61% | -17.66% | 300% | 305% |
| Class D | -13.96% | -17.39% | -17.06% | 321% | 298% |
| Class F | -11.65% | -16.78% | -16.78% | 291% | 311% |
| Class TGM | -13.73% | -20.40% | -21.15% | 260% | 244% |

1. **Recommendations**

The AHG recommends to:

* Continue to develop ECM software
* Improve 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 pointed out that it might be desirable to shift some of the surplus chroma gain to luma.

[JVET-X0007](https://jvet-experts.org/doc_end_user/current_document.php?id=11152) JVET AHG report: Coding of HDR/WCG material (AHG7) [A. Segall, E. François, W. Husak, S. Iwamura, D. Rusanovskyy]

This document summarizes the activity of AHG7: Coding of HDR/WCG Material between the 23rd meeting (7 – 16 July 2021) and the 24th meeting (6 – 15 October 2021), both held by teleconference.

* Mandates

The AHG was established with the following mandates:

* Study and evaluate available HDR/WCG test content.
* Study objective metrics for quality assessment of HDR/WCG material, including investigation of the correlation between subjective and objective results.
* Compare the performance of the VTM and HM for HDR/WCG content.
* Study the luma/chroma bit allocation in the HDR CTC, especially for HLG content.
* Contribute to activities in merging HDR related CTC documents, in coordination with AHG3.
* Study additional aspects of coding HDR/WCG content.
* Activities

The primary activity of the AhG was related to the mandates of (i) comparing the performance of the VTM for HDR/WCG content and (ii) contribute to activities in merging HDR related CTC documents. This work is described in the following subsection.

* ***Anchor Generation***

The AhG generated CTC anchors for the VTM according to JVET-V2011. A summary of the performance is provided below, and more detailed information may be found in the included XLS data.

* **VTM 14.0 versus VTM 13.0**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra** | | | | | | | | | |
|  | **Over VTM-13.0** | | | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 93% | 102% |
| Class H2 |  |  |  |  |  | 0.00% | 0.00% | 0.00% | 94% | 103% |
| **Overall** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 93% | 102% |
|  |  |  |  |  |  |  |  |  |  |  |
|  | **Random Access** | | | | | | | | | |
|  | **Over VTM-13.0** | | | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | 0.04% | -0.17% | -0.19% | 0.00% | 0.26% | -0.18% | -0.12% | 0.29% | 99% | 104% |
| Class H2 |  |  |  |  |  | -0.03% | 0.10% | 0.48% | 99% | 105% |
| **Overall** | 0.04% | -0.17% | -0.19% | 0.00% | 0.26% | -0.12% | -0.04% | 0.36% | 99% | 104% |

* ***CTC* Merging**

The AhG prepared a draft version of a merged CTC document. The draft is included with this report.

* **Recommendations**

The AHG recommends the following:

* Progress the work of merging the HM and VTM HDR CTC documents
* Close the AhG

It is pointed out that, as HDR is becoming more mainstream in terms of applications, it might be useful to make the HDR CTC more prominent in the future (also in explorations), or establish a merge with SDR.

In this context, it was later commented that it would be interesting to use the results of the verification test to investigate a correlation with the diverse metrics used in HDR. Contributions on this would be welcome.

[JVET-X0008](https://jvet-experts.org/doc_end_user/current_document.php?id=11153) 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]

This document summarizes the activity of AHG8: High bit depth, high bit rate, and high frame rate coding between the 23rd meeting by teleconference (7-16 July 2021) and the 24th meeting by teleconference (6-15 October 2021).

1. **Mandates**

The AHG was established with the following mandates:

* 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 conformance testing for operation range extensions in coordination with AHG5.
* Identify suitable test material for testing of high bit depth, high bit rate, and high frame rate coding in coordination with AHG4 and AHG7.
* Study VVC entropy decoding throughput and latency in the cases of high bit depth, high bit rate, and high frame rate coding.

1. **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 23rd and 24th meetings.

The major area of work for the AHG in this meeting cycle has been the development of conformance testing. A revision of JVET-W2026 “Conformance testing for VVC operation range extensions (Draft 1)” was uploaded on 1st September and a notification was sent to the reflector. The AHG would like to thank Iole Moccagatta for assistance with these activities.

In total there are 19 high bit depth related contributions. The following section lists these contributions.

1. **Contributions**

The contributions can be split into two sections. There are 9 contributions related to high bit depth profiles and HLS, and 5 contributions are related to tools for high bit depths. In addition, 5 cross-checks have been registered.

***Profiles / HLS***

[JVET-X0050](https://jvet-experts.org/doc_end_user/current_document.php?id=11043) “AHG2: On editing notes for the 2nd edition draft text of VVC”, Y.-K. Wang (Bytedance), M. Zhou (Broadcom)

[JVET-X0073](https://jvet-experts.org/doc_end_user/current_document.php?id=11066) “AHG2: On specifying the range extensions profiles”, Y.-K. Wang (Bytedance)

[JVET-X0075](https://jvet-experts.org/doc_end_user/current_document.php?id=11068) “AHG8: Level refinement for VVC operation range extension profiles”, T. Ikai, T. Chujoh, T. Aono (Sharp)

[JVET-X0076](https://jvet-experts.org/doc_end_user/current_document.php?id=11069) “AHG8: GCI flags for VVC operation range extension profiles”, T. Ikai, T. Chujoh, T. Aono (Sharp)

[JVET-X0093](https://jvet-experts.org/doc_end_user/current_document.php?id=11086) “AHG2/AHG8: Comments on VVC operation range extensions”, B. Choi, S. Wenger, S. Liu (Tencent)

[JVET-X0095](https://jvet-experts.org/doc_end_user/current_document.php?id=11088) “AHG8: General Constraints Information (GCI) flags for operation range extensions”, B. Choi, S. Wenger, S. Liu (Tencent)

[JVET-X0106](https://jvet-experts.org/doc_end_user/current_document.php?id=11099) “On constraints for intra profiles in VVC”, T. Tsukuba, S. Keating (Sony)

[JVET-X0108](https://jvet-experts.org/doc_end_user/current_document.php?id=11101) “Coded Picture Buffer sizes and MinCr for high bit-depth profiles”, S. Keating, A. Browne, K. Sharman (Sony)

[JVET-X0109](https://jvet-experts.org/doc_end_user/current_document.php?id=11102) “On maximum bit-rates for high bit-depth profiles”, S. Keating, A. Browne, K. Sharman (Sony)

***Tools***

[JVET-X0127](https://jvet-experts.org/doc_end_user/current_document.php?id=11120) “AHG8: Modification of History Based Rice Parameter Derivation”, Y. Yu, H. Yu, Z. Xie, F. Wang, L. Xu, D. Wang(OPPO)

[JVET-X0128](https://jvet-experts.org/doc_end_user/current_document.php?id=11121) “AHG8: On History-Based Rice Parameter Derivations for Wavefront Parallel Processing”, Y. Yu, H. Yu, Z. Xie, F. Wang, L. Xu, D. Wang(OPPO)

[JVET-X0129](https://jvet-experts.org/doc_end_user/current_document.php?id=11122) “AHG8: Independent Rice Parameter Derivation for high bit depth and high bit rate extensions”, Y. Yu, H. Yu, Z. Xie, F. Wang, L. Xu, D. Wang(OPPO)

[JVET-X0136](https://jvet-experts.org/doc_end_user/current_document.php?id=11129) “AHG8: On significance, GT1, and GT2 flag coding for high bit depths”, A. Browne, S. Keating, K. Sharman (Sony)

[JVET-X0137](https://jvet-experts.org/doc_end_user/current_document.php?id=11130) “AHG8 and AHG10: On derivation of sh\_reverse\_last\_sig\_coeff\_flag and sh\_ts\_residual\_coding\_rice\_idx\_minus1”, A. Browne, S. Keating, K. Sharman (Sony)

***Cross-checks***

[JVET-X0157](https://jvet-experts.org/doc_end_user/current_document.php?id=11164) “Cross-check report of JVET-X0127: AHG8: Modification of History Based Rice Parameter Derivation”, D. Rusanovskyy

[JVET-X0158](https://jvet-experts.org/doc_end_user/current_document.php?id=11165) “Cross-check report of JVET-X0128: AHG8: On History-Based Rice Parameter Derivations for Wavefront Parallel Processing”, D. Rusanovskyy

[JVET-X0159](https://jvet-experts.org/doc_end_user/current_document.php?id=11166) “Cross-check report of JVET-X0129: AHG8: Independent Rice Parameter Derivation for high bit depth and high bit rate extensions”, D. Rusanovskyy

[JVET-X0173](https://jvet-experts.org/doc_end_user/current_document.php?id=11181) “Cross-check of JVET-X0136: AHG8: On significance, GT1, and GT2 flag coding for high bit depths”, Z. Xie, Y. Yu (OPPO)

[JVET-X0174](https://jvet-experts.org/doc_end_user/current_document.php?id=11182) “Cross-check of JVET-X0137: AHG8 and AHG10: On derivation of sh\_reverse\_last\_sig\_coeff\_flag and sh\_ts\_residual\_coding\_rice\_idx\_minus1”, Z. Xie, Y. Yu (OPPO)

1. **Benchmarks**

***VTM 14.0 vs VTM13.0***

**Standard QP Range**

The standard QP range is from QP 22 to QP 37. The most significant changes for this QP range were the adoption of JVET-W0129 which enables using true original samples for ALF, and the adoption of JVET-W0136 which disables the use of VVC v2 coding tools for the Main 12 profile which is automatically selected for these simulations.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access** | | | | | | | | | |
|  | **Over VTM13.0** | | | | | | | | | |
|  |  |  | **wPSNR** |  |  | **PSNR** |  |  |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | 0.20% | -0.21% | -0.18% | 0.06% | 0.55% | -0.15% | 0.08% | 0.50% | 103% | 102% |
| Class H2 |  |  |  |  |  | -0.12% | 0.03% | 0.38% | 102% | 102% |
| **Overall** | 0.20% | -0.21% | -0.18% | 0.06% | 0.55% | -0.13% | 0.06% | 0.44% | 103% | 102% |
|  |  |  |  |  |  |  |  |  |  |  |
|  | **All Intra** | | | | | | | | | |
|  | **Over VTM13.0** | | | | | | | | | |
|  |  |  | **wPSNR** | | | **PSNR** | | |  |  |
|  | DE100 | PSNR-L100 | Y | U | V | Y | U | V | EncT | DecT |
| Class H1 | 0.03% | 0.03% | 0.03% | 0.03% | 0.03% | 0.03% | 0.03% | 0.03% | 98% | 99% |
| Class H2 |  |  |  |  |  | 0.02% | 0.03% | 0.03% | 98% | 99% |
| **Overall** | 0.03% | 0.03% | 0.03% | 0.03% | 0.03% | 0.02% | 0.03% | 0.03% | 98% | 99% |

**Low QP Range**

The low QP range is from QP -13 to QP 12 for 12 bit processing and QP -33 to QP -8 for 16 bit processing. The gains in this QP range can be attributed to the adoption of JVET-W0046 which allows the signalling of the last significant coefficient position to be reversed.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **HDR PQ** |  |  | **AI** |  |  |  |  |  |
|  |  |  | **Over VTM13.0** |  |  |  |  |  |
|  | wPsnrY | wPsnrU | wPsnrV | psnrY | psnrU | psnrV | EncT | DecT |
| PQ444 | -0.02% | -0.03% | -0.03% | -0.02% | -0.03% | -0.03% | 101% | 100% |
| PQ422 | -0.03% | -0.05% | -0.06% | -0.03% | -0.05% | -0.06% | 99% | 98% |
| **Overall** | -0.02% | -0.04% | -0.05% | -0.02% | -0.04% | -0.05% | 100% | 99% |
|  |  |  |  |  |  |  |  |  |
|  |  |  | **LDB** |  |  |  |  |  |
|  |  |  | **Over VTM13.0** |  |  |  |  |  |
|  | wPsnrY | wPsnrU | wPsnrV | psnrY | psnrU | psnrV | EncT | DecT |
| PQ444 | 0.00% | 0.04% | 0.02% | 0.00% | 0.03% | 0.01% | 100% | 98% |
| PQ422 | -0.02% | 0.00% | 0.00% | -0.01% | 0.00% | 0.00% | 100% | 99% |
| **Overall** | -0.01% | 0.02% | 0.01% | -0.01% | 0.02% | 0.01% | 100% | 99% |
|  |  |  |  |  |  |  |  |  |
|  |  |  | **RA** |  |  |  |  |  |
|  |  |  | **Over VTM13.0** |  |  |  |  |  |
|  | wPsnrY | wPsnrU | wPsnrV | psnrY | psnrU | psnrV | EncT | DecT |
| PQ444 | 0.01% | -0.01% | 0.00% | 0.02% | -0.01% | 0.00% | 100% | 101% |
| PQ422 | 0.00% | 0.01% | 0.04% | 0.00% | 0.01% | 0.03% | 100% | 100% |
| **Overall** | 0.01% | 0.00% | 0.02% | 0.01% | 0.00% | 0.02% | 100% | 100% |
|  |  |  |  |  |  |  |  |  |
| **Overall PQ** | -0.01% | -0.01% | -0.01% | -0.01% | -0.01% | -0.01% | 100% | 99% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **HDR HLG** |  |  | **AI** |  |  |
|  |  |  | **Over VTM13.0** |  |  |
|  | psnrY | psnrU | psnrV | EncT | DecT |
| HLG444 | -0.21% | -0.24% | -0.30% | 101% | 98% |
| HLG422 | -0.23% | -0.27% | -0.32% | 101% | 98% |
| **Overall** | -0.22% | -0.25% | -0.31% | 101% | 98% |
|  |  |  |  |  |  |
|  |  |  | **LDB** |  |  |
|  |  |  | **Over VTM13.0** |  |  |
|  | psnrY | psnrU | psnrV | EncT | DecT |
| HLG444 | -0.17% | -0.16% | -0.24% | 102% | 101% |
| HLG422 | -0.20% | -0.20% | -0.32% | 101% | 100% |
| **Overall** | -0.18% | -0.18% | -0.28% | 102% | 100% |
|  |  |  |  |  |  |
|  |  |  | **RA** |  |  |
|  |  |  | **Over VTM13.0** |  |  |
|  | psnrY | psnrU | psnrV | EncT | DecT |
| HLG444 | -0.17% | -0.16% | -0.23% | 101% | 100% |
| HLG422 | -0.22% | -0.22% | -0.35% | 101% | 100% |
| **Overall** | -0.19% | -0.19% | -0.29% | 101% | 100% |
|  |  |  |  |  |  |
| **Overall HLG** | -0.20% | -0.21% | -0.29% | 101% | 100% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **SVT RGB** |  |  | **AI** |  |  |
|  |  |  | **Over VTM13.0** |  |  |
|  | psnrG | psnrB | psnrR | EncT | DecT |
| SVT16 | -0.29% | -0.28% | -0.28% | 99% | 98% |
| SVT12 | -0.33% | -0.29% | -0.29% | 99% | 99% |
| **Overall** | -0.31% | -0.28% | -0.29% | 99% | 98% |
|  |  |  |  |  |  |
|  |  |  | **LDB** |  |  |
|  |  |  | **Over VTM13.0** |  |  |
|  | psnrG | psnrB | psnrR | EncT | DecT |
| SVT16 | -0.28% | -0.26% | -0.26% | 104% | 102% |
| SVT12 | -0.29% | -0.26% | -0.26% | 103% | 100% |
| **Overall** | -0.28% | -0.26% | -0.26% | 104% | 101% |
|  |  |  |  |  |  |
|  |  |  | **RA** |  |  |
|  |  |  | **Over VTM13.0** |  |  |
|  | psnrG | psnrB | psnrR | EncT | DecT |
| SVT16 | -0.27% | -0.25% | -0.25% | 103% | 101% |
| SVT12 | -0.26% | -0.23% | -0.23% | 101% | 100% |
| **Overall** | -0.27% | -0.24% | -0.24% | 102% | 100% |
|  |  |  |  |  |  |
| **Overall RGB** | -0.29% | -0.26% | -0.26% | 102% | 100% |

**Lossless**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **PQ** | **All Intra** | | | **Low delay B** | | | **Random Access** | | |
| **ratio** | | bit-rate savings | **ratio** | | bit-rate savings | **ratio** | | bit-rate savings |
| VTM 13.0 | VTM 14.0 | VTM 13.0 | VTM 14.0 | VTM 13.0 | VTM 14.0 |
| PQ444 | 2.6 | 2.6 | -0.32% | 3.1 | 3.1 | -0.41% | 3.1 | 3.1 | -0.39% |
| PQ422 | 2.4 | 2.4 | -0.30% | 2.9 | 2.9 | -0.44% | 2.9 | 2.9 | -0.42% |
| **Overall** | **2.5** | **2.5** | **-0.31%** | **3.0** | **3.0** | **-0.42%** | **3.0** | **3.0** | **-0.41%** |
| Enc Time[%] | 104% | | | 99% | | | 97% | | |
| Dec Time[%] | 98% | | | 99% | | | 97% | | |
|  |  |  |  |  |  |  |  |  |  |
| **HLG** | **All Intra** | | | **Low delay B** | | | **Random Access** | | |
| **ratio** | | bit-rate savings | **ratio** | | bit-rate savings | **ratio** | | bit-rate savings |
| VTM 13.0 | VTM 14.0 | VTM 13.0 | VTM 14.0 | VTM 13.0 | VTM 14.0 |
| HLG444 | 1.8 | 1.8 | -0.31% | 2.0 | 2.0 | -0.43% | 2.0 | 2.0 | -0.42% |
| HLG422 | 1.7 | 1.7 | -0.29% | 1.9 | 1.9 | -0.47% | 1.9 | 1.9 | -0.46% |
| **Overall** | **1.7** | **1.7** | **-0.30%** | **1.9** | **2.0** | **-0.45%** | **1.9** | **2.0** | **-0.44%** |
| Enc Time[%] | 102% | | | 99% | | | 101% | | |
| Dec Time[%] | 96% | | | 98% | | | 101% | | |
|  |  |  |  |  |  |  |  |  |  |
| **SVT** | **All Intra** | | | **Low delay B** | | | **Random Access** | | |
| **ratio** | | bit-rate savings | **ratio** | | bit-rate savings | **ratio** | | bit-rate savings |
| VTM 13.0 | VTM 14.0 | VTM 13.0 | VTM 14.0 | VTM 13.0 | VTM 14.0 |
| SVT16 | 1.2 | 1.2 | -0.15% | 1.2 | 1.2 | -0.20% | 1.2 | 1.2 | -0.19% |
| SVT12 | 1.3 | 1.3 | -0.25% | 1.3 | 1.3 | -0.26% | 1.3 | 1.3 | -0.26% |
| **Overall** | **1.2** | **1.2** | **-0.20%** | **1.3** | **1.3** | **-0.23%** | **1.3** | **1.3** | **-0.22%** |
| Enc Time[%] | 100% | | | 103% | | | 104% | | |
| Dec Time[%] | 101% | | | 102% | | | 102% | | |

***VTM 14.0 vs HM 16.23***

**Standard QP Range**

The standard QP range is from QP 22 to QP 37. HM uses a GOP size of 16 whilst VTM uses a GOP size of 32 for this QP range.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Random Access** | | | | |
|  | **Over HM16.23** | | | | |
|  | **PSNR** |  |  |  |  |
|  | Y | U | V | EncT | DecT |
| Class H1 | -30.26% | -69.93% | -79.50% | 367% | 176% |
| Class H2 | -33.60% | -57.88% | -56.84% | 356% | 169% |
| **Overall** | -31.93% | -63.90% | -68.17% | 362% | 173% |
|  |  |  |  |  |  |
|  | **All Intra** | | | | |
|  | **Over HM16.23** | | | | |
|  | **PSNR** | | |  |  |
|  | Y | U | V | EncT | DecT |
| Class H1 | -11.85% | -67.95% | -54.93% | 2488% | 184% |
| Class H2 | -22.45% | -49.09% | -49.93% | 2965% | 171% |
| **Overall** | -17.15% | -58.52% | -52.43% | 2716% | 178% |

**Low QP Range**

The low QP range is from QP -13 to QP 12 for 12 bit processing and QP -33 to QP -8 for 16 bit processing. Both HM and VTM use a GOP size of 16 for this QP range.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **HDR PQ** |  |  | **AI** |  |  |
|  |  |  | **Over HM16.23** |  |  |
|  | psnrY | psnrU | psnrV | EncT | DecT |
| PQ444 | -7.30% | -8.46% | -9.19% | 3249% | 168% |
| PQ422 | -9.97% | -14.18% | -14.80% | 2714% | 161% |
| **Overall** | -8.63% | -11.32% | -12.00% | 2981% | 165% |
|  |  |  |  |  |  |
|  |  |  | **LDB** |  |  |
|  |  |  | **Over HM16.23** |  |  |
|  | psnrY | psnrU | psnrV | EncT | DecT |
| PQ444 | -6.73% | -4.97% | -6.61% | 313% | 165% |
| PQ422 | -7.96% | -11.75% | -12.42% | 413% | 164% |
| **Overall** | -7.34% | -8.36% | -9.52% | 363% | 164% |
|  |  |  |  |  |  |
|  |  |  | **RA** |  |  |
|  |  |  | **Over HM16.23** |  |  |
|  | psnrY | psnrU | psnrV | EncT | DecT |
| PQ444 | -6.86% | -5.17% | -6.48% | 361% | 165% |
| PQ422 | -8.21% | -12.15% | -12.42% | 500% | 163% |
| **Overall** | -7.54% | -8.66% | -9.45% | 431% | 164% |
|  |  |  |  |  |  |
| **Overall PQ** | -7.84% | -9.45% | -10.32% | 1258% | 164% |

**Lossless**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **PQ** | **All Intra** | | | **Low delay B** | | | **Random Access** | | |
| **ratio** | | bit-rate savings | **ratio** | | bit-rate savings | **ratio** | | bit-rate savings |
| HM 16.23 | VTM 14.0 | HM 16.23 | VTM 14.0 | HM 16.23 | VTM 14.0 |
| PQ444 | 2.5 | 2.6 | -1.58% | 3.1 | 3.1 | 1.18% | 3.1 | 3.1 | 0.96% |
| PQ422 | 2.3 | 2.4 | -2.72% | 2.9 | 2.9 | -0.21% | 2.9 | 2.9 | -0.36% |
| **Overall** | **2.4** | **2.5** | **-2.15%** | **3.0** | **3.0** | **0.49%** | **3.0** | **3.0** | **0.30%** |
| Enc Time[%] | 7610% | | | 1063% | | | 1100% | | |
| Dec Time[%] | 171% | | | 160% | | | 159% | | |
|  |  |  |  |  |  |  |  |  |  |
| **HLG** | **All Intra** | | | **Low delay B** | | | **Random Access** | | |
| **ratio** | | bit-rate savings | **ratio** | | bit-rate savings | **ratio** | | bit-rate savings |
| HM 16.23 | VTM 14.0 | HM 16.23 | VTM 14.0 | HM 16.23 | VTM 14.0 |
| HLG444 | 1.8 | 1.8 | -0.95% | 1.9 | 2.0 | -0.76% | 2.0 | 2.0 | -0.83% |
| HLG422 | 1.7 | 1.7 | -1.40% | 1.9 | 1.9 | -1.11% | 1.9 | 1.9 | -1.18% |
| **Overall** | **1.7** | **1.7** | **-1.17%** | **1.9** | **2.0** | **-0.94%** | **1.9** | **2.0** | **-1.00%** |
| Enc Time[%] | 8084% | | | 1119% | | | 1118% | | |
| Dec Time[%] | 164% | | | 148% | | | 150% | | |
|  |  |  |  |  |  |  |  |  |  |
| **SVT** | **All Intra** | | | **Low delay B** | | | **Random Access** | | |
| **ratio** | | bit-rate savings | **ratio** | | bit-rate savings | **ratio** | | bit-rate savings |
| HM 16.23 | VTM 14.0 | HM 16.23 | VTM 14.0 | HM 16.23 | VTM 14.0 |
| SVT16 | 1.2 | 1.2 | -0.43% | 1.2 | 1.2 | -0.01% | 1.2 | 1.2 | -0.06% |
| SVT12 | 1.3 | 1.3 | -0.75% | 1.3 | 1.3 | -0.21% | 1.3 | 1.3 | -0.27% |
| **Overall** | **1.2** | **1.2** | **-0.59%** | **1.3** | **1.3** | **-0.11%** | **1.3** | **1.3** | **-0.16%** |
| Enc Time[%] | 8781% | | | 1158% | | | 1139% | | |
| Dec Time[%] | 142% | | | 149% | | | 150% | | |

1. **Recommendations**

The AHG recommends the following:

* To review all related contributions;
* To continue high bit depth, high bit rate, and high frame rate studies.

[JVET-X0009](https://jvet-experts.org/doc_end_user/current_document.php?id=11154) JVET AHG report: SEI message studies (AHG9) [J. Boyce, S. McCarthy, C. Fogg, P. de Lagrange, A. Luthra, G. J. Sullivan, A. Tourapis, Y.-K. Wang, S. Wenger]

This document summarizes the activity of AHG9: SEI message studies between the 23rd JVET Meeting (teleconference, 7–16 Jul. 2021) and the 24th JVET Meeting (teleconference, 6–15 Oct. 2021).

1. **Introduction**

At the 23rd JVET meeting, the AHG on SEI message studies was established with the following mandates:

* Study the SEI messages in VSEI, VVC, HEVC and AVC.
* Collect software and showcase information for SEI messages, including encoder and decoder implementations and bitstreams for demonstration and testing.
* Identify potential needs for additional SEI messages.
* Investigate the possible need of mandatory post processing in the context of SEI messages
* Study SEI messages defined in HEVC and AVC for potential use in the VVC context.
* Coordinate with AHG3 for software support of SEI messages.

1. **Related contributions**

A total of 5 contributions, not including cross-checks, are identified relating to AHG9, of which:

* 3 contributions relate to the mandate to study the SEI messages in VSEI, VVC, HEVC, and AVC;
* 0 contributions relate to the mandate to collect software and showcase information of SEI messages;
* 2 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
* 0 contributions relate to the mandate to study SEI messages defined in HEVC and AVC for potential use in the VVC context.

The following is a list of contributions related to AHG9.

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

[JVET-X0059](https://jvet-experts.org/doc_end_user/current_document.php?id=11052) AHG2/AHG9: Comments on the 2nd edition draft text for VSEI [Y.-K. Wang (Bytedance)]

[JVET-X0096](https://jvet-experts.org/doc_end_user/current_document.php?id=11089) AHG2/AHG9: On Multiview View Position (MVP) SEI message. [B. Choi, A. Hinds, X. Zhang, S. Wenger, S. Liu (Tencent)]

[JVET-X0101](https://jvet-experts.org/doc_end_user/current_document.php?id=11094) AHG9: On the CREI SEI message. [R. Skupin, C. Bartnik, A. Wieckowski, K. Suehring, Y. Sanchez, B. Bross, T. Schierl (HHI)]

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

[JVET-X0092](https://jvet-experts.org/doc_end_user/current_document.php?id=11085) AHG9: Down-sample phase indication (SEI message) [J. Samuelsson, A. Tourapis, D. Podborski, K. Rapaka, D. Singer (Apple)]

[JVET-X0112](https://jvet-experts.org/doc_end_user/current_document.php?id=11105) AHG9: On post-filter SEI [ M. M. Hannuksela, E. B. Aksu, F. Cricri, H. R. Tavakoli, M. Santamaria (Nokia)]

1. **Activities**

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

1. **Recommendations**

The AHG recommends to:

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

[JVET-X0010](https://jvet-experts.org/doc_end_user/current_document.php?id=11155) JVET AHG report: Encoding algorithm optimization (AHG10) [P. de Lagrange, R. Sjöberg, A. Tourapis]

This document summarizes the activities of AHG10: Encoding algorithm optimization, between the 23rd meeting (teleconference, 7-16 July 2021) and the 24th meeting (teleconference, 6-15 October 2021).

1. **Mandates**

At the 23rd meeting of the ITU-T/ISO/IEC Joint Video Experts Team (JVET), an ad hoc group on Encoding algorithm optimization was established with the following mandates:

* 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.

The regular JVET e-mail reflector was used for discussions ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de)). No e-mail related to AHG10 activity was sent to the JVET reflector during the AHG period.

1. **Overview of input documents related to the AHG**

The following input documents were identified to be related to this AHG:

[***JVET-X0061***](https://jvet-experts.org/doc_end_user/current_document.php?id=11054) ***AHG10: Fast skip of TT split partitioning on top of VTM-14.0 and ECM reference software***

The contribution presents a fast method of conditionally skipping ternary-tree (TT) split on top of VTM-14.0 and ECM-reference software. The overall simulation results compared to VTM-14.0 (VTM-14.0 as anchor) are reported as follows:

* AI: 0.09%/0.13%/0.17% with 96% EncT and 101% DecT
* RA: 0.18%/0.19%/0.27% with 95% EncT and 99% DecT
* LB: 0.09%/-0.17%/0.41% with 97% EncT and 104% DecT
* LP: 0.12%/0.07%/0.24% with 97% EncT and 102% DecT

Further information is expected to be available comparing it to ECM-2.0, but it was not available at the time of this report.

It is indicated that JVET-W0086 explored unsymmetric partitioning methods denoted as unsymmetric binary-tree (UBT) partitioning and unsymmetric quad-tree (UQT) In the proposal, a fast-partitioning method was also proposed on top of ECM-1.0. The method is also applicable for TT split on top of both VTM and ECM. The proposed method on JVET-W0086 used the RD cost of the non-split partition and the binary-tree (BT) partition to early skip the RD process of the UBT and UQT. This fast method could be further utilized for ternary-tree (TT) partition. In this proposal, the fast method in is used to early skip the TT partitioning on top of VTM-14.0 and ECM-2.0. The proposed fast skip checking mechanism will be performed before TT split partition searching process. For horizontal TT split partition, the RD cost comparison between non-split partition and horizontal BT split partition and the RD cost comparison between horizontal BT split and vertical BT split are calculated to decide whether the horizontal TT split partition search procedure should be performed or not. For vertical TT split partition, the above similar process is also utilized to early skip the vertical TT split partition search procedure.

[***JVET-X0063***](https://jvet-experts.org/doc_end_user/current_document.php?id=11056) ***AHG10: Deblocking filter setting for VTM***

In this contribution, a set of encoder only optimizations for deblocking filter are proposed for VTM. The deblocking filter control parameters BetaOffset\_div2 and TcOffset\_dive2 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-14.0, the overall performance are as follows:

* AI: -0.74%/-0.69%/-0.44% with 99%EncT/99%DecT
* RA: -0.32%/-0.58%/-0.50% with 100%EncT/100%DecT
* LDB: -0.62%/-0.41%/0.02% with 100%EncT/100%DecT
* LDP: -0.34%/-0.26%/-0.09% with 100%EncT/100%DecT

In the ECM the deblocking filter (DBF) control parameters BetaOffset\_div2 and TcOffset\_div2 vary with the temporal level for better coding efficiency. Two new syntax elements pps\_beta\_offset\_div2\_num\_minus1 and pps\_beta\_Tc\_div2\_num\_minus1 are also signaled in ECM bitstream. However, in VTM CTC, BetaOffset\_div2 and TcOffset\_div2 are set to the same value for all temporal layers. To align the deblocking settings of ECM and VTM while avoiding new syntax elements in VTM, this contribution proposes to use temporal layer dependent DBF control parameters in VTM.

[***JVET-X0116***](https://jvet-experts.org/doc_end_user/current_document.php?id=11109) ***AHG10: Suggestion to enable GOP-based temporal filtering for low-delay configurations in CTC for HM and VTM***

GOP-based temporal filtering / Motion Compensated Temporal Filtering (MCTF) is enabled in CTC for random-access for HM and VTM. For low-delay MCTF has so far been disabled due to bad performance for smaller resolutions. In this contribution we show that by enabling MCTF for low-delay configurations in HM-16.24rc1 and VTM-14.0 a BD-rate impact of -0.7% and -0.8% can be achieved for HM respectively VTM for low-delay B and -1.1% for low-delay P. The contribution also indicated that they have also checked visually that it works well.

It is proposed therefore suggest to also enable temporal filtering for low-delay configurations in CTC for HM and VTM.

MCTF was firstly introduced in VTM/HM using two pictures before and two pictures after the current picture for every eight picture in random access configuration indicated on JVET-O0549. It was later further improved in JVET-V0056 and then also using four pictures before and four pictures after the current picture in random access configuration. For low-delay configuration it has been present in reference software to only use pictures before the current picture, but it has been disabled in CTC because of bad performance for small resolutions, e.g. Class C and Class D. Tis contribution presents results for the improved MCTF which is part of VTM-14.0 and HM-16.24rc1 but here with more conservative setting and indicating general improvements on the small resolutions for Low delay B Main10 and Low delay P Main10.

|  |
| --- |
| ***JVET-X0137 AHG8 and AHG10: On derivation of sh\_reverse\_last\_sig\_coeff\_flag and sh\_ts\_residual\_coding\_rice\_idx\_minus1*** |

This document describes methods for determining the signalled values **sh\_reverse\_last\_sig\_coeff\_flag** and **sh\_ts\_residual\_coding\_rice\_idx\_minus1** which are used to decode coefficient values in the tools adopted in VVC v2 described in JVET-V0121 and JVET-W0046.

The methods described operate on the original image data and therefore can be used before coefficients are coded. This allows the values of the signalled parameters to be determined using image data for the first frame of a sequence unlike the current implementations in VTM which rely on QP and bit depth based formulas for the first frame of a sequence.

Some gains are reported on the contribution, but limited set of results were available at the time of the creation of this report.

***JVET-X0143 AHG10: VTM Encoder Changes for ALF Usage with Subpicture***

This document proposes several VTM software changes which enable the use of ALF in applications where extraction and merging of subpictures occur, e.g., in subpicture-based viewport adaptive streaming (VAS). Currently, ALF should be disabled in such applications when encoding using VTM software. The proposed changes include: 1) making the maximum number of ALF APSs per picture configurable, 2) allowing both luma and chroma in case of having one ALF APS, and 3) introducing an offset value to the ALF APS identifiers. This offset along with limiting the maximum number of ALF APSs per picture allows controlling the allocation of ALF APS identifiers in sub-streams which may be merged into a single stream. Moreover, the contribution proposes a VTM encoder configuration that enables a constant joint chroma coding residual (JCCR) sign, since JCCR sign is a picture-level parameter and hence needs to be the same in all encoded sub-streams which may be merged into a single stream. Furthermore, the contribution proposes several changes to the VTM subpicture merger tool to improve the performance of merging operation. Finally, the contribution illustrates the benefit of enabling ALF in a subpicture-base VAS use case.

|  |
| --- |
| ***JVET-X0104 Update on the progress of the optimized VVC encoder implementation, Ali266*** |

This contribution reports updated information on runtime and coding efficiency of Ali266, an optimized software VVC encoder developed by Alibaba. Ali266 was first presented previously in JVET-W0127, it supports two presets, with its fast preset aimed at real-time encoding applications.

In this document, it reported further improvement in Ali266’s encoding speed at its fast preset, and also the addition of 8-bit encoding to Ali266 with the continued support of 10-bit encoding. It is reported that a 1.9x encoding speedup has been achieved in the past three months for video sequences at 1920x1080 or 1080x1920 resolutions. Currently, when configured with fast preset, Ali266 achieves average encoding speeds of 36.3 fps for classes 1920x1080 / 1080x1920 and 63.4 fps for classes 1280x720 / 720x1080. Based on such data, it is expected that Ali266 is a VVC encoder capable of real-time encoding for 1080p sequences, and Ali266 is more ready for real-world commercial applications.

Besides the speed improvement, Ali266’s encoding performance is also improved. For the JVET CTC content, Ali266 at slow preset relative to VTM-13.0 achieves PSNR BD-rates for {Y, U, V} of {10.95%, -6.11%, -4.50%} under the Random Access configuration with an average speedup factor of 227 times over VTM-13.0.

1. **Recommendation**

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

[JVET-X0011](https://jvet-experts.org/doc_end_user/current_document.php?id=11156) 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]

This document summarizes the activities of AHG11: Neural Network-based video coding between the 23rd meeting (7 – 16 July 2021) and the 24th meeting (6 – 15 October 2021), both held by teleconference.

1. **Introduction**

The AHG was established with the following mandates:

* 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 JVET-W2023.
* Solicit input contributions on NN-based video coding technologies.
* Refine the test conditions for neural network-based video coding, 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 video 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.
* Review the outcome of the expert viewing conducted at the meeting, refine the methodology, and prepare viewing for the next meeting.
* Generate and distribute anchor encodings and develop improvements of the JVET-W2016 common test conditions for NNVC technology.
* Coordinate with other relevant groups, including SC29/AG5 on visual quality assessment.

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. Two emails were exchanged on the reflector.

***EE Coordination***

The AHG finalized, conducted and discussed the EE on NN based video coding. This was accomplished via the reflector. And, the final version of the EE description was announced on August 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 August 16, 2021.

Furthermore, the following updates were made to the AHG Git repository to better support the CTC activity:

1. A patch to support the MS-SSIM2 calculation was developed and added to the Git repository. The patch supports the full size calculation for RPR experiments.
2. The TVD (Tencent Video Dataset) was added to the list of training data, as previously agreed.
3. The YouTube UGC (User Generate Content) dataset was added to the list of training data, as previously agreed.
4. The license conditions for print\_complexity.py were clarified and aligned with the statement used in other JVET software.

The AHG11 coordinators would like to thank Johannes Sauer for their generous contributions.

***Anchor Encoding***

Anchors for the NN-based video coding activity were created 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, it is intended to perform remote viewing sessions to understand the visual benefit of the approaches.

***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-X0041 | AHG11: CNN-based In-Loop Filter with Knowledge Distillation | Yes | Partial  (Class C, D) | Partial  (Class C, D) | Partial  (Class C, D) | BVI-DVC | - |
| JVET-X0054 | AHG11: neural network based in-loop filter with adaptive model selection | Yes | Yes | Yes | Yes | BVI-DVC  TVD | DIV2K |
| JVET-X0055 | AHG11: neural network based in-loop filter with constrained storage and low complexity | Yes | Yes | Yes | Yes | BVI-DVC  TVD | DIV2K |
| JVET-X0082 | EE1-related: Training Using Knowledge Distillation for Deep In-Loop Filters | Yes | No | No | Waiting | - | DIV2K |
| JVET-X0084 | EE1-related: Improved RDO Considering Deep In-Loop Filter | No  *Missing description table* | Waiting | Waiting | Yes | Unknown | Unknown |
| JVET-X0094 | AHG11: A Deep In-Loop Filter with Frame Level Flag | Yes | Waiting | No | Yes | - | DIV2K |
| JVET-X0126 | AHG11: Neural Network-based Adaptive Model Selection for CNN In-Loop Filtering | Yes | Yes | No | Yes | - | DIV2K |
| **Post-Processing** | | | | | | | |
| JVET-X0111 | AHG11: MPEG NNR compressed bias update for the CNN based post-filter of EE1-1.1 | Yes | Partial  (Class B,C,D,F) | No | No | BVI-DVC | JVET-SCC |
| **Super-Resolution** | | | | | | | |
| JVET-X0080 | EE1-related: CNN-based Super Resolution for Video Coding Using Decoded Information with Simplified Models | Yes | Partial  (Class A1, A2) | No | Partial  (Class A1, A2) | BVI-DVC | DIV2K |
| JVET-X0081 | EE1-related: CNN-based Super Resolution for Video Coding Using Separate Networks for Chroma Components | Yes | Partial  (Class A1, A2) | No | Partial  (Class A1, A2) | BVI-DVC | DIV2K |
| JVET-X0097 | AHG11: A CNN-based Super Resolution Method | No  *Different QP range* | No | No | Yes | BVI-DVC | - |
| JVET-X0113 | AHG11: CNN-based Low Complexity Super Resolution | Yes | Partial  (Class A1, A2) | No | No | BVI-DVC | - |
| **Inter-Prediction** | | | | | | | |
| JVET-X0060 | AHG11: NN-based Reference Frame Interpolation for VVC Hierarchical Coding Structure | Yes | Partial  (Class B,C,D) | No | No | - | Vimeo |
| JVET-X0102 | AHG11: Deep neural network for inter bi-prediction | No  *VTM-11.0 Anchor* | Partial  (Class B,C,D) | No | No | BVI-DVC | - |
| **Intra-Prediction** | | | | | | | |
| JVET-X0125 | AHG11: Autoencoder-based intra prediction with auxiliary feature | No  *VTM-11.0 Anchor* | Yes | No | Yes | BVI-DVC | DIV2K |
| JVET-X0130 | AHG11: Cross-component prediction based on a neural network model | Yes | Yes | No | Yes | BVI-DVC | - |
| **End-to-End** | | | | | | | |
| JVET-X0043 | [AHG11 & AHG6] DOVC: Deep Omnidirectional Video Compression | N/A | N/A | N/A | N/A | - | JVET 360 |

1. **Input contributions**

There are 31 input contriubtions related to the AHG mandates. Thirteen of the contributions are directly related to the EE activity, while the remaining 18 contributions are related to AHG11 but not directly part of the EE. The list of input contributions is provided below.

***EE Input Contributions***

|  |  |  |
| --- | --- | --- |
| **Reporting** | | |
| 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 |
| **EE Technology** | | |
| JVET-X0052 | EE1-1.3: neural network based in-loop filter | L. Wang,W. Jiang,X. Xu,S. Liu (Tencent) |
| JVET-X0053 | EE1-1.5: neural network based in-loop filter using depthwise separable convolution and regular convolution | L. Wang,S. Lin,X. Xu,S. Liu,Xiang Li (Tencent) |
| JVET-X0064 | EE1-2.2: CNN-based Super Resolution for Video Coding Using Decoded Information | C. Lin,Y. Li,K. Zhang,L. Zhang (Bytedance) |
| JVET-X0065 | EE1-1.2: Test on Deep In-Loop Filter with Adaptive Model Selection and External Attention | Y. Li,K. Zhang,L. Zhang (Bytedance) |
| JVET-X0066 | EE1-1.6: Combined Test of EE1-1.2 and EE1-1.4 | Y. Li,K. Zhang,L. Zhang (Bytedance),H. Wang,J. Chen,K. Reuze,A.M. Kotra,M. Karczewicz (Qualcomm) |
| JVET-X0074 | EE1-2.4: 1.5x/2.0x Upsample method for NN-Based Super-Resolution Post-Filters | K. Takada,Y. Yasugi,T. Chujoh,T. Ikai (Sharp) |
| JVET-X0107 | EE1-2.3: Neural Network-based Super Resolution | A. M. Kotra,K. Reuz√©,J. Chen,H. Wang,M. Karczewicz,J. Li (Qualcomm) |
| JVET-X0110 | EE1-1.1: Content-adaptive neural network post-processing filter | M. Santamaria,J. Lainema,F. Cricri,R. G. Youvalari,H. Zhang,A. Zare,H. R. Tavakoli,M. Hannuksela (Nokia) |
| JVET-X0117 | EE1-2.1: Super Resolution with existing VVC functionality. | Elena Alshina,Johannes Sauer (Huawei) |
| JVET-X0118 | EE1-3.1: BD-rate gains vs complexity of NN-based intra prediction | T. Dumas,F. Galpin,P. Bordes,F. Le Leannec (InterDigital) |
| JVET-X0140 | EE1-1.4: Tests on Neural Network-based In-Loop Filter with Constrained Computational Complexity | H. Wang,J. Chen,K. Reuze,A.M. Kotra,M. Karczewicz (Qualcomm) |
| **Cross Checks** | | |
| JVET-X0123 | Crosscheck of JVET-X0118 (EE1-3.1: BD-rate gains vs complexity of NN-based intra prediction) | Johannes Sauer (Huawei) |

***Non-EE Input Contributions***

|  |  |  |
| --- | --- | --- |
| **Reporting** | | |
| JVET-X0011 | 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 |
| **Loop Filtering** | | |
| JVET-X0041 | AHG11: CNN-based In-Loop Filter with Knowledge Distillation | H.-K Kang,D.-W. Kim,S.-W. Jung,H. Kwon,S. Y. Jeong (ETRI) |
| JVET-X0054 | AHG11: neural network based in-loop filter with adaptive model selection | L. Wang, X. Xu, S. Liu (Tencent) |
| JVET-X0055 | AHG11: neural network based in-loop filter with constrained storage and low complexity | L. Wang,X. Xu,S. Liu (Tencent) |
| JVET-X0082 | EE1-related: Training Using Knowledge Distillation for Deep In-Loop Filters | Y. Li,K. Zhang,L. Zhang (Bytedance) |
| JVET-X0084 | EE1-related: Improved RDO Considering Deep In-Loop Filter | J. Li,Y. Li,K. Zhang,L. Zhang (Bytedance) |
| JVET-X0094 | AHG11: A Deep In-Loop Filter with Frame Level Flag | X. Zhang,C. Fang,S. Peng,D. Jiang,J. Lin (Dahua) |
| JVET-X0126 | AHG11: Neural Network-based Adaptive Model Selection for CNN In-Loop Filtering | Z. Dai,Y. Yu,H. Yu,K. Sato,L. Xu,Z. Xie,D. Wang(OPPO) |
| **Post Filtering** | | |
| JVET-X0111 | AHG11: MPEG NNR compressed bias update for the CNN based post-filter of EE1-1.1 | M. Santamaria,J. Lainema,F. Cricri,R. G. Youvalari,H. Zhang,A. Zare,G. Rangu,H. R. Tavakoli,H. Afrabandpey,M. Hannuksela (Nokia) |
| **Super Resolution** | | |
| JVET-X0080 | EE1-related: CNN-based Super Resolution for Video Coding Using Decoded Information with Simplified Models | C. Lin,Y. Li,K. Zhang,L. Zhang (Bytedance) |
| JVET-X0081 | EE1-related: CNN-based Super Resolution for Video Coding Using Separate Networks for Chroma Components | C. Lin,Y. Li,K. Zhang,L. Zhang (Bytedance) |
| JVET-X0097 | AHG11: A CNN-based Super Resolution Method | S. Peng,C. Fang,X. Zhang,D. Jiang,J. Lin (Dahua) |
| JVET-X0113 | AHG11: CNN-based Low Complexity Super Resolution | Eun Yeo,Jewon Kang (Ewha W. University),Dongcheol Kim,Kyungyong Kim,Ju-Hyung Son,JinSam Kwak (WILUS Inc.) |
| **Inter Prediction** | | |
| JVET-X0060 | AHG11: NN-based Reference Frame Interpolation for VVC Hierarchical Coding Structure | Z. Liu,X. Xu,S. Liu (Tencent),Y. Guo,Z. Chen (Wuhan Univ.) |
| JVET-X0102 | AHG11: Deep neural network for inter bi-prediction | Z. Liu,X. Xu,S. Liu (Tencent),Y. Guo,Z. Chen (Wuhan Univ.) |
| **Intra Prediction** | | |
| JVET-X0125 | AHG11: Autoencoder-based intra prediction with auxiliary feature | L. Xu,Y. Yu,H. Yu,K. Sato,Z. Dai,Z. Xie,D. Wang(OPPO) |
| JVET-X0130 | AHG11: Cross-component prediction based on a neural network model | Y.Y. Lee,T.M. Bae,D. Ruiz Coll,K. Goswami,A. Filippov,V. Rufitskiy (Ofinno) |
| **End-to-End** | | |
| JVET-X0043 | [AHG11 & AHG6] DOVC: Deep Omnidirectional Video Compression | Qipu Qin,Cheolkon Jung,Zou Dan,Ming Li (OPPO) |

1. **Recommendations**

The AHG recommends:

* Review all input contributions.
* Establish a BoG for selecting and conducting viewing of the EE results early in the meeting, in coordination with SC29/AG5.
* Discuss if DIV2K should be formally included in the training set.
* Continue investigating neural network-based video coding tools, including coding performance and complexity.

Plan for BoG (A. Segall) for Thu. morning

[JVET-X0012](https://jvet-experts.org/doc_end_user/current_document.php?id=11157) JVET AHG report: Enhanced compression beyond VVC capability (AHG12) [M. Karczewicz, Y. Ye, L. Zhang, B. Bross, X. Li, K. Naser, H. Yang]Formularende

This document reports the work of the JVET ad hoc group on enhanced compression beyond VVC capability (AHG12) between the 23rd meeting by teleconference, 7-16 July 2021, and the 24th meeting by teleconference, 6–15 October 2021.

1. **Introduction**

The mandates given to the AHG are:

* Solicit and study non-neural-network video coding tools with enhanced compression capabilities beyond VVC.
* Discuss and propose refinements to the ECM2 algorithm description JVET-W2025.
* Study the performance and complexity tradeoff of these video coding tools.
* Coordinate with AHG6 on ECM software development.
* Refine test conditions in JVET-W2017, 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-W2024 in coordination with the EE coordinators.
* Coordinate with AHG11 to study the interaction with neural network-based coding tools.

The regular JVET e-mail reflector was used for discussions ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de)). A few emails were exchanged regarding EE2.

1. **Activities**

The Common Test Conditions were updated to reduce simulation time for low delay configuration scenarios (JVET-W2017). The primary activity of the AHG was the “Exploration experiment on enhanced compression beyond VVC capability” (JVET-W2024). The indication of the achievable improvement over VVC are the results of the EE2 software base (ECM-2.0) as agreed at JVET-W meeting. The combined improvement of the ECM-2.0 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 | -6.76% | -10.85% | -12.55% | 306% | 235% |
| Class A2 | -6.43% | -9.83% | -6.78% | 294% | 226% |
| Class B | -5.92% | -9.95% | -11.25% | 337% | 248% |
| Class C | -6.73% | -8.79% | -9.19% | 329% | 243% |
| Class E | -7.23% | -9.70% | -9.20% | 329% | 286% |
| Overall | -6.54% | -9.78% | -9.92% | 321% | 247% |
| Class D | -5.70% | -7.02% | -6.59% | 332% | 256% |
| Class F | -10.50% | -13.32% | -14.04% | 244% | 285% |
| Class TGM | -15.50% | -17.44% | -17.29% | 233% | 290% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Random Access Main 10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -13.50% | -15.91% | -20.31% | 342% | 504% |
| Class A2 | -14.37% | -17.39% | -16.47% | 321% | 584% |
| Class B | -12.47% | -17.52% | -17.43% | 355% | 548% |
| Class C | -14.37% | -16.46% | -16.52% | 351% | 488% |
| Class E |  |  |  |  |  |
| Overall | -13.56% | -16.89% | -17.57% | 345% | 529% |
| Class D | -15.35% | -16.36% | -15.88% | 358% | 530% |
| Class F | -13.20% | -16.71% | -16.88% | 319% | 438% |
| Class TGM | -14.41% | -17.97% | -18.41% | 325% | 307% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Low Delay B Main 10 | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -10.33% | -21.17% | -20.57% | 301% | 326% |
| Class C | -11.78% | -16.02% | -16.53% | 326% | 286% |
| Class E | -10.56% | -13.79% | -14.32% | 266% | 296% |
| Overall | -10.87% | -17.61% | -17.66% | 300% | 305% |
| Class D | -13.96% | -17.39% | -17.06% | 321% | 298% |
| Class F | -11.65% | -16.78% | -16.78% | 291% | 311% |
| Class TGM | -13.73% | -20.40% | -21.15% | 260% | 244% |

1. **Contributions**

In addition to thirteen EE2 contributions, 40 related contributions were received which can be subdivided as follows:

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

JVET-X0086, "EE2-related: Adaptive Filter Shape Selection for ALF", W. Yin, K. Zhang, L. Zhang (Bytedance), N. Hu, V. Seregin, M. Karczewicz (Qualcomm)

JVET-X0105, "AHG12: Edge Classifier for Cross-component Sample Adaptive Offset (CCSAO)”, A. M. Kotra, N. Hu, V. Seregin, M. Karczewicz (Qualcomm)

JVET-X0152, "AHG12: CCSAO classification with edge information", C.-W. Kuo, X. Xiu, Y.-W. Chen, H.-J. Jhu, W. Chen, N. Yan, X. Wang (Kwai)

***Intra (14)***

JVET-X0072, "EE2-related: PDPC-skip scheme for angular intra modes", Jinrong Zhang, Chuan Zhou, Zhuoyi Lv (Vivo)

JVET-X0099, "Non-EE2: Extension of TIMD to intra chroma coding", Y. Wang, K. Zhang, L. Zhang, H. Liu (Bytedance)

JVET-X0100, "EE2-related: On propagating intra prediction mode for IBC", L.-F. Chen, X. Li, L. Li, S. Liu (Tencent)

JVET-X0114, "EE2-related: Fix on issues of TIMD mode", C. Zhou, Z. Lv, J. Zhang (vivo)

JVET-X0115, "EE2-related: Optimization on the second mode derivation of DIMD blending mode", C. Zhou, Z. Lv, J. Zhang (vivo)

JVET-X0122, "Non-EE2: Unification of negative modes processing in TIMD", A. Filippov, V. Rufitskiy, K. Goswami, D. Ruiz Coll, Y.Y. Lee, T.M. Bae, E. Dinan (Ofinno)

JVET-X0124, "AHG12: On signalling of intra template matching", Z. Xie, Y. Yu, H. Yu, L. Xu, F. Wang, D. Wang (OPPO)

JVET-X0131, "Non-EE2: Low-Complexity Improvements of Intra Coding for Screen Content", D. Ruiz Coll, T.M. Bae, A. Filippov, V. Rufitskiy, Y.Y. Lee, K. Goswami (Ofinno)

JVET-X0135, "Non-EE2: Adaptive intra MTS", B. Ray, V. Seregin, M. Karczewicz (Qualcomm)

JVET-X0139, "AHG12: removing a discontinuity in the discrete angle comparison in DIMD", T. Dumas, P. Bordes, F. Galpin, F. Le Leannec (InterDigital)

JVET-X0142, "Non-EE2: Extended MRL candidate list", K. Cao, Y.-J. Chang, B. Ray, V. Seregin, M. Karczewicz (Qualcomm)

JVET-X0148, "AHG12: On the PDPC handling in DIMD and TIMD", H.-J. Jhu, X. Xiu, Y.-W. Chen, W. Chen, C.-W. Kuo, N. Yan, X. Wang (Kwai)

JVET-X0149, "AHG12: Removal of floating operations in DIMD and TIMD", X. Xiu, J.-H. Jhu, W. Chen, C.-W. Kuo, N. Yan, Y.-W. Chen, X. Wang (Kwai)

JVET-X0156, "Non-EE2: Fix on issues of DIMD Histogram of Gradient generation", K. Cao, V. Seregin, M. Karczewicz (Qualcomm)

***Inter (20)***

JVET-X0056, "EE2-Related: Complexity reduction for decoder side motion derivation", H. Huang, V. Seregin, Z. Zhang, C.-C. Chen, M. Karczewicz (Qualcomm)

JVET-X0058, "AHG12: Bilateral matching SMVD mode", D. Kim, H. Kwon, J. Kim, W. Lim, S. Y. Jeong (ETRI)

JVET-X0078, "EE2-related: Modified GPM with inter and intra prediction", Y. Kidani, H. Kato, K. Kawamura (KDDI)

JVET-X0085, "Non-EE2: Template Matching-based Reordering for Extended MMVD Design", M. Salehifar, Y. He, K. Zhang, N. Zhang, L. Zhang (Bytedance)

JVET-X0087, "Non-EE2: Template Matching Based Merge Candidate List Construction (TM-MCLC)", L. Zhao, K. Zhang, N. Zhang, L. Zhang (Bytedance)

JVET-X0088, "Non-EE2: History-based Affine Model Inheritance", K. Zhang, L. Zhang, Z. Deng, N. Zhang, Y. Wang (Bytedance)

JVET-X0089, "Non-EE2: Modifications of IBC Merge/AMVP List Construction", N. Zhang, K. Zhang, L. Zhang, J. Xu (Bytedance)

JVET-X0090, "Non-EE2: On combination of CIIP, OBMC and LMCS", R.-L. Liao, X. Li, J. Chen, Y. Ye (Alibaba)

JVET-X0091, "Non-EE2: On TMVP improvement", R.-L. Liao, J. Chen, Y. Ye, X. Li (Alibaba)

JVET-X0119, "Non-EE2: On pairwise merge candidate", Laroche G., Onno P., Bellessort R. (Canon)

JVET-X0121, "EE2-related: bug fixes for enabling RPR in ECM", P. Bordes, F. Galpin, F. Leleannec, E. Francois (InterDigital)

JVET-X0132, "EE2-Related: On MVD sign prediction", Y. Zhang, B. Ray, H. Huang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-X0133, "Non-EE2: MV candidate type-based ARMC", Y.-J. Chang, H. Huang, V. Seregin, C.-C. Chen, M. Karczewicz (Qualcomm)

JVET-X0134, "Non-EE2: On the number of TM merge candidates", Y.-J. Chang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-X0141, "EE2-3.1-related: CIIP with template matching", Z. Deng, K. Zhang, L. Zhang (Bytedance), X. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba)

JVET-X0145, "EE2-related: Template matching CIIP on top of EE2-3.1", Y.-J. Chang, V. Seregin, M. Karczewicz (Qualcomm)

JVET-X0146, "Non-EE2: Decoder side motion derivation using sample's spatial correlation", H. Huang, V. Seregin, C.-C. Chen, M. Karczewicz (Qualcomm)

JVET-X0147, "EE2-related: intra mode derivation based on TIMD for GPM inter/intra", H. Jang, S. Kim, J. Lim (LGE), Y. Kidani, H. Kato, K. Kawamura (KDDI)

JVET-X0151, "AHG12: Non-adjacent spatial neighbors for affine merge mode", W. Chen, X. Xiu, Y.-W. Chen, H.-J. Jhu, C.-W. Kup, N. Yan, X. Wang (Kwai)

JVET-X0166, "EE2-related: Combination of JVET-X0078 (Test 7/8), JVET-X0147 (Proposal-2), and GPM direct motion storage", Y. Kidani, H. Kato, K. Kawamura (KDDI), H. Jang, S. Kim, J. Lim (LGE), Z. Deng, K. Zhang, L. Zhang (Bytedance)

***Coefficient Coding (2)***

JVET-X0120, "AHG12: On sign prediction", M. G. Sarwer, Y. Yan, J. Chen, R. -L. Liao (Alibaba)

JVET-X0150, "AHG12: Enhanced sign prediction", X. Xiu, Y.-W. Chen, N. Yan, C.-W. Kuo, H.-J. Jhu, W. Chen, X. Wang (Kwai)

***Test Conditions (1)***

JVET-X0138, “Proposal for a new Low Latency & Controlled Complexity (LLCC) common test conditions", Gaelle Martin-Cocher, Tangi Poirier, Saurabh Puri (Interdigital)

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.

# Project development (22)

## Deployment and advertisement of standards (0)

Contributions in this area were discussed in session 13 at 0810–0825 UTC on Monday 11 Oct. 2021 (chaired by JRO).

[JVET-X0020](https://jvet-experts.org/doc_end_user/current_document.php?id=11158) Deployment status of the HEVC standard [G. J. Sullivan] [late]

This information contribution contains a survey of deployed products and services using the HEVC standard and the formal specifications in which it is supported, along with a brief introduction to the standard written for broad readership. Revision marking is included to show changes relative to JVET-V0020-v2 of April 2021.

As of approximately July 2021, a survey of 503 respondents from “a very wide range of industries … from educational institutions and production companies to traditional broadcast channels and cable networks” by Bitmovin, Streaming Media, and Help Me Stream Research Foundation appeared to report (double-checking of the interpretation of this report is pending) [28]:

a. About 65% were currently using HEVC

b. About 80% were expecting to be using it in one year

Additional broadcast usage reported in Benin, China, Croatia, Czech Republic, Netherlands, Poland, Seychelles, South Africa; additional service by Amazon; product usage by GoPro, various new professional encoders.

[JVET-X0021](https://www.jvet-experts.org/doc_end_user/current_document.php?id=11159) Deployment status of the VVC standard [G. J. Sullivan] [late]

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).

Revision marking is included to show changes relative to JVET-W0021-v5 of July 2021.

Pointing to new contributions in 4.9.

## Text development and errata reporting (6)

Contributions in this area were discussed in session 5 at 0500–0700 UTC on Thursday 7 Oct. 2021 (chaired by JRO), and in session 17 at 0805–0830 UTC on Tuesday 12 Oct. 2021 (chaired by JRO).

[JVET-X0050](https://jvet-experts.org/doc_end_user/current_document.php?id=11043) AHG2: On editing notes for the 2nd edition draft text of VVC [Y.-K. Wang (Bytedance), M. Zhou (Broadcom)]

This contribution discusses the editing notes for the 2nd edition draft text of VVC, with suggestions for some of them.

1 On signalling of general\_lower\_bit\_rate\_constraint\_flag

During the discussion, another suggestion is made to infer the flag for certain profiles. T. Ikai is asked to provide a corresponding text (see contribution JVET-X0187, which however became obsolete by the decision to remove that flag).

2 On semantics of sh\_entry\_point\_offset\_minus1[ i ]

Regarding the first editing note, indeed it would make more sense to just remove ", and the number of subsets shall be 1", or change it to be ", i.e., the number of subsets is equal to 1", as it has been established above that the number of subsets is equal to NumEntryPoints + 1. A similar change should be made to HEVC.

This change is agreed. The two solutions are equivalent, left to editors which one is selected. Also include in HEVC related errata items.

Regarding the second editing note, the requirement would be violated when the encoder puts wrong values for the instances of sh\_entry\_point\_offset\_minus1[ i ] that for example would split coded bits of one CTU into different subsets or the coded bits of some CTUs of a slice in one subset and the coded bits of other CTUs of the same slice into another subset in this case.

Agreed that no action should be taken. May be further studied to impose it as semantics constraint rather than encoder constraint.

3 On the semantics of amvr\_flag[ x0 ][ y0 ]

Agreed. Replace “difference” by “differences” (3x), and clarify that this refers to current CU.

4 On CpbVclFactor, CpbNalFactor, FormatCapabilityFactor and MinCrScaleFactor for 12-bit profiles

Other contributions on this topic in 4.12 – see further discussion there.

5 On the informative tables showing effect of level limits on picture rate

Agreed to add mentioning the new profiles in the informative table.

6 On the addition of specifications for use of some SEI messages specified in the VSEI spec

No action necessary.

[JVET-X0093](https://jvet-experts.org/doc_end_user/current_document.php?id=11086) AHG2/AHG8: Comments on VVC operation range extensions [B. Choi, S. Wenger, S. Liu (Tencent)]

The following modifications are proposed to improve the VVC operation range extension specifications:

**Proposal-1**: Add the constraint that vps\_max\_layers\_minus1 shall be equal to 0 in the format range extension profiles.

Multilayer operation is already disabled by the ptl\_multilayer\_enabled\_flag equal 0. Not clear if any action is necessary – it was clarified offline that there is no issue with semantics.

**Proposal-2**: Update the value range of the syntax element sps\_bitdepth\_minus8 in VPS to cover the extended bitdepth range.

Agreed that a change is necessary. Another option could be to just remove the constraint, as it is anyway specified in the annex. Agreed to follow the HEVC style, change “2” to “8”, but without the reference to the annex.

**Proposal-3**: Add the description that clarifies the meaning of “-“ in the Table A.1.

Agreed that a change is necessary. See further discussion under X0073.

**Proposal-4:** Replace “active SPSs” with “referenced SPSs”.

Agreed.

[JVET-X0059](https://jvet-experts.org/doc_end_user/current_document.php?id=11052) AHG2/AHG9: Comments on the 2nd edition draft text for VSEI [Y.-K. Wang (Bytedance)]

This contribution mainly discusses the editing notes for the 2nd edition draft text of VSEI, along with a few other assertedly minor issues, mainly for triggering discussions at the meeting, with suggestions for some of the items.

1 On persistency flag for the annotated regions SEI message

It was was agreed at the JVET-T meeting in October 2020 not to add the persistency flag. Thus this editing note can be removed.

2 On semantics of the alpha channel information SEI message

The editor’s note can be removed, as there is no issue. Agreed that the additional bug fix suggested under this item should be implemented.

3 On semantics of alpha\_channel\_use\_idc

Further clarification necessary. AVC does have something somewhat equivalent by its value 3. It was added in HEVC likely for supporting some kind of proprietary usage. However, this leads to the fact that this does not even need to be an alpha map, whereas AVC specifies it is alpha, but does not specify specific blending. It was agreed that it can be left as is.

4 On semantics of alpha\_channel\_bit\_depth\_minus8, alpha\_transparent\_value, and alpha\_opaque\_value

Was carried over from HEVC. More clarification necessary by looking at the equations how blending is performed. It was agreed that this should be specified as it was specified in AVC which seems to be more correct but could be further improved editorially (als include this into the errata report for AVC and HEVC). It was further agreed to remove constraints on opaque value needing to be greater than transparent value, remove constraints on opaque value needing to be greater than transparent value.

5 On semantics of alpha\_channel\_clip\_type\_flag

Was carried over from HEVC. More clarification necessary by looking at the equations how blending is performed. It was agreed that “-“ should be changed to “+” per editor's note (2x).

6 On semantics of the multiview view position SEI message

Agreed to make the suggested change.

[JVET-X0096](https://jvet-experts.org/doc_end_user/current_document.php?id=11089) AHG2/AHG9: On Multiview View Position (MVP) SEI message [B. Choi, A. Hinds, X. Zhang, S. Wenger, S. Liu (Tencent)]

Items 2 and 3 of this document belong to this category. (item 3 was removed from v2 of this doc).

Two modifications on multiview view position SEI message, which was adopted into VSEI, are proposed.

Proposal-1: Extend the dimension of view positions from 1-D array to 2-D matrix.

Proposal-2: Add a constraint that a multiview view position SEI message shall not be contained in a scalable nesting SEI message.

Agreed on item/proposal 2.

[JVET-X0076](https://jvet-experts.org/doc_end_user/current_document.php?id=11069) AHG8: GCI flags for VVC operation range extension profiles [T. Ikai, T. Chujoh, T. Aono (Sharp)] [late]

This contribution proposes to add the following general constraint flags for VVC v2 profiles.

* gci\_no\_extended\_precision\_processing\_constraint\_flag
* gci\_no\_ts\_residual\_rice\_coding\_constraint\_flag
* gci\_no\_rrc\_rice\_extension\_constraint\_flag
* gci\_no\_persistent\_rice\_adaptation\_constraint\_flag
* gci\_no\_reverse\_last\_sig\_coeff\_constraint\_flag

A gating flag, gci\_range\_extension\_flag, is used to structuralize syntax structure and it is asserted that the proposed structure also could guide how to extend general\_constraint\_info( )) for potential future profiles.

It is clarified that the purpose of GCI flags is indicating the usage of tools in a bitstream at high level, not for enabling/disabling them (which is done in SPS/PPS). Therefore, it makes sense to include the flags. However, the gating flag might complicate parsing and definition of further extensions. GCI parsing should be kept simple. A v1 decoder would just ignore the additional flags.

Decision: Adopt JVET-X0076 (5 new GCI flags as listed above), but without gating flag.

[JVET-X0095](https://jvet-experts.org/doc_end_user/current_document.php?id=11088) AHG8: General Constraints Information (GCI) flags for operation range extensions [B. Choi, S. Wenger, S. Liu (Tencent)]

(add abstract)

Identical to the decision on JVET-X0076.

## Test conditions (1)

Contributions in this area were discussed in session 13 at 0825–0900 UTC on Monday 11 Oct. 2021 (chaired by JRO).

[JVET-X0138](https://jvet-experts.org/doc_end_user/current_document.php?id=11131) Proposal for a new Low Latency & Controlled Complexity (LLCC) common test conditions [G. Martin-Cocher, T. Poirier, S. Puri (InterDigital)]

This document proposes to define a low latency & controlled complexity (LLCC) test condition for use in applicable technical contributions to JVET meetings. Once finalized the LLCC configuration should be mandatory to test for class F sequences. Additional sequences would be desirable. LLCC is proposed as a revision of the current ECM LDP configuration. A first set of tools for this configuration is proposed for consideration. Implementation of GDR and modification of ALF and Affine in ECM are suggested to refine the LLCC configuration.

Why only class F? Is not class E also important for low latency applications.

Question: Comparison is only made versus ECM LP config. What would be the performance of the enabled ECM tools relative to VTM in a similar setting.

It is agreed that the CTC are somewhat outdated in terms of current low delay applications. LB may hardly be used in the real world. The same may however be true for other CTC (e.g., scalable, RPR).

Low complexity should include both encoder and decoder aspects.

Should this also include rate control?

New AHG (T. Poirier) on Low Latency and Controlled Complexity.

## Verification testing (1)

Contributions in this area were discussed in session 20 at 1400–XXXX UTC on Wednesday 13 Oct. 2021 (chaired by JRO).

[JVET-X0202](https://jvet-experts.org/doc_end_user/current_document.php?id=11212) AHG4: preparation of spatial scalability verification tests [P. de Lagrange, F. Urban, E. François (Interdigital), W. Hamidouche (INSA)] [late]

This contribution reports ongoing work to prepare visual tests for VVC in spatial scalability mode. The purpose of these tests is two-fold. Firstly, they aim at providing an 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 and JVET-V0167). Secondly, these tests can be used as a basis for JVET verification tests of VVC in scalable mode.

Only spatial scalability with 2x, HD->4K. Anchor is single layer 4K, and single layer 4K as secondary anchor.

Content SDR, HDR-PQ, HDR-HLG.

No MCTF used, as its impact on spatial scalability is minor if compared to single-layer or simulcast.

Performance compared to VVC single layer is worse for VVC scalable than for SHVC vs. HEVC single layer (from the previous SHVC verification test), where however the latter used a shorter RA period of 0.5 s. Scalable VVC vs. SHVC is clearly better.

Why is the target to use similar rates for base and enhancement layer? This might not be realistic in applications.

Very valuable information for preparing a scalable VVC verification test. However, different from the test currently prepared for the Brazilian call, also comparison against SHVC should be included.

Comparison against LCEVC is also mentioned, but it is expressed by experts that LCEVC is not a scalable codec in the sense of competing against simulcast, as the rate between base and enhancement is largely unbalanced, and base layer quality may not be guaranteed to be viewable.

Target establishing a test plan by the next meeting, also considering which parts could be re-used, or need to be adapted from the ongoing “Brazilian call” test preparation.

## Test material (0)

Contributions in this area were discussed in session x at XXXX–XXXX UTC on XXday X Oct. 2021 (chaired by XXX).

## Quality assessment (4)

Contributions in this area were discussed in session 20 at 1400–1500 UTC on Wednesday 13 Oct. 2021 (chaired by JRO).

See the AHG4 report discussion in section 3 and the verification testing discussions in section 4.4.

[JVET-X0051](https://jvet-experts.org/doc_end_user/current_document.php?id=11044) AG5-Related: Full Reference Video Quality Analysis [P. Topiwala (FastVDO)]

This contribution offers an approach to objective video quality assessment, based on feature extraction and learning based assessment, which can be applied to full reference video quality analysis (FR VQA). It is asserted that the performance is quite competitive with leading algorithms. Our approach strives to both minimize the computationally burden of the feature extraction stage and, using advanced yet fast machine-learning based assessment stage, to achieve high performance. Performance is measured in terms of the common Pearson and Spearman Rank Correlation Coefficients (PCC and SRCC, respectively). Asserted advantages in the full reference case are against the well-known Video Multi-Algorithm Fusion (VMAF) algorithm, where it is reported that our method can exceed 90% accuracy on several test databases, without hyperparameter search, and without prior training. That is, it is asserted that the relatively modest complexity of VMAF can be retained while achieving gains. With parametric search, or advancing to a neural network regressor, it is reported that further gains can be achieved. We present a framework that can be utilized in both full reference and no reference cases, but focus our results on the FR case here. Work on the NR case will be reported on elsewhere.

Was previously presented and discussed in AG 5.

Trained on each dataset individually.

It is mentioned that it could be interesting performing some tool-on/off tests, and investigate if the metric goes into the correct direction (which has not been always the case with VMAF).

[JVET-X0186](https://jvet-experts.org/doc_end_user/current_document.php?id=11196) Subjective Quality Assessment of VVC and HEVC Standards for 8K Video Resolution [C. Bonnineau (TDF/bcom), W. Hamidouche (INSA), N. Sidaty, J-F. Travers (TDF), O. Déforges (INSA)] [late]

Presented in session 21 1540 UTC.

This contribution provides a comparative subjective quality evaluation between the VTM-11 reference software (VVC) [1] and the HM-16.20 reference software (HEVC) [2] for 8K resolution videos. In addition, we evaluate the perceived quality improvement offered by 8K over UHD 4K resolution. The compression performance of both VVC and HEVC standards has been conducted in random access (RA) coding configuration. This test was performed on six video scenes with various spatial and temporal characteristics collected from two different sources: the Japanese organization ITE and Fraunhofer HHI [3]. Objective measurements using PSNR, MS-SSIM [4], and VMAF [5] metrics are provided. Subjectively, VVC offers an average of around 41% of bitrate reduction over HEVC for the same visual quality. In addition, a significant visual difference between uncompressed 4K and 8K has been noticed for most tested scenes.

The results of this activity are planned to be published and could be referenced. See <https://arxiv.org/pdf/2109.06555.pdf>.

[JVET-X0203](https://jvet-experts.org/doc_end_user/current_document.php?id=11213) Updated draft Guidelines for Verification Testing of Visual Media Specifications [M. Wien, L Yu, V. Baroncini]

Was presented and discussed in joint meeting with AG 5 Monday 11 Oct. 1620 UTC

[JVET-X0204](https://jvet-experts.org/doc_end_user/current_document.php?id=11214) Draft guidelines for remote experts viewing sessions (v2) [J. Jung, M. Wien, V. Baroncini]

Was presented and discussed in joint meeting with AG 5 Monday 11 Oct. 1620 UTC

## Conformance test development (2)

Contributions in this area were discussed in session x at XXXX–XXXX UTC on XXday X Oct. 2021 (chaired by XXX).

See the AHG5 report discussion in section 3, and BoG report JVET-X0XXX.

[JVET-X0161](https://jvet-experts.org/doc_end_user/current_document.php?id=11168) AHG5: Editors update on VVC conformance testing [J. Boyce, E. Alshina, F. Bossen, K. Kawamura, I. Moccagatta, W. Wan] [late]

[JVET-X0185](https://jvet-experts.org/doc_end_user/current_document.php?id=11195) AHG5: Editors update on conformance testing for VVC operation range extensions [D. Rusanovskyy, H.-J. Jhu, I. Moccagatta, M. Sarwer, Y. Yu, T. Zhou] [late]

## Software development (0)

Contributions in this area were discussed in session x at XXXX–XXXX UTC on XXday X Oct. 2021 (chaired by XXX).

## Implementation studies (2)

Contributions in this area were discussed in session x at 0900–XXXX UTC on Monday 11 Oct. 2021 (chaired by JRO).

[JVET-X0044](https://jvet-experts.org/doc_end_user/current_document.php?id=11037) 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 1.1.0 of the open VVC software encoder VVenC released in August 2021 and version 1.2.0 of the open VVC software decoder VVdeC released in September 2021.

Main changes for VVenC v1.1.0 since version 1.0.0 include:

* Faster presets: 20% for faster, 15% for fast, 7% for medium.
* Bi-prediction with CU-level weights (BCW) added as new coding tool
* Improved 2-pass Rate Control (RC) and Screen Content Coding
* Various bugfixes and improvements

Without QP adaptation for subjective optimization and 8 threads the following PSNR-based YUV BD-rates compared to HM-16.23 (GOP16+MCTF) as well as speedup factors compared to HM-16.23 and VTM-13.0 (GOP32+MCTF) are reported for different presets:

Faster HD: −2.5%, 160x (HM), 1200x (VTM) UHD: −9.3%, 190x (HM), 1300x (VTM)

Fast HD: −21.7%, 83x (HM), 630x (VTM) UHD: −23.8%, 97x (HM), 700x (VTM)

Medium HD: −31.3%, 23x (HM), 170x (VTM) UHD: −34.3%, 32x (HM), 230x (VTM)

Slow HD: −35.1%, 7.7x (HM), 59x (VTM) UHD: −38.0%, 11x (HM), 76x (VTM)

Slower HD: −38.7%, 1.7x (HM), 13x (VTM) UHD: −41.5%, 2.3x (HM), 16x (VTM)

With QP adaptation for subjective optimization and 8 threads, the following MS-SSIM-based YUV BD-rates compared to HM-16.23 (GOP16+MCTF) as well as a speedup factors compared to HM-16.23 and VTM-13.0 (GOP32+MCTF) are reported for different presets:

Faster HD: −14.4%, 150x (HM), 1200x (VTM) UHD: −11.0% 170x (HM), 1400x (VTM)

Fast HD: −28.4%, 80x (HM), 630x (VTM) UHD: −26.2% 93x (HM), 750x (VTM)

Medium HD: −36.3%, 22x (HM), 180x (VTM) UHD: −36.9%, 30x (HM), 240x (VTM)

Slow HD: −39.9%, 7.6x (HM), 60x (VTM) UHD: −40.7%, 10x (HM), 82x (VTM)

Slower HD: −42.5%, 1.6x (HM), 13x (VTM) UHD: −43.9%, 2.2x (HM), 18x (VTM)

Main changes for VVdeC v1.2.0 since version 1.1.2 include:

* WebAssembly to enable decoding and playback in web browsers.
* Various fixes and improvements:
  + Added and improved some SSE41 SIMD
  + Minor improvements to ALF, DMVR and LMCS processing
  + Added option to remove padding from YUV on output

Real-time playback in browser plugin up to 1080p60.

YUV PSNR is 6:1:1 using weights, same for MS-SSIM

[JVET-X0104](https://jvet-experts.org/doc_end_user/current_document.php?id=11097) Update on the progress of the optimized VVC encoder implementation, Ali266 [X. Dong, S. Fang, Z. Huang, J. Liu, S. Xu, R. Yang, L. Yu, J. Chen, R.-L. Liao, Y. Ye (Alibaba)]

software VVC encoder developed by Alibaba. Ali266 was first presented in JVET-W0127 [1], it supports 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 this document, we report further improvement in Ali266’s encoding speed at its fast preset, and also the addition of 8-bit encoding to Ali266 with the continued support of 10-bit encoding. We report that a 1.9x encoding speedup has been achieved in the past three months for video sequences at 1920x1080 or 1080x1920 resolutions. Currently, when configured with fast preset, Ali266 achieves average encoding speeds of 36.3 fps for classes 1920x1080 / 1080x1920 and 63.4 fps for classes 1280x720 / 720x1080. Based on such data, it is our new conclusion that Ali266 is a VVC encoder capable of real-time encoding for 1080p sequences, and Ali266 is more ready for real-world commercial applications.

Besides the speed improvement, Ali266’s encoding performance is also improved. For the JVET CTC content, Ali266 at slow preset achieves BD rates for {Y, U, V, YUV} of {10.95%, -6.11%, -4.50%, 7.11%} under the Random Access configuration with an average speedup factor of 227 times over VTM-13.0. For the internal VOD content, Ali266 at fast preset achieves BD rates for {Y, U, V, YUV} of {-36.25%, -45.73%, -43.02%, -38.10%} over the widely used open source HEVC software encoder x265 at medium preset, which is better than the result that was previously reported as {-32.87%, -46.07%, -42.70%, -35.35%} in JVET-W0127 over the same anchor.

How many threads were used? 12.

Typical bit rates for fast preset where own content was used? Not exactly known.

Fast preset with 8 or 10 bit? Both is possible, but results based on 8 bit.

Fast preset still slower than X265 with medium preset (about factor 2).

## Complexity analysis (0)

Contributions in this area were discussed in session x at XXXX–XXXX UTC on XXday X Oct. 2021 (chaired by XXX).

## AHG10: Encoding algorithm optimization (3)

Contributions in this area were discussed in session 20 at 1300–1350 UTC on Wednesday 13 Oct. 2021 (chaired by JRO).

[JVET-X0167](https://jvet-experts.org/doc_end_user/current_document.php?id=11175) Cross-check of JVET-X0061: AHG10: Fast skip of TT split partitioning on top of VTM-14.0 and ECM reference software [A. Henkel (HHI)] [late]

No need for presentation, relates to a withdrawn contribution.

[JVET-X0063](https://jvet-experts.org/doc_end_user/current_document.php?id=11056) AHG10: Deblocking filter setting for VTM [H. Zhang, X. Li, S. Liu (Tencent)]

In this contribution, a set of encoder only settings of deblocking filter are proposed for VTM. Basically, the deblocking filter control parameters BetaOffset\_div2 and TcOffset\_dive2 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-14.0, the overall performance are as follows:

* AI: -0.74%/-0.69%/-0.44% with 99%EncT/99%DecT
* RA: -0.32%/-0.58%/-0.50% with 100%EncT/100%DecT
* LDB: -0.62%/-0.41%/0.02% with 100%EncT/100%DecT
* LDP: -0.34%/-0.26%/-0.09% with 100%EncT/100%DecT

In ECM, a similar setting is used, but it is further controlled by additional syntax elements. In VTM, the different settings for different temporal layers can be controlled via slice header. One intent of the proposal is making the comparison of VTM and ECM more fair.

Visual quality impact should be investigated before such a change is applied. Deblocking settings are known to be sensitive in terms of visual quality, even when showing BD gains.

Only requires a config file change, no encoder software change.

It is pointed out that also ECM has not been investigated in terms of visual quality, but even if visual problems were found in VTM, it might not be the case in ECM. Both VTM and ECM should be checked regarding visual impact of the method.

Further study (coordinate with M. Wien to prepare a visual investigation at next meeting).

[JVET-X0171](https://jvet-experts.org/doc_end_user/current_document.php?id=11179) Crosscheck of JVET-X0063 (AHG10: Deblocking filter setting for VTM) [[H.-J. Jhu (Kwai)](mailto:jhuhong-jheng@kwai.com)] [late]

[JVET-X0116](https://jvet-experts.org/doc_end_user/current_document.php?id=11109) AHG10: Suggestion to enable GOP-based temporal filtering for low-delay configurations in CTC for HM and VTM [K. Andersson, J. Enhorn, P. Wennersten (Ericsson)]

GOP-based temporal filtering / Motion Compensated Temporal Filtering (MCTF) is enabled in CTC for random-access for HM and VTM. For low-delay MCTF has so far been disabled due to bad performance for smaller resolutions. In this contribution we show that by enabling MCTF for low-delay configurations in HM-16.24rc1 and VTM-14.0 a BD-rate impact of -0.7% and -0.8% can be achieved for HM respectively VTM for low-delay B and -1.1% for low-delay P. We have also checked visually that it works well.

We therefore suggest to also enable temporal filtering for low-delay configurations in CTC for HM and VTM.

Only past frames are used in low delay, so it is still low delay. It is configurable in the software

The MCTF software is not changed, only parameters are modified. It is pointed out that the newest version of MCTF had never been investigated for low delay configs.

It should be considered to use this also in ECM for low delay configurations.

Visual quality is very similar (perhaps minor improvement in some cases). The gain is lower than in RA; so it might be less likely to see visual quality impact.

Decision(CTC): Adopt JVET-X0116 for both HM and VTM for LDB and LDP configurations.

[JVET-X0183](https://jvet-experts.org/doc_end_user/current_document.php?id=11193) Cross-check of JVET-X0116 (AHG10: Suggestion to enable GOP-based temporal filtering for low-delay configurations in CTC for HM and VTM) [[A. Wieckowski (HHI)](mailto:adam.wieckowski@hhi.fraunhofer.de)] [late]

[JVET-X0143](https://jvet-experts.org/doc_end_user/current_document.php?id=11136) AHG10: VTM Encoder Changes for ALF Usage in Viewport-adaptive Streaming [A. Zare, A. Aminlou, A. Hallapuro, M. M. Hannuksela (Nokia)]

This document proposes several VTM software changes which enable the use of ALF in applications where extraction and merging of subpictures occur, e.g., in subpicture-based viewport adaptive streaming (VAS). Currently, ALF should be disabled in such applications when encoding using VTM software. The proposed changes include: 1) making the maximum number of ALF APSs per picture configurable, 2) allowing both luma and chroma in case of having one ALF APS, and 3) introducing an offset value to the ALF APS identifiers. This offset along with limiting the maximum number of ALF APSs per picture allows controlling the allocation of ALF APS identifiers in sub-streams which may be merged into a single stream. Moreover, the contribution proposes a VTM encoder configuration that enables a constant joint chroma coding residual (JCCR) sign, since JCCR sign is a picture-level parameter and hence needs to be the same in all encoded sub-streams which may be merged into a single stream. Furthermore, the contribution proposes several changes to the VTM subpicture merger tool to improve the performance of merging operation. Finally, the contribution illustrates the benefit of enabling ALF in a subpicture-based VAS use case.

It is confirmed by cross-checker and software coordinators that the code is clean and appropriate.

Several experts supported that this functionality is important for subpicture merging, or parallel encoding of subpictures.

It is also desirable that the software supports as much as possible the capabilities of VVC, and does not limit number of APSs below the number allowed in VVC profiles.

No impact on any CTC

Decision (SW): Adopt JVET-X0143

[JVET-X0201](https://jvet-experts.org/doc_end_user/current_document.php?id=11211) Crosscheck of JVET-X0143 (AHG10: VTM Encoder Changes for ALF Usage with Subpicture) [K. Sühring (HHI)] [late]

## Profile/tier/level specification (7)

Contributions in this area were discussed in session 6 at 0720–0920 UTC on Thursday 7 Oct. 2021 (chaired by JRO), at 0500-0550 on Friday 8 Oct. 2021 (chaired by JRO), and at 0835–0920 UTC on Tuesday 12 Oct. 2021 (chaired by JRO)

[JVET-X0073](https://jvet-experts.org/doc_end_user/current_document.php?id=11066) AHG2: On specifying the range extensions profiles [Y.-K. Wang (Bytedance)]

This contribution discusses a few asserted issues in the current text for or related to specifying the range extensions profiles, and proposes the following changes:

1. Change the value range for sps\_bitdepth\_minus8 in its semantics to be 0..8 (Agreed – same change also needs to be applied in the corresponding VPS syntax element).
2. Add the following constraint: In a bitstream conforming to the Main 12 Intra, Main 12 4:4:4 Intra, or Main 16 4:4:4 Intra profile, the value of sh\_slice\_type shall be equal to 2 for all slices (Agreed in principle, but it is no longer necessary having this constraint due to the constraint proposed in X0106).
3. The specified general\_profile\_idc values for all the range extensions profiles are increased by 1 (It was later agreed to change the LSBs for 16 bit from 100 to 011 which allows a more flexible signalling for possible future extension – reserve all three LSBs for possibly signalling up to 8 different bit depths – see X0073 v2).
4. Replace the current decoder capability requirements for the range extension profiles with the following (Agreed on this – necessary to specify that corresponding v1 profile bitstreams need to be decoded):

Decoders conforming to a format range extensions profile at a specific level (identified by a specific value of general\_level\_idc) of a specific tier (identified by a specific value of general\_tier\_flag) shall be capable of decoding all bitstreams and sublayer representations for which all of the following conditions apply:

* Any of the following conditions apply:
  + The decoder conforms to the Main 12 profile, and the bitstream is indicated to conform to the Main 10, Main 10 Still Picture, Main 12, Main 12 Intra, or Main 12 Still Picture profile.
  + The decoder conforms to the Main 12 4:4:4 profile, and the bitstream is indicated to conform to the Main 10, Main 10 Still Picture, Main 10 4:4:4, Main 10 4:4:4 Still Picture, Main 12, Main 12 Intra, Main 12 Still Picture, Main 12 4:4:4, Main 12 4:4:4 Intra, or Main 12 4:4:4 Still Picture profile.
  + The decoder conforms to the Main 16 4:4:4 profile, and the bitstream is indicated to conform to Main 10, Main 10 Still Picture, Main 10 4:4:4, Main 10 4:4:4 Still Picture, or any of the format range extensions profile.
  + The decoder conforms to the Main 12 Intra profile, and the bitstream is indicated to conform to the Main 10 Still Picture, Main 12 Intra, or Main 12 Still Picture profile, or the gci\_all\_rap\_pictures\_flag is equal to 1, and the bitstream is indicated to conform to Main 10 or Main 12 profile. (similar for all other intra profiles in this bullet list).
  + The decoder conforms to the Main 12 4:4:4 Intra profile, and the bitstream is indicated to conform to the Main 10 Still Picture, Main 10 4:4:4 Still Picture, Main 12 Intra, Main 12 4:4:4 Intra, Main 12 Still Picture, or Main 12 4:4:4 Still Picture profile.
  + The decoder conforms to the Main 16 4:4:4 Intra profile, and the bitstream is indicated to conform to the Main 10 Still Picture, Main 10 4:4:4 Still Picture, Main 12 Intra, Main 12 4:4:4 Intra, Main 16 4:4:4 Intra, Main 12 Still Picture, Main 12 4:4:4 Still Picture, or Main 16 4:4:4 Still Picture profile.
  + The decoder conforms to the Main 12 Still Picture profile, and the bitstream is indicated to conform to the Main 10 Still Picture or Main 12 Still Picture profile.
  + The decoder conforms to the Main 12 4:4:4 Still Picture profile, and the bitstream is indicated to conform to the Main 10 Still Picture, Main 10 4:4:4 Still Picture, Main 12 Still Picture or Main 12 4:4:4 Still Picture profile.
  + The decoder conforms to the Main 16 4:4:4 Still Picture profile, and the bitstream is indicated to conform to the Main 10 Still Picture, Main 10 4:4:4 Still Picture, Main 12 Still Picture, Main 12 4:4:4 Still Picture, or Main 16 4:4:4 Still Picture profile.
* The bitstream is indicated to conform to a tier that is lower than or equal to the specified tier.
* The bitstream is indicated to conform to a level that is not level 15.5 and is lower than or equal to the specified level.

Decoders conforming to the Main 12 Still Picture profile at a specific level of a specific tier shall also be capable of decoding of the first picture of a bitstream when both of the following conditions apply:

* That bitstream is indicated to conform to the Main 10, Main 12, or Main 12 Intra profile, to conform to a tier that is lower than or equal to the specified tier, and to conform to a level that is not level 15.5 and is lower than or equal to the specified level.
* That picture is an IRAP picture or is a GDR picture with ph\_recovery\_poc\_cnt equal to 0, is in an output layer, and has ph\_pic\_output\_flag equal to 1.

Decoders conforming to the Main 12 4:4:4 Still Picture profile at a specific level of a specific tier shall also be capable of decoding of the first picture of a bitstream when both of the following conditions apply:

* That bitstream is indicated to conform to the Main 10, Main 10 4:4:4, Main 12, Main 12 Intra, Main 12 4:4:4, or Main 12 4:4:4 Intra profile, to conform to a tier that is lower than or equal to the specified tier, and to conform to a level that is not level 15.5 and is lower than or equal to the specified level.
* That picture is an IRAP picture or is a GDR picture with ph\_recovery\_poc\_cnt equal to 0, is in an output layer, and has ph\_pic\_output\_flag equal to 1.

Decoders conforming to the Main 16 4:4:4 Still Picture profile at a specific level of a specific tier shall also be capable of decoding of the first picture of a bitstream when both of the following conditions apply:

* That bitstream is indicated to conform to the Main 10, Main 10 4:4:4, Main 12, Main 12 Intra, Main 12 4:4:4, Main 12 4:4:4 Intra, Main 16 4:4:4, or Main 16 4:4:4 Intra profile, to conform to a tier that is lower than or equal to the specified tier, and to conform to a level that is not level 15.5 and is lower than or equal to the specified level.
* That picture is an IRAP picture or is a GDR picture with ph\_recovery\_poc\_cnt equal to 0, is in an output layer, and has ph\_pic\_output\_flag equal to 1.

1. Change the caption of Table A.1 to be "Greatest allowed values for syntax elements in the format range extensions profiles" and just include the greatest allowed value in each cell (Agreed to specify ranges of allowed values instead - this approach is better than clarifying the meaning of “-“ as suggested in X0093 proposal 3).
2. Discuss whether to rename and flip the meaning of the general\_lower\_bit\_rate\_constraint\_flag and merge Table A.2 into Table A.1. If not, change the caption of Table A.2 to be "Allowed values for general\_lower\_bit\_rate\_constraint\_flag for format range extensions profiles" (Agreed to use the first approach and merge tables. Later, it was decided to remove that flag).

The exact text changes for the proposed changes are included in an attachment to this contribution, with change marks relative to JVET-W2005-v2.

It is further discussed that additional consideration may be useful if picture reordering is at all necessary in case of intra profiles.

[JVET-X0106](https://jvet-experts.org/doc_end_user/current_document.php?id=11099) On constraints for intra profiles in VVC [T. Tsukuba, S. Keating (Sony)]

This contribution proposes adding constraints on slice\_type for Intra profiles in VVC.

First aspect was resolved by imposing ph\_inter\_slice\_allowed\_flag=0. X0106 had proposed to use a GCI mechanism which is a less appropriate solution.

Second aspect (carried over from HEVC) is to only allow IRAP NAL units in intra profiles.

It is agreed to include the following text or editorial equivalent:

“In bitstreams conforming to the Main 12 Intra, Main 12 4:4:4 Intra, or Main 16 4:4:4 Intra profiles, all pictures shall be IRAP pictures or GDR pictures with recovery\_poc\_count=0.”

There is a further editorial suggestion in the proposal for alignment with Main10 profile definition, which should also be integrated.

[JVET-X0075](https://jvet-experts.org/doc_end_user/current_document.php?id=11068) AHG8: Level refinement for VVC operation range extension profiles [T. Ikai, T. Chujoh, T. Aono (Sharp)] [late]

This contribution proposes to refine level definition for VVC operation range profiles about HbrFactor and CpbVcl/NalFactor and MinCrScaleFactor.

Proposes to inherit the values of the factors for 12 bit from HEVC (increase by factor 1.5).

[JVET-X0079](https://jvet-experts.org/doc_end_user/current_document.php?id=11072) Proposals on maximum bit rate for HEVC and VVC [T. Tsukuba, M. Ikeda, T. Suzuki (Sony)]

This contribution proposes to add new code points to support higher max bit rate for Main 4:2:2 and 4:4:4 Intra profiles for both HEVC and VVC. The authors believe that the current max bit rate is not sufficient to support high end production applications. Therefore, the additional code points to support higher max bit rate are proposed. The modifications to the spec text are summarized in the Annex. This proposal does not affect on the current existing code points.

Current max bit rates not sufficient for 8K in particular in intra profiles (as shown in JVET-X0109).

* Option 1: Modify the calculation to derive max bit rate
* Option 2: Add new “tier” above high tier
* Option 3: Add new level to support higher bit rate

Option 1 and option two require syntax change (identical syntax but different semantics)

It is discussed if the following solution would resolve the problem:

For HEVC: Introducing a new level 6.3 (option 3)

For VVC v2 draft: Modify 6.3 for high tier (6.3 is new in v2), and defining different scaling factors and different max bit rate for high tier (also different for intra and inter), and in the latter case removing the lower\_bitrate\_constraint flag

(For HEVC, the lower\_bitrate\_constraint flag exists and needs to be used in case that the bit rate is not sufficient in high tier. This means that a decoder can only know what its capabilities must be when it decodes that flag)

Text review for VVC v2 according to the suggested method above was done in session 17 (v3 of X0079)

GCI syntax needs to be aligned with the other flags that are added (as per X0076 adoption)

All GCI syntax elements should start with gci\_...

If gci\_all\_rap\_pictures\_flag is true, a sentence must be added that an intra profile decoder must be able decoding it.

Decision: Adopt JVET-X0079 (v3 with the suggested refinements above)

Further, it was agreed to target new edition of HEVC, including new level, and errata items. Could start as DIS in January (plus request), FDIS July 2022. Draft of new level (only as JVET document)

[JVET-X0108](https://jvet-experts.org/doc_end_user/current_document.php?id=11101) Coded Picture Buffer sizes and MinCr for high bit-depth profiles [S. Keating, A. Browne, K. Sharman (Sony)]

This contribution describes proposed changes to maximum coded picture buffer (CPB) size for 12-bit profiles in order to guarantee that the buffer can always store a full picture when compressed at the minimum compression ratio (MinCr).

Similar to X0075, but slightly larger values.

It is pointed out during the discussion that the values from HEVC (factor 1.5 for 12 bit relative to 10 bit) might be challenging for 12 bit consumer applications. A possible solution might be using a factor of 1.2 fo main tier and 1.5 (or even higher) for high tier.

[JVET-X0109](https://jvet-experts.org/doc_end_user/current_document.php?id=11102) On maximum bit-rates for high bit-depth profiles [S. Keating, A. Browne, K. Sharman (Sony)]

This contribution compares the maximum specified bit-rates at certain levels of VVC with the bit-rates and PSNRs of VTM14.0 for appropriate CTCs. Comparison is also made with HEVC/HM bit-rates.

Powerpoint presentation should be provided – all plots shown in it are also in the xls spreadsheets.

Maximum bitrates of high tier are found to be sometimes not sufficient for enough quality (PSNR below 40 dB), in particular for HLG and SVT content.

It is mentioned that the general\_lower\_bitrate\_constraint\_flag from X0073 might also allow definition of a kind of “super tier”, however would be in SPS and coming later than profile/level/tier information.

[JVET-X0187](https://jvet-experts.org/doc_end_user/current_document.php?id=11197) Inference rule on general\_lower\_bit\_rate\_constraint\_flag [T. Ikai (Sharp)] [late]

No need to be presented, as the flag would be removed (see notes under X0079).

## Proposed modification of system interface (0)

Contributions in this area were discussed in session x at XXXX–XXXX UTC on XXday X Oct. 2021 (chaired by XXX).

# Low-level tool technology proposals (84)

## AHG8: High bit rate and high bit depth coding for VVC (5)

### General (0)

Contributions in this area were discussed in session x at XXXX–XXXX UTC on XXday X Oct. 2021 (chaired by XXX).

### Entropy Coding for High Bit Depth and High Bit Rate Coding (5)

Contributions in this area were discussed in session 9 at 0600–0715 and session 10 at 0735–0800 UTC on Friday 8 Oct. 2021 (chaired by JRO).

[JVET-X0127](https://jvet-experts.org/doc_end_user/current_document.php?id=11120) AHG8: Modification of History Based Rice Parameter Derivation [Y. Yu, H. Yu, Z. Xie, F. Wang, L. Xu, D. Wang (OPPO)]

In the history-based Rice parameter derivation method adopted in JVET-V meeting (JVET-V0106), a history value is transferred through a counter to update localSumAbs so that a more accurate Rice parameter can be derived to code remaining level. The counter may be updated once per TU and the updated counter will be used to derive history value for next TU. This contribution proposes a modification to the Rice parameter derivation process. Once the counter is updated, a history value is derived and used immediately for coding the remaining level within the current TU till a new history value is derived. There is no any complexity increase and the specification change is also very simple. Under the high bit-depth common test condition (CTC), the simulation results show the following coding gains in BD-bitrate.

Low QP:

PQ: AI: -0.01%, -0.01%, -0.02%; LDB: 0.00%, -0.02%, 0.01%; RA: -0.01%, -0.01%, -0.01%;

HLG: AI: -0.02%, -0.01%, -0.02%; LDB: 0.00%, -0.01%, 0.01%; RA: 0.00%, 0.00%, 0.00%;

SVT: AI: -0.03%, -0.03%, -0.03%; LDB: -0.02%, -0.02%, -0.02%; RA: -0.02%, -0.02%, -0.02%;

Lossless:

PQ: AI: -0.03%; LDB: -0.04%, RA: -0.04%;

HLG: AI: -0.02%; LDB: -0.02%, RA: -0.02%;

SVT: AI: -0.02%; LDB: -0.02%, RA: -0.02%;

It is asserted (according to the crosschecker’s and other experts opinion) that this proposal introduces an additional dependency on previously decoded coefficients in same TU, which is undesirable for pipelining and parallelization of the decoding process.

The very small additional compression gain does not justify this.

No action.

[JVET-X0157](https://jvet-experts.org/doc_end_user/current_document.php?id=11164) Cross-check report of JVET-X0127: AHG8: Modification of History Based Rice Parameter Derivation [D. Rusanovskyy (Qualcomm)] [late]

[JVET-X0128](https://jvet-experts.org/doc_end_user/current_document.php?id=11121) AHG8: On History-Based Rice Parameter Derivations for Wavefront Parallel Processing [Y. Yu, H. Yu, Z. Xie, F. Wang, L. Xu, D. Wang (OPPO)]

A history-based Rice parameter derivation method (JVET-V0107) was adopted in JVET V-meeting. A counter per color component is utilized to maintain a history record with data relating to the Rice parameter and may be updated TU by TU. The counter in the preceding CTU row will be passed to the first TU in the current CTU row. As a result, this process will break the parallelism of WPP when WPP is enabled at the same time. This contribution proposes two methods to address the issue. The first method proposes that the counter per color component is saved after the last TU of the first CTU in each CTU row is coded. For each CTU row except the first CTU row, prior to the first TU coding, the synchronization process is applied, and the counter is synchronized with the saved counter from the preceding CTU row. The second method proposes that a fixed initial value for counter is used prior to the first TU coding in each CTU row. Under the high bit-depth common test condition (CTC), the simulation results show minor loss in BD-bitrate for both methods. Comparing method 2 with method 1, method 2 does not require any extra memory to save extra variable while method 2 has the same coding performance with method 1.

Method 1:

Low QP

PQ: AI: 0.04%, 0.04%, 0.02%; LDB: 0.09%, 0.08%, 0.03%; RA: 0.09%, 0.08%, 0.01%;

HLG: AI: 0.02%, 0.01%, 0.01%; LDB: 0.08%, 0.04%, 0.00%; RA: 0.10%, 0.06%, 0.02%;

SVT: AI: 0.01%, 0.01%, 0.01%; LDB: 0.02%, 0.02%, 0.01%; RA: 0.03%, 0.02%, 0.01%;

Lossless:

PQ: AI: 0.01%; LDB: 0.01%; RA: 0.01%;

HLG: AI: 0.02%; LDB: 0.03%; RA: 0.03%;

SVT: AI: 0.01%; LDB: 0.02%; RA: 0.02%;

Method 2:

Low QP

PQ: AI: 0.04%, 0.04%, 0.02%; LDB: 0.09%, 0.08%, 0.03%; RA: 0.09%, 0.08%, 0.01%;

HLG: AI: 0.02%, 0.01%, 0.01%; LDB: 0.08%, 0.04%, 0.00%; RA: 0.10%, 0.06%, 0.02%;

SVT: AI: 0.01%, 0.01%, 0.01%; LDB: 0.02%, 0.02%, 0.01%; RA: 0.03%, 0.02%, 0.01%;

Lossless:

PQ: AI: 0.01%; LDB: 0.01%; RA: 0.01%;

HLG: AI: 0.02%; LDB: 0.03%; RA: 0.03%;

SVT: AI: 0.01%; LDB: 0.02%; RA: 0.02%;

It is reported that method 1 is already implemented in software but missing in the specification text. However, method 2 is simpler in particular in terms of text.

Decision: Adopt JVET-X0128 method 2

[JVET-X0158](https://jvet-experts.org/doc_end_user/current_document.php?id=11165) Cross-check report of JVET-X0128: AHG8: On History-Based Rice Parameter Derivations for Wavefront Parallel Processing [D. Rusanovskyy (Qualcomm)] [late]

[JVET-X0129](https://jvet-experts.org/doc_end_user/current_document.php?id=11122) AHG8: Independent Rice Parameter Derivation for high bit depth and high bit rate extensions [Y. Yu, H. Yu, Z. Xie, F. Wang, L. Xu, D. Wang (OPPO)]

In the history-based Rice parameter derivation method adopted in JVET-V meeting, a history value is transferred through a counter to update localSumAbs so that a more accurate Rice parameter can be derived to code remaining levels. The counter may be updated once in each TU based upon its value in the previous TU. This results in dependency among TUs. This contribution proposes an independent Rice parameter derivation method. If the last significant coefficient in a TU has abs\_remainder[] part in its level coding, a fixed value is used as HistValue in the computation of localSumAbs. After the last significant coefficient is coded, the absolute value of last significant coefficient is used, as a history value, to compute localSumAbs for coding all remaining abs\_remainder[ ] or dec\_abs\_level[ ] within the current TU. The part of the counter StatCoeff updating process that uses its value carried over from the previous TU is totally removed and therefore there is no any dependency between the current TU and the previous TUs. Under the high bit-depth common test condition (CTC), the simulation results show the following minor coding loss in BD-bitrate.

Low QP

PQ: AI: 0.08%, 0.04%, 0.04%; LDB: 0.01%, 0.00%, 0.02%; RA: 0.02%, 0.03%, 0.01%;

HLG: AI: 0.09%, 0.04%, 0.03%; LDB: 0.01%, 0.00%, 0.01%; RA: 0.02%, 0.02%, 0.02%;

SVT: AI: 0.15%, 0.10%, 0.10%; LDB: 0.07%, 0.07%, 0.07%; RA: 0.06%, 0.06%, 0.06%;

Lossless:

PQ: AI: 0.55%; LDB: 0.48%, RA: 0.45%;

HLG: AI: 0.53%; LDB: 0.51%, RA: 0.50%;

SVT: AI: 0.63%; LDB: 0.31%, RA: 0.30%;

The dependency problem in current method is not critical, and the proposal would incur losses, which are not negligible in particular for the case of lossless coding. No support to make such a change by other experts.

No action.

[JVET-X0159](https://jvet-experts.org/doc_end_user/current_document.php?id=11166) Cross-check report of JVET-X0129: AHG8: Independent Rice Parameter Derivation for high bit depth and high bit rate extensions [D. Rusanovskyy (Qualcomm)] [late]

[JVET-X0136](https://jvet-experts.org/doc_end_user/current_document.php?id=11129) AHG8: On significance, GT1, and GT2 flag coding for high bit depths [A. Browne, S. Keating, K. Sharman (Sony)]

This document describes a study of the use of significance, GT1 and GT2 flags in regular residual coding at low QP and lossless operating points for bit depths beyond 10 bit. An analysis using VTM 14.0 is provided which indicates the fraction of these flags that are set for a range of different QPs within and beyond the low QP high bit depth CTC. This analysis outlines how as QPs decrease, the fraction of the flags that are set increases to almost 100%. This results in the number of coefficients at least partially coded using flags reduces as QPs become smaller. A modification is described which effectively shifts the significance, GT1 and GT2 tests left by a number of bits signalled from the encoder. A non-normative algorithm is also described which computes a suitable value for this shift in a pre-encode step. The analysis, previously done for VTM 14.0, is repeated for the modification and indicates that the fraction of flags that are set does not rise as QPs decrease, unlike VTM 14.0. Consequently the number of coefficients coded using the significance, GT1 and GT2 flags is maintained.

The following results are reported for the modification against an anchor of VTM 14.0:

|  |  |  |  |
| --- | --- | --- | --- |
| Lossless | AI | LDB | RA |
| PQ | -0.20% | -0.24% | -0.24% |
| HLG | -0.83% | -0.92% | -0.92% |
| SVT-12 | -1.17% | -1.35% | -1.37% |
| SVT-16 | -1.18% | -1.51% | -1.53% |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Low QP | AI | | | LDB | | | RA | | |
|  | Y/G | U/B | V/R | Y/G | U/B | V/R | Y/G | U/B | V/R |
| PQ | -0.01% | 0.02% | 0.02% | 0.02% | 0.01% | 0.03% | 0.02% | 0.02% | 0.03% |
| HLG | -0.07% | -0.08% | -0.08% | 0.02% | 0.02% | 0.02% | 0.02% | 0.03% | 0.04% |
| SVT-12 | -0.20% | -0.19% | -0.21% | -0.05% | -0.05% | -0.05% | -0.04% | -0.02% | -0.02% |
| SVT-16 | -0.75% | -0.77% | -0.78% | -0.76% | -0.73% | -0.73% | -0.71% | -0.67% | -0.67% |

It is shown that the number of significant, gt1, gt2 by tendency approaches 100% when QP decreases to extreme low (negative) QPs. Even though CABAC should work very efficient in that case, the number of context coded bins would grow very large but anyway reaches the limits, and number of escape coded bins grows very large.

The analysis shows that significance, gt1 and gt2 coding does not work efficiently at low bit rates.

The proposal improves significantly in particular for lossless coding, but also would require significant changes to regular residual coding.

Probably, even more efficient solutions could be found for lossless coding with a design specifically dedicated for that – but globally lossless coding is not a main target of VVC.

Not for v2 – further study recommended.

[JVET-X0173](https://jvet-experts.org/doc_end_user/current_document.php?id=11181) Cross-check of JVET-X0136: AHG8: On significance, GT1, and GT2 flag coding for high bit depths [[Z. Xie](mailto:xiezhihuang@oppo.com), [Y. Yu (OPPO)](mailto:yue.yu@oppo.com)] [late]

[JVET-X0137](https://jvet-experts.org/doc_end_user/current_document.php?id=11130) AHG8 and AHG10: On derivation of sh\_reverse\_last\_sig\_coeff\_flag and sh\_ts\_residual\_coding\_rice\_idx\_minus1 [A. Browne, S. Keating, K. Sharman (Sony)]

This document describes methods for determining the signalled values **sh\_reverse\_last\_sig\_coeff\_flag** and **sh\_ts\_residual\_coding\_rice\_idx\_minus1** which are used to decode coefficient values in the tools adopted in VVC v2 described in JVET-V0121 and JVET-W0046.

The methods described here operate on the original image data and therefore can be used before coefficients are coded. This allows the values of the signalled parameters to be determined using image data for the first frame of a sequence unlike the current implementations in VTM which rely on QP and bit depth based formulas for the first frame of a sequence.

Gains are reported as follows

For the **sh\_reverse\_last\_sig\_coeff\_flag** modification:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Low QP | AI | | | LDB | | | RA | | |
|  | Y/G | U/B | V/R | Y/G | U/B | V/R | Y/G | U/B | V/R |
| PQ | 0.01% | -0.01% | 0.00% | 0.02% | 0.00% | 0.03% | 0.01% | 0.00% | 0.01% |
| HLG | -0.01% | -0.02% | -0.02% | -0.03% | -0.13% | -0.06% | -0.03% | -0.10% | -0.05% |
| SVT-12 | 0.00% | 0.00% | 0.00% | -0.01% | -0.01% | 0.01% | -0.01% | -0.01% | -0.01% |
| SVT-16 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Mid QP | AI | | | LDB | | | RA | | |
|  | Y/G | U/B | V/R | Y/G | U/B | V/R | Y/G | U/B | V/R |
| PQ | -0.04% | 0.02% | -0.08% |  |  |  |  |  |  |
| HLG | -0.01% | 0.00% | 0.00% |  |  |  |  |  |  |
| SVT-12 | -0.05% | -0.03% | -0.04% |  |  |  |  |  |  |
| SVT-16 | 0.00% | 0.00% | 0.00% |  |  |  |  |  |  |

For the **sh\_ts\_residual\_coding\_rice\_idx\_minus1 modification**:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Low QP | AI | | | LDB | | | RA | | |
|  | Y/G | U/B | V/R | Y/G | U/B | V/R | Y/G | U/B | V/R |
| SCC | -0.39% | -0.38% | -0.37% |  |  |  |  |  |  |
| PQ | -0.01% | -0.02% | -0.03% | 0.01% | -0.02% | 0.00% | -0.01% | -0.01% | -0.01% |
| HLG | 0.00% | -0.02% | -0.02% | 0.00% | -0.01% | -0.02% | -0.01% | 0.01% | 0.00% |
| SVT-12 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| SVT-16 | 0.00% | 0.01% | 0.00% | 0.01% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Very | AI | | | LDB | | | RA | | |
| Low QP | Y/G | U/B | V/R | Y/G | U/B | V/R | Y/G | U/B | V/R |
| PQ | -0.07% | -0.31% | -0.31% |  |  |  |  |  |  |
| HLG | -0.40% | -0.75% | -0.73% |  |  |  |  |  |  |
| SVT-12 | -0.01% | -0.00% | -0.01% |  |  |  |  |  |  |
| SVT-16 | 0.00% | 0.00% | 0.00% |  |  |  |  |  |  |

Non-normative proposal (encoder only)

Parameters computed from current frame (rather than previous frame as the existing method does). “Pre-encoding”, histogram based on pixel data, no two-pass encoding

Very low QP: -23 … 2

Hardly any change for normal QP range

Effective mainly for 1st frame, where no adaptation is applied with current method. Results with anchor which is modified processing all I-frames as “1st frame”.

Algorithm could be just for first frame, AI coding, and random access points.

Decision(SW): Adopt JVET-X0137 (not replacing the existing method, but applying it for the frames where the existing method is not used)

[JVET-X0174](https://jvet-experts.org/doc_end_user/current_document.php?id=11182) Cross-check of JVET-X0137: AHG8 and AHG10: On derivation of sh\_reverse\_last\_sig\_coeff\_flag and sh\_ts\_residual\_coding\_rice\_idx\_minus1 [[Z. Xie](mailto:xiezhihuang@oppo.com), [Y. Yu (OPPO)](mailto:yue.yu@oppo.com)] [late]

### Adaptation of other tools for high bit rate and high bit depth (0)

Contributions in this area were discussed in session x at XXXX–XXXX UTC on XXday X Oct. 2021 (chaired by XXX).

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

### General (0)

Contributions in this area were discussed in session x at XXXX–XXXX UTC on XXday X Oct. 2021 (chaired by XXX).

### EE1 contributions: Neural network-based video coding (12)

Contributions in this area were discussed in session 3 at 1300–1415 UTC on Wednesday 6 Oct. 2021 (chaired by JRO), in session 7 at 1300-1510 UTC on Thursday 7 Oct. 2021 (chaired by JRO), and in session 11 at at 1300-1315 UTC on Friday 8 Oct. 2021 (chaired by JRO).

[JVET-X0023](https://jvet-experts.org/doc_end_user/current_document.php?id=11173) 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]

This contribution provides a summary report for the Exploration Experiment 1 on Neural Network-based Video Coding.In total 11 tests have been completed within this EE between the JVET-W and JVET-X meetings to study and evaluate NNVC technologies, analyze their performance and complexity aspects. Several variants of NN-based in-loop filters with complexity 30...800 kMAC/pxl providing average bit-saving in a range 2...10% (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 1...6% bit-saving in random access configuration for 4K content. An example of cross-platform bit-exact decodable implementation of NN-based Intra was demonstrated and cross-checked. This tool improves Intra coding by more than 3%.

1. **Test conditions and evaluation criteria**

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. It should be noticed that solutions presented in Super Resolution category are still far from target rates, and RD-curves often crossing. So, objective performance results for this category are less reliable than in other tests.

1. **Objective test results summary and observations**

Performance test results in RA configuration for all proposals are shown in Table 1.

**Table 1**. Test results in Random Access cfg (all tests).

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  | Random Access (AVG, 5QPs) | | | | | | | |
|  |  |  |  |  | PSNR | | | MS-SSIM | | | Time | |
| NN-filters | | | | | | | | | | | | |
| Doc# | Total CONV | Total memory, MB | kMAC /pxl | Pre-cision | Y | U | V | Y | U | V | ×Enc | ×Dec |
| JVET-X0110-1 | 6 | 3 | 151 | F | -2.5% | -6% | -4% | -1.9% | -7% | -4% | 1.2 | 392 |
| JVET-X0110-2 | 6 | 3 | 151 | F | -0.4% | 0% | -1% | -0.9% | -1.9% | -1.2% | 1.2 | 361 |
| JVET-X0065 | 37 | 146 | 520 | F | -10.1% | -24% | -23% | x | x | x | 1.8 | 390 |
| JVET-X0052 | 27 | 23 | 786 | F | -4.2% | -11% | -10% | -4.0% | -10% | -8% | 1.4 | 643 |
| JVET-X0140-1 | 35 | 2 | 34 | F | -5.8% | -7% | -5% | x | x | x | 1.2 | 41 |
| JVET-X0140-2 | 38 | 30 | 624 | F | -8.4% | -22% | -22% | x | x | x | 2.1 | 352 |
| JVET-X0053 | 13 | 0.2 | 29 | F | -1.1% | -5% | -4% | -1.2% | -5% | -3% | 1.1 | 12 |
| JVET-X0066 | 41 | 25 | 539 | F | -9.8% | -21% | -21% | x | x | x | 1.8 | 404 |
| Super Res | | | | | | | | | | | | |
| JVET-X0117 | 2 | 0.00 | 0.03 | int16 | -1.0% | 0.6% | 0.5% | -2.0% | -0.1% | -0.6% | 1.3 | 0.9 |
| JVET-X0064 | 36 | 22 | 764 | F | x | x | x | x | x | x | x | x |
| JVET-X0107 | 16 | 10 | 344 | F | 91% | 45% | 36% | x | x | x | 1.2 | 0.4 |
| JVET-X0074-1 | 10 | 1 | 85 | F | 58% | 71% | 66% | x | x | x | 0.8 | 0.5 |
| JVET-X0074-2 | 10 | 1 | 85 | F | 60% | 80% | 70% | x | x | x | x | x |
| NN-Intra | | | | | | | | | | | | |
| JVET-X0118-1 | 4 | 62 | 100 | F | -1.6% | -1% | -1% | -1.5% | 0% | 0% | 1.4 | 5 |
| JVET-X0118-3 | 4 | 31 | 100 | int16 | -1.6% | -1% | -1% | -1.5% | -0.1% | -0.1% | 1.1 | 3 |
| JVET-X0118-4 | 4 | 11 | 17 | int16 | -1.4% | -1% | -1% | -1.3% | 0% | 0% | 1.1 | 3 |

Following observations can be done.

Observation #1. As was suggested in [3] plot BD-rate gain vs kMAC/pxl and total memory size was made (Fig. 1). The dash line indicates capability of NVIDIA RTX3080 for 4K@60 processing. This graph is not directly comparable with one in [3], since average BD-rate gain was computed at slightly different conditions (not all data was available). But trend is very clear: proponents try to move their test point in this graph toward top-left corner. In [1] the solution with highest performance had ~1500 kMAC/pxl, currently the most complicated one is below 1000 kMAC/pxl, maintaining the gain at the level ~10% relatively to VTM.

Total memory size for all NN parameters is another important complexity measure. Almost all EE participants are switching between different NN models depending on situation (QP, block size or RDO choice). This adaptation is the source of additional performance improvements. Finding good balance between number of degrees of freedom and compression performance is an important aspect of AhG11/EE1 research. Fig. 2 shows compression performance gain vs total memory size needed for all NN parameters. For example, JVET-X0065 and JVET-X0066 are almost not distinguishable in Fig. 1, but form Fig. 2 one easily sees difference in terms of total memory size. Similar trend is also observed in Fig. 2, the test points are being pushed toward top-left corner so that less memory is required (the one with the largest memory size requires less than 150MB) but the majority of the gain is preserved.

**Fig. 1.** An average BD-rate gain in random access configuration vs kMAC/pxl.

**Fig. 2.** An average BD-rate gain in random access configuration vs totally memory size in MB.

Observation #2. There is strong correlation between kMAC/pxl number and decoder run-time increment factor relatively to VTM. Using fine-tuned model parameters leads to a slightly higher decoding run time, which is likely due to a more frequent use of the NN filter (Fig. 3).

**Fig. 3.** Decoding run-time increment factor (relatively to VTM) vs kMAC/pxl.

Observation #3. Filters with deeper neural network (higher total number of convolution layers) tend to have better complexity-performance trade-off (Fig. 4). Data shown only for NN-filter category, but NN-Intra show the same trend.

**Fig. 4.** An average BD-rate gain (for NN-filters) in random access configuration vs kMAC/pxl categorized depending on NN depth.

Observation #4. In NN-filter and NN-Intra categories BD-rate gain measured based on PSNR and MS-SSIM is very close. In super-resolution category MS-SSIM gain is almost twice higher than PSNR gain, indicating that MS-SSIM is less sensitive to re-sizing artifacts. It would be interesting to compare trend shown by BD-rate results in PSNR and MS-SSIM with visual tests results.

1. **Super resolution category specific**

Not all proponents is this category provided results for all test sequences, so not all can appear in graph on Fig. 1. Proposals in super-resolution category mostly useful for high resolution sequence coding. Table 2 and Fig. 5 summarize performance for UHD classes (A1, A2) in test set only.

Still comparison is not very accurate due to curves overlap. In Test 2.1 some efforts were taking to solve this problem. In order to avoid RD-curves overlap RPR was enabled only if it outperforms full-size coding (this is multi-pass encoding reflected in ×1.3 higher encoder run-time). In order to match target rate closer QP for down sampled sequence coding was selected per sequence, per parget rate. It should be discussed during the meeting if this type of solving RD-curves crossing problem can be recommended for further tests in Super-resolution category.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  | Random Access (4K, 5QPs) | | | | | | | |
| Labels on plot | Total CONV | Total memory, MB | kMAC /pxl | Pre-cision | PSNR | | | MS-SSIM | | | Time | |
| JVET-X0117 | 2 | 0.00 | 0.03 | int16 | -2.4% | 2% | 1% | -5.1% | 0% | -2% | 1.3 | 0.8 |
| JVET-X0064 | 34-36 | 22 | 764 | F | -6.1% | 6% | -2% | x | x | x | 1.0 | 0.3 |
| JVET-X0107 | 16 | 10 | 344 | F | -0.8% | 50% | 28% | x | x | x | 1.1 | 0.4 |
| JVET-X0074-1 | 10 | 1 | 85 | F | -3.9% | 28% | 14% | x | x | x | 0.8 | 0.5 |
| JVET-X0074-2 | 10 | 1 | 85 | F | -0.9% | 51% | 24% | x | x | x | 0.5 | 0.3 |

**Table 2**. Test results in for Random Access cfg (UHD only).

**Fig. 5.** Only UHD BD-rate gain in random access configuration vs kMAC/pxl.

1. **NN-Intra category specific**

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 Table 3. With not that deep NN-structure and moderate kMAC/pxl (and decoding run time). NN-Intra tools shows performance competitive with NN-filters in all-intra configuration test (Fig. 5).

In opposite to enchantment filters NN-Intra operates on block level. It is unlikely that NN-dedicated HW would be used inside the picture (if the rest of the picture doesn’t require NN-operations). Proponent of this EE test provided c++ compatible implementation of their NN algorithm (both float point and integer implementation are supported). This resulted in successful cross-check. It was observed that it is possible to get matched results only for integer implementation of NN-Intra. Also it was confirmed identical behavior of decoder under Linux and Windows for integer version of this EE test. NN-library [4] provided by proponent of this EE test and integer implantation of NN algorithms definitely can be recommended for further wide use in this EE.

BD-rate gain in “all-intra” test vs kMAC/pxl for NN-solution is different categories is shown on Fig. 6.

**Table 3**. Test results in for All Intra cfg (all ).

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  | All Intra (AVG, 5QPs) | | | | | | | |
|  |  |  |  |  | PSNR | | | MS-SSIM | | | Time | |
| NN-filters | | | | | | | | | | | | |
| Doc# | Total CONV | Total memory, MB | kMAC /pxl | Pre-cision | Y | U | V | Y | U | V | ×Enc | ×Dec |
| JVET-X0110-1 | 6 | 151 | 151 | F | x | x | x | x | x | x | x | x |
| JVET-X0110-2 | 6 | 151 | 151 | F | x | x | x | x | x | x | x | x |
| JVET-X0065 | 37 | 520 | 520 | F | -7.4% | -20% | -20% | x | x | x | 1.3 | 189 |
| JVET-X0052 | 27 | 786 | 786 | F | -5.3% | -10% | -10% | -4.8% | -9% | -9% | 1.5 | 387 |
| JVET-X0140-1 | 35 | 34 | 34 | F | -5.2% | -6% | -5% | x | x | x | 1.2 | 23 |
| JVET-X0140-2 | 38 | 624 | 624 | F | -7.6% | -18% | -20% | x | x | x | 1.8 | 203 |
| JVET-X0053 | 13 | 29 | 29 | F | -1.6% | -5% | -4% | -1.4% | -4% | -4% | 1.1 | 29 |
| JVET-X0066 | 41 | 539 | 539 | F | -7.4% | -19% | -19% | x | x | x | 1.4 | 238 |
| Super Res | | | | | | | | | | | | |
| JVET-X0117 | 2 | 0.03 | 0.03 | F | -0.5% | 0.9% | 0.7% | -1.5% | 0.6% | 0.3% | 1.0 | 0.9 |
| JVET-X0064 | 34-36 | 764 | 764 | F | x | x | x | x | x | x | x | x |
| JVET-X0107 | 16 | 344 | 344 | F | 29% | 20.3% | 2.4% | x | x | x | 1.3 | 0.5 |
| JVET-X0074-1 | 10 | 85 | 85 | F | 65% | 81% | 70% | x | x | x | 0.9 | 0.6 |
| JVET-X0074-2 | 10 | 85 | 85 | F | 65% | 95% | 73% | x | x | x | x | x |
| NN-Intra | | | | | | | | | | | | |
| JVET-X0118-1 | 4 | 62 | 100 | F | -3.3% | -3% | -3% | -3.3% | -2% | -2% | 3.7 | 40 |
| JVET-X0118-3 | 4 | 31 | 100 | int16 | -3.3% | -3% | -3% | -3.3% | -2% | -2% | 2.2 | 19 |
| JVET-X0118-4 | 4 | 11 | 17 | int16 | -2.9% | -2% | -2% | -2.9% | -2% | -2% | 2.2 | 18 |

**Fig. 6.** BD-rate gain in all-intra access configuration vs kMAC/pxl.

1. **Recommendation**

* Review all EE1 documents for more detailed information.
* Conduct BoG for preparing viewing test, better understanding test results and refining assessment methodology.
* Organize and conduct viewing during JVET-X meeting for prepared materials.

Comments/questions:

* For those who reported MSSIM (not mandatory), the difference of BD rate savings seem to be in a similar range as for PSNR.
* In SR, for one proposal (X0117), MSSIM gain is twice as large compared to PSNR (only proposal in SR where curves are not crossing). This is however not based on NN technology, just using RPR technology for upsampling, and for each rate point making a decision by multi-pass coding which coded resolution would be better. Similar approach might be used in NN based approaches.
* Memory requirements as reported by proponents might require some additional verification -> to be discussed in BoG
* First results of an intra prediction approach, including 16 bit integer implementation which reduces complexity.
* Impact of training was not thoroughly investigated, most proponents did not report about learning curves, etc. -> BoG to discuss how this could be improved, to get more information how this impacts the performance
* It is suggested to use log scale for the plots of kMAC/pxl
* It would be desirable to easify the study of overlap/orthogonality of tools investigated in EE1 with ECM, not mandatory but possibly only for best performing tools (e.g. study this as an AHG mandate). This could be realistic as the codebase of most proposals (VTM11) is similar, and even though decoder runtime is very high, it seems realistic due to the relative moderate increase of encoder runtime of NN based proposals.
* Present/discuss possible advantages of potentially using the library proposed in JVET-W0181 in BoG (was recommended for further study in last meeting, but now it has been used in some of the EE parts).

[JVET-X0052](https://jvet-experts.org/doc_end_user/current_document.php?id=11045) EE1-1.3: neural network based in-loop filter [L. Wang, W. Jiang, X. Xu, S. Liu (Tencent)]

This contribution reports the EE results of JVET-W0113[ ], the neural network based in-loop filter is exploited to replace the Deblocking and SAO modules. Compared with the EE1 anchor, the proposed method can reportedly achieve {4.20%, 11.08, 10.16%}, {2.96%, 9.73%, 9.23%} and {5.34%, 10.03%, 10.04%} BD-rate savings with RA, LB and AI configurations, respectively.

The reconstructed image after LMCS is fed into the network. Then, the output of NN filter is processed by ALF and CCALF. Deblocking and SAO are disabled. In detail, as shown in Fig. 1, U and V channels are up-sampled and down-sampled in the pre-processing and post-processing, respectively.

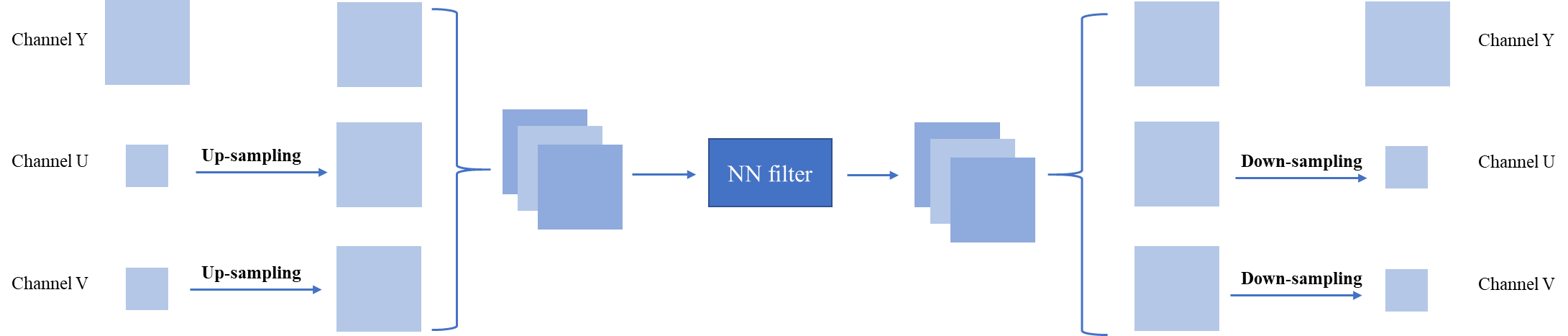


Fig. 1 the pre-processing and post-processing units

Fig[.](file:///C:\Users\ohm\AppData\Local\Temp\Temp1_JVET-X0052-v2.zip\JVET-X0052.docx#_Network_information_in) 2 shows the network structure of the proposed NN filter. The first and last layers are regular convolutional layers with kernel size 3. The skip layer is a regular convolutional layer with kernel size 5. The structure of residual block (ResBlock) consists of 3 regular convolutional layers, where N denotes the number of ResBlock. N is set to 8 in this contribution. In the ResBlock, the first layer is a 1x1 convolutional layer followed by a ReLU activation function, the second layer is a 1x1 convolutional layer and the third layer is 3x3 convolutional layer. As for the internal convolutional layers, the number of feature maps is set as 64.

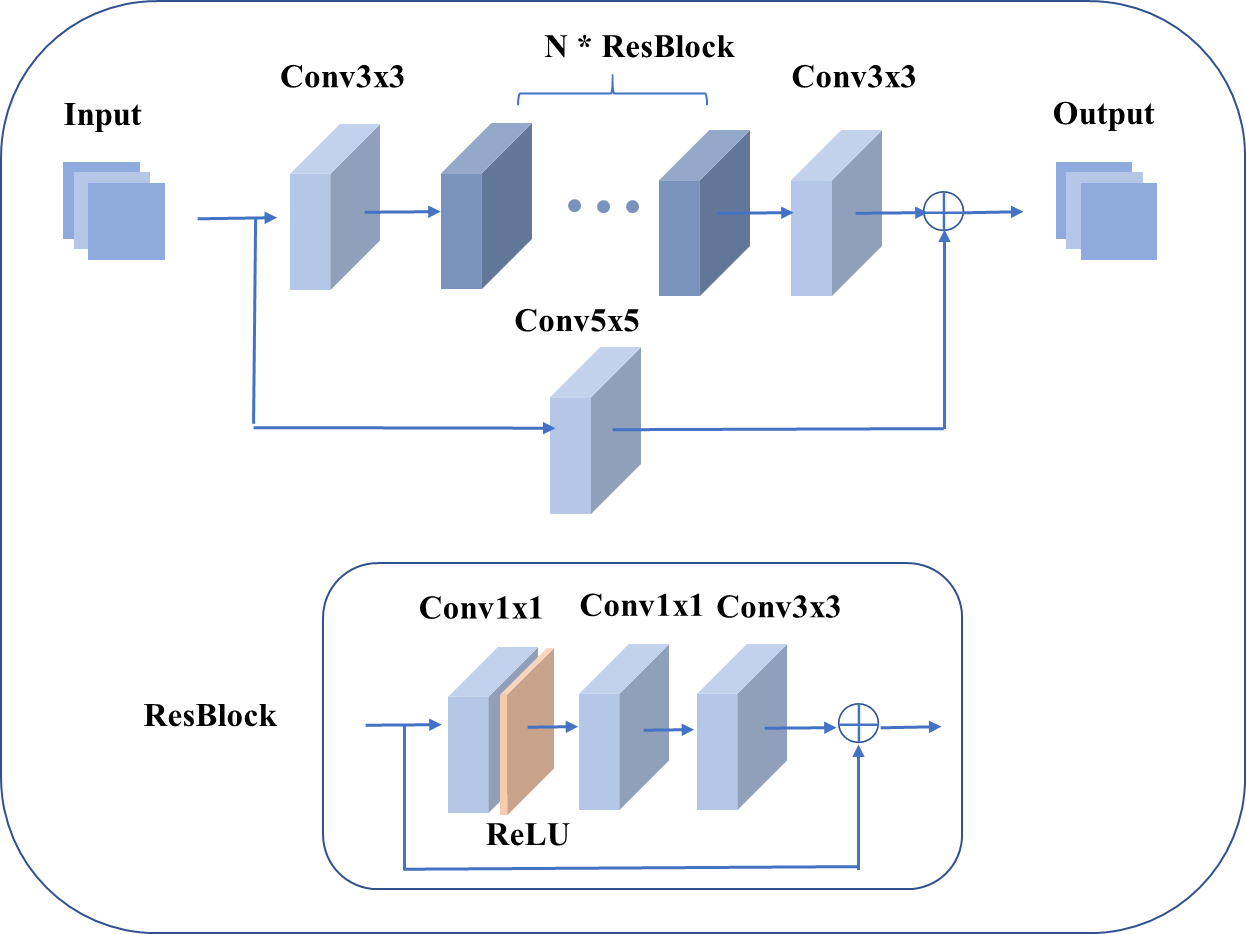


Fig. 2: Network architecture

Q: Is the 5x5 convolution needed? Some other proposals of the same company don’t use it. Actually, the filter can be disabled at CTU or slice level, so it should not do any harm at least and only be used if beneficial.

It is also pointed out that it should be straightforward to port the method to the library provided by Interdigital as that supports PyTorch.

[JVET-X0053](https://jvet-experts.org/doc_end_user/current_document.php?id=11046) EE1-1.5: neural network based in-loop filter using depthwise separable convolution and regular convolution [L. Wang, S. Lin, X. Xu, S. Liu, Xiang Li (Tencent)]

This contribution reports the EE results of JVET-W0151, the neural-network-based in loop filter using depthwise separable convolution (DSC) in EE1-1.5. With VTM11+ V0056[2][3] as anchor and test platform, it is reported that luma BD-Rate of 1.14%, 0.99%, and 1.62% can be achieved from RA, LB and AI configurations, respectively.

Significantly less complexity than X0052, but Deblocking and SAO are still used (NN filter between the two).

It is pointed out that it would be desirable to investigate for which frames the filter is actually used (as the decoding time had been reduced significantly by allowing to disable it).

[JVET-X0064](https://jvet-experts.org/doc_end_user/current_document.php?id=11057) EE1-2.2: CNN-based Super Resolution for Video Coding Using Decoded Information [C. Lin, Y. Li, K. Zhang, L. Zhang (Bytedance)]

This contribution reports the EE1-2.2 test results of JVET-W0099. JVET-W0099 presents a convolutional neural network-based super resolution scheme for video coding. Compared with VTM-11.0 with new MCTF enabled, the proposed method reportedly shows on average {-6.15%, 5.58%, -2.31%} and {-10.04%, 20.44%, -6.16%} BD-rate changes for {Y, U, V} under RA and AI configurations, respectively.

Luma: Two different models for I-slice and B-slice (I-slice model larger); chroma only one model.

Is the better performance in intra due to the larger model with 96 channels? The proponent mentions that the performance does not drop much when using 64 channels.

Why only L1 as training criterion? No specific reason (except that it is commonly used in SR), L2 fine-tuning might improve results

Processing is currently done in 128x128 blocks. Would line-wise processing be possible?

[JVET-X0065](https://jvet-experts.org/doc_end_user/current_document.php?id=11058) EE1-1.2: Test on Deep In-Loop Filter with Adaptive Model Selection and External Attention [Y. Li, K. Zhang, L. Zhang (Bytedance)]

No need for presentation, largely identical with JVET-W0100 which was presented in previous meeting.

[JVET-X0066](https://jvet-experts.org/doc_end_user/current_document.php?id=11059) EE1-1.6: Combined Test of EE1-1.2 and EE1-1.4 [Y. Li, K. Zhang, L. Zhang (Bytedance), H. Wang, J. Chen, K. Reuzé, A.M. Kotra, M. Karczewicz (Qualcomm)]

This contribution presents the experimental results of the EE1-1.6, i.e., the combined test of EE1-1.2 and EE1-1.4. In this combined test, convolutional neural network-based models are utilized for in-loop filtering. Compared with VTM-11.0 + NewMCTF, the test results reportedly show on average {9.80%, 21.20%, 21.17%}, {8.52%, 14.33%, 14.24%}, and {7.39%, 19.28%, 19.37%} BD-rate reductions for {Y, Cb, Cr} under RA, LDB, and AI configurations, respectively.

In the JVET-W meeting, several deep learning-based in-loop filtering methods were proposed to further improve the coding efficiency of VVC. Those methods are then further studied in EE1. Specifically, EE1-1.2 investigates a deep in-loop filter with adaptive model selection and external attention, while EE1-1.4 investigates a neural network-based in-loop filter with residual scaling. This contribution presents the experimental results of EE1-1.6, i.e. the combined test of EE1-1.2 and EE1-1.4.

Deblocking is disabled.

4 models are used (compared to 24 in X0065). It is highly beneficial that the previous block-level model selection was removed.

Training is using L1 until convergence, and then switches to L2.

No preprocessing of training data (compressed using VTM)

It is pointed out that the proposal was not cross-checked yet.

The proponents suggested to select their proposal as a “base software” for the NN exploration. Other experts expressed that at this moment the status of the exploration does not yet justify an official software package development. It would be more desirable to investigate the interrelationship with ECM (see discussion on this elsewhere).

It is also pointed out that there was an agreement at the previous meeting (and confirmed in the current meeting, see notes under BoG report), that for a proposal considered for “adoption” also the training process should be cross-checked.

The proposal could also be kept in the EE on loop filters without further modification, and get a cross-check done (for inference at least, and possibly for training).

[JVET-X0074](https://jvet-experts.org/doc_end_user/current_document.php?id=11067) EE1-2.4: 1.5x/2.0x Upsample method for NN-Based Super-Resolution Post-Filters [K. Takada, Y. Yasugi, T. Chujoh, T. Ikai (Sharp)]

This contribution reports an EE test result of JVET-W0132, a super-resolution post-filter using linear interpolation in feature domain to support arbitrary scaling factor. In this experiments, the performance of NN based 1.5 x and 2.0x scaling are studied compared to RPR anchor. It is reported that

* In 2.0x, the bd-rate gain is 1.64% in class A1 and 6.03% in class A2 against RPR anchor (2.0x)
* In 1.5x, the bd-rate gain is 1.71% in class A1 and 3.39% in class A2 against RPR anchor (1.5x)

The number of parameters of the used model was 336k.

Model is simplified ESRGAN, linear interpolation (bicubic) in feature domain (not pixel domain)

Presentation deck should be uploaded.

[JVET-X0107](https://jvet-experts.org/doc_end_user/current_document.php?id=11100) EE1-2.3: Neural Network-based Super Resolution [A. M. Kotra, K. Reuzé, J. Chen, H. Wang, M. Karczewicz, J. Li (Qualcomm)]

No need for presentation, largely identical with JVET-W0105 which was presented in previous meeting.

[JVET-X0184](https://jvet-experts.org/doc_end_user/current_document.php?id=11194) Cross-check of JVET-X0107 (EE1-2.3: Neural Network-based Super Resolution) [[K. Takada (Sharp)](mailto:keiichiro.takada@sharp.co.jp)] [late] [miss]

[JVET-X0110](https://jvet-experts.org/doc_end_user/current_document.php?id=11103) EE1-1.1: Content-adaptive neural network post-processing filter [M. Santamaria, J. Lainema, F. Cricri, R. G. Youvalari, H. Zhang, A. Zare, H. R. Tavakoli, M. Hannuksela (Nokia)]

This Exploration Experiment (EE) contribution studies the content-adaptive Convolutional Neural Network (CNN) post-processing filter presented in JVET-W0057. Two experiments were carried out. In the first one (EE1-1.1.1), content-adaptive CNNs were used. In the second one (EE1-1.1.2), pre-trained CNNs were used. A scaling factor determines the strength of the filter to be applied to the output. The scaling factor is selected by the encoder and signalled as a slice-level parameter within the bitstream. In addition, there is Coding Tree Unit (CTU) level control that allows the encoder to switch the filter on and off separately for luma and chroma. The evaluation was done on top of the VVC Test Model (VTM) 11.0 with Neural Network for Video Coding (NNVC) technology 1.0, using the Random Access (RA) configuration. It is reported that the proposed content-adaptation technique increases the overall BD-rate gains (over all classes) from -0.36% (Y), -0.13% (U), -0.82% (V) to -2.51% (Y), -5.86% (U), -3.86% (V).

Presentation deck should be uploaded.

Fine-tuning is done per sequence/QP. This is not included in the encoding run times, and it is also a two-pass encoding.

Number of parameters to be transmitted for fine-tuned parameters (once per sequence) is around 1%.

[JVET-X0117](https://jvet-experts.org/doc_end_user/current_document.php?id=11110) EE1-2.1: Super Resolution with existing VVC functionality [E. Alshina, J. Sauer (Huawei)]

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 gain (in PSNR) 1.0 % in random access and 0.5% in all intra configuration can be achieved. Gain is higher in MS-SSIM metric (2.0% in random access and 1.5% in all intra configuration).

Encoding time is approx. 125% as always both resolutions are encoded to decide which one is better.Almost only used for class A1, and only at lowest rate point.

[JVET-X0118](https://jvet-experts.org/doc_end_user/current_document.php?id=11111) EE1-3.1: BD-rate gains vs complexity of NN-based intra prediction [T. Dumas, F. Galpin, P. Bordes, F. Le Léannec (InterDigital)]

This contribution reports the EE1 results for the contribution JVET-W0081 presented during the 23rd JVET meeting.

Mean BD-rate reductions of -2.87% -2.32% -2.37% and -1.38% -0.55% -0.82% in AI and RA configurations respectively are reported when comparing VTM-11.0-NNVC including the low complexity version of the neural network-based intra prediction mode with respect to VTM-11.0-NNVC. The complexity (in MAC/pixel) of the low complexity version of the neural network-based intra prediction mode is about 12 times smaller than its regular version.

Presentation deck should be uploaded.Numbers above are for “test 4” (lowest complexity version, integer 16 bit).

For test 1 (highest complexity) version luma gains are 3.3%, 1.6% for AI/RA.

Cross-checker reports perfect matching for integer version.

Candidate for testing on top of ECM.

[JVET-X0123](https://jvet-experts.org/doc_end_user/current_document.php?id=11116) Crosscheck of JVET-X0118 (EE1-3.1: BD-rate gains vs complexity of NN-based intra prediction) [J. Sauer (Huawei)]

[JVET-X0140](https://jvet-experts.org/doc_end_user/current_document.php?id=11133) EE1-1.4: Tests on Neural Network-based In-Loop Filter with constrained computational complexity [H. Wang, J. Chen, K. Reuzé, A. M. Kotra, M. Karczewicz (Qualcomm)]

No need for presentation, largely identical with JVET-W0130/-W0131 which were presented in previous meeting.

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

Contributions in this area were discussed in session 11 at 1315–1500 UTC on Friday 8 Oct. 2021 (chaired by JRO), and session 14 at 1300–1400 UTC on Monday 11 Oct. 2021 (chaired by JRO).

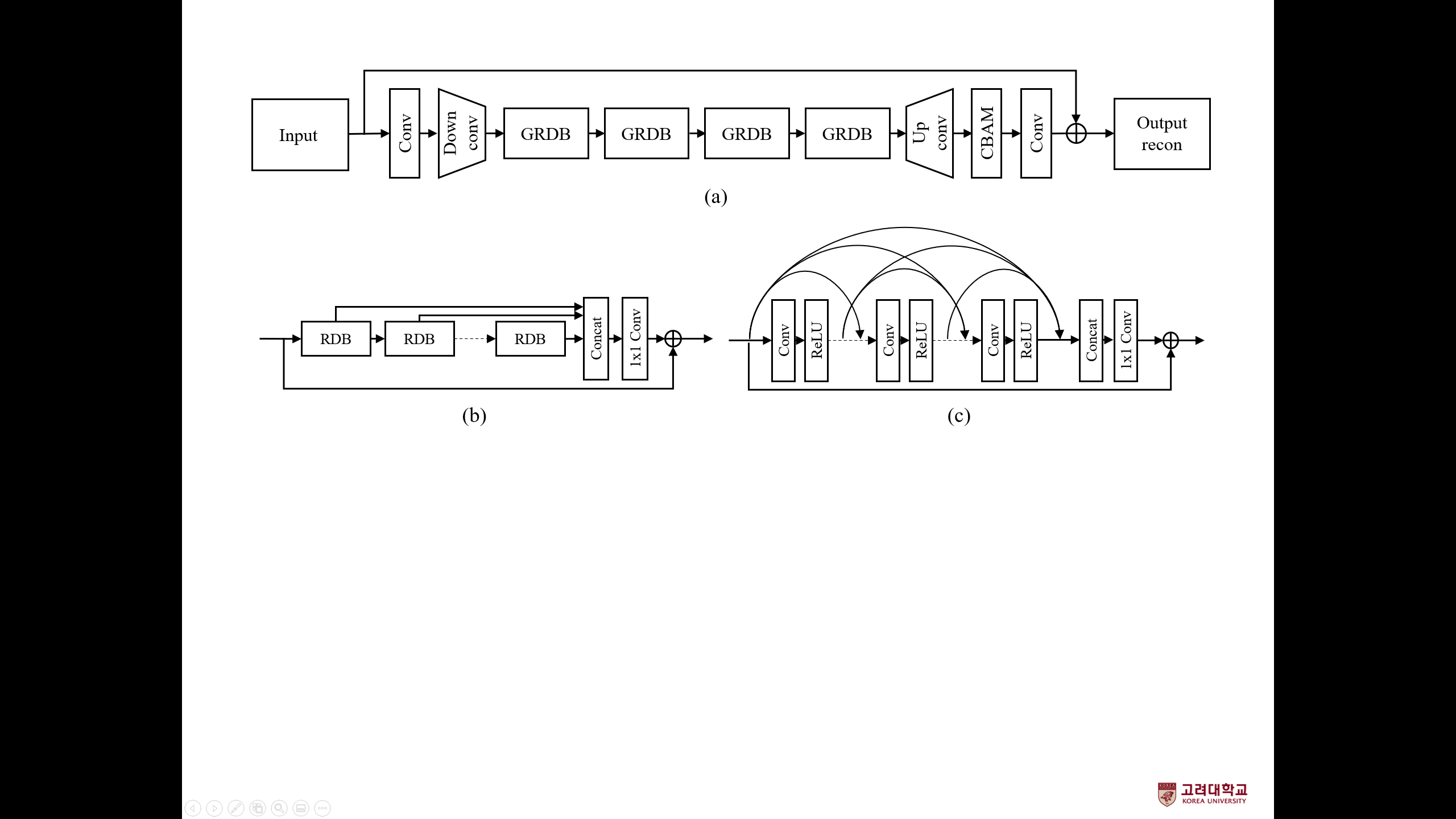
[JVET-X0041](https://jvet-experts.org/doc_end_user/current_document.php?id=11034) AHG11: CNN-based In-Loop Filter with Knowledge Distillation [H.-K. Kang, D.-W. Kim, S.-W. Jung (Korea Univ.), H. Kwon, S. Jeong (ETRI)]

This contribution presents a convolutional neural network (CNN)-based in-loop filtering method with knowledge distillation (KD). The proposed KD-based training strategy for in-loop filtering CNN enables a small-sized student network to perform favorably to a large-size teacher network. Compared with VTM-11.0-NNVC, the proposed method shows on average {7.43%, 9.71%, and 12.76%}, {5.91%, 9.52%, and 9.97%}, and {4.93%, -0.06%, and -0.50%} BD-rate reductions for {Y, Cb, Cr} on C and D classes, under AI, RA, and LDB configurations, respectively.

The architecture of the proposed CNN filter is shown in Fig.1. The backbone network is based on grouped residual dense network (GRDN), consisting of 4 grouped residual dense blocks (GRDBs) containing *N* residual dense blocks (RDBs). Each RDB has *M* cascaded convolutional layers followed by the ReLU activation layer. Convolutional block attention module (CBAM) is adopted for emphasizing spatial and channel-wise feature refinement.

The input size of the proposed CNN filter is equal to the size of the reconstructed frame before in-loop filtering process. The CNN filter restores the reconstructed input frame for each component (Y, U, V).

In KD, a large network, called teacher, transfers its knowledge to a small network, called student. Therefore, the teacher and the student with different block sizes are used during training. In this contribution, the teacher has 4 RDBs (*N*=4) in each GRDB and 8 convolution layers (*M*=8) in each RDB. On the other hand, the student has 2 RDBs (*N*=2) and 4 convolution layers (*M*=4) in each RDB. Both teacher and student have convolution layers with 64 channels. With this network design, the number of parameters of student is about seven times less than the one of the teacher’s parameters.

 *Fig. 1*. *(a)* *Architecture of GRDN. GRDN consists of 4 GRDBs. Conv denotes 3*ⅹ*3 convolution layer. (b) Architecture of GRDB. GRDB contains N RDBs. 1*ⅹ*1 Conv denotes 1*ⅹ*1 convolution. (c) Architecture of RDB. Each Conv has 64 channels.*

Note: Positive BD rate numbers in the abstract mean reduction

Loss of “student” compared to “teacher” is around 1% for AI, 0.5% for RA; 7x reduction of kMAC/pix

Two rounds of training were used for the student.

Would be interesting to see results for classes A and B, where gain is usually higher when high-resolution video is used for training (which was the case here).

No interest in joining the EE at this time.

[JVET-X0054](https://jvet-experts.org/doc_end_user/current_document.php?id=11047) AHG11: neural network based in-loop filter with adaptive model selection [L. Wang, X. Xu, S. Liu (Tencent)]

In this contribution, a neural network based in-loop filter is proposed. According to the slice type, the QP value and the color component, multiple models are trained for the proposed filter. Especially for B slices, adaptive model selection is utilized. In addition to the reconstruction image, other side information is fed into the network, such as prediction image, partition image. Compared with EE1 anchor, the proposed method can reportedly achieve {8.19%, 24.57, 24.32%}, {6.86%, 25.46%, 26.51%} and {7.05%, 19.77%, 19.56%} BD-rate savings with RA, LB and AI configurations, respectively.22 models in total

Q: Gain of attention model which is new? Not known.

The proposal is built on top of W0113, but combines some elements from W0100. Difference of model selection compared to W0100? Similar, also uses block selection at block level, with additional input parameters such as partitioning. Different from W0100, chroma is used as input to the luma filter. In total more complex than W0113, also more memory.

Is it known how much gain the additional elements give? Partitioning gives approx. 0.7%

[JVET-X0055](https://jvet-experts.org/doc_end_user/current_document.php?id=11048) AHG11: neural network based in-loop filter with constrained storage and low complexity [L. Wang, X. Xu, S. Liu (Tencent)]

A neural network based in-loop filter with constrained storage and low complexity is proposed in this contribution. In all, only two models are used in the proposed filter design, for which each model is designed with relatively low complexity. In addition to the reconstruction image, other side information is fed into the network, such as the prediction image, partition image, slice QP and based QP. Based on the EE1 anchor, two NN filter tests are proposed with different complexity levels. For the filter test 1, { 7.15%, 19.33%, 19.11%}, {6.07%, 18.25%, 18.98%} and { 7.32%, 18.49%, 18.33%} BD-rate savings are reported with RA, LB and AI configurations, respectively. For the filter test 2, { 4.04%, 7.79%, 5.83%}, { 3.31%, 7.33%, 5.27%} and {4.82%, 8.48%, 6.84%} BD-rate savings are reported with RA, LB and RA configurations, respectively.

For B slices, the filter is operated in parallel with DBF/SAO, and the better is selected (slice level switching).

For I slices, it replaces DBF/SAO.

2 models per filter. Less complex than X0055

As elements proposed in X0054 and X0055, usage of partitioning information could be interesting for EE investigation.

[JVET-X0080](https://jvet-experts.org/doc_end_user/current_document.php?id=11073) EE1-related: CNN-based Super Resolution for Video Coding Using Decoded Information with Simplified Models [C. Lin, Y. Li, K. Zhang, L. Zhang (Bytedance)]

This EE related contribution studies the performance of the network tested in EE1-2.2 with simplified models. In this simplified version, the number of feature maps is reduced, resulting in roughly 43% lower kMAC/pixel comparing to the models tested in EE1-2.2. The simulation results reportedly show {-6.15%, 5.58%, -2.31%} and {-9.09%, 20.44%, 6.16%} BD rate savings for {Y, Cb, Cr} under RA and AI configurations, respectively.

In RA, almost no change, in AI approx. 1% performance drop relative to EE1-2.2.

Results are only for classes A1/A2. In particular for chroma, results are quite divergent over different sequences. Averaging BD rates might not be meaningful there.

One participant suggests that further reduction in complexity might be achievable.

Why loss in AI? Uses a different model than for B slice.

Investigate in EE.

[JVET-X0081](https://jvet-experts.org/doc_end_user/current_document.php?id=11074) EE1-related: CNN-based Super Resolution for Video Coding Using Separate Networks for Chroma Components [C. Lin, Y. Li, K. Zhang, L. Zhang (Bytedance)]

This EE related contribution is asserted to improve the performance of EE1-2.2 by using separate networks for Cb and Cr components. For anchor QPs in {22, 27, 32, 37, 42}, the simulation results show {-6.15%, 2.53%, -6.33%} and {-10.04%, 14.90%, -9.37%} BD-rate savings for {Y, Cb, Cr} under RA and AI configurations, respectively. For anchor QPs in {27, 32, 37, 42, 47}, the simulation results show {-9.73%, -2.32%, -7.44%} and {-11.99%, 7.79%, -9.61%} BD-rate savings for {Y, Cb, Cr} under RA and AI configurations, respectively.

Luma model as in EE1-2-2.

Still both chroma components (and luma) are input to the separate networks. One expert mentions that this might not be necessary.

It is pointed out that the BD numbers might be affected by curves crossing. It might be useful (as the rate should be exactly the same, it is postprocessing), if by tendency the PSNR is improved for chroma, using X0064 as anchor (in that case the curves are no longer crossing).

Using RPR as anchor might also provide better interpretation of results (for all SR proposals).

Also RPR has sometimes large losses in chroma. It is asked whether for chroma it might be better just not to downsample chroma. However, the current RPR configuration does not allow downsampling to 4:4:4.

Investigate in EE.

[JVET-X0082](https://jvet-experts.org/doc_end_user/current_document.php?id=11075) EE1-related: Training Using Knowledge Distillation for Deep In-Loop Filters [Y. Li, K. Zhang, L. Zhang (Bytedance)]

This contribution presents a training optimization technique on top of the deep learning-based in-loop filtering method presented in JVET-W0100. The proposed method is asserted to improve the performance by introducing knowledge distillation to the training process. Compared with VTM-11.0+NewMCTF, the test results reportedly show on average {9.25%, 23.31%, 23.50%} BD-rate reductions for {Y, Cb, Cr} under AI configurations.

Unlike X0041, the purpose of knowledge distillation here is improving performance rather than reducing complexity (“student” becomes better than “teacher” model from W0100).

Student is trained from scratch, but part of the loss function is for generating the same output as the teacher. This may lead to faster convergence.

Might it be advantageous to adjust the weight (influence of teacher) during training?

Training becomes longer in total – would increase of “normal” training duration (of the teacher) also improve performance? Not known.

Proponents intend further study before joining EE.

[JVET-X0084](https://jvet-experts.org/doc_end_user/current_document.php?id=11077) EE1-related: Improved RDO Considering Deep In-Loop Filter [J. Li, Y. Li, K. Zhang, L. Zhang (Bytedance)]

This contribution presents an encoder optimization technique on top of the deep learning-based in-loop filtering method presented in JVET-W0100. The proposed method is asserted to improve the compression efficiency at encoder by introducing CNN-based filtering into the rate-distortion optimization (RDO) process. Compared with VTM11.0\_nnvc, the proposed method reportedly shows on average {9.61%, 22.52%, 22.71%}, {12.81%, 27.64%, 27.42%}, and {x%, x%, x%} BD-rate reductions for {Y, Cb, Cr} components, under AI, RA, and LDB configurations, respectively.

A simplified version of NN filter is used in RDO. What is the size? What is the drop in performance? Cannot answer precisely

Encoding time increases by 10+ %, additional bit rate reduction is around 0.5%, compared to W0100.

It is noted that this is similar gain as with RDO considering conventional loop filters such as DBF.

In the current exploration, optimization of the loop filters themselves seems more important than additional tricks of RDO decisions.

[JVET-X0094](https://jvet-experts.org/doc_end_user/current_document.php?id=11087) AHG11: A Deep In-Loop Filter with Frame Level Flag [X. Zhang, C. Fang, S. Peng, D. Jiang, J. Lin (Dahua)]

This contribution presents a convolutional neural network-based in-loop filtering method with QP based models. The proposed method is developed from the previous contribution JVET-W0059. In this contribution, models with a longer training time and a frame-level flag for each component are utilized.

Compared with VTM-11.0-NNVC, the proposed method reportedly shows on average {5.80%, 16.54%, 18.19%} and {%, %, %} BD-rate reductions for {Y, Cb, Cr}, under AI and RA configuration, respectively.

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.



(a)



(b)



(c)

Results on RA not ready yet, but by tendency lower gain than AI.

Question for further study (EE?): What would be the standalone gain of luma filter?

[JVET-X0097](https://jvet-experts.org/doc_end_user/current_document.php?id=11090) AHG11: A CNN-based Super Resolution Method [S. Peng, C. Fang, X. Zhang, D. Jiang, J. Lin (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.

In the proposed method, the channel attention mechanism is added into the residual block.

Compared with VTM-11.0-NNVC, the experimental results of the proposed method are as follow:

Test1: AI 6.26% BD-rate gain for Class A1&A2.

Test2: AI 6.36% BD-rate gain for Class A1&A2.

Test 1: Five models for five different QP ranges

Test 2: Two models only for low and high QP ranges

Test 2 has slightly better performance.

Attention module is added to the residual block, however only with luma as input (whereas other approaches also take additional info such as prediction residual). This could be interesting to compare, and also investigate whether it gives subjective benefit.

Chroma is using RPR filters, no NN technology

Investigate in EE.

[JVET-X0111](https://jvet-experts.org/doc_end_user/current_document.php?id=11104) AHG11: MPEG NNR compressed bias update for the CNN based post-filter of EE1-1.1 [M. Santamaria, J. Lainema, F. Cricri, R. G. Youvalari, H. Zhang, A. Zare, G. Rangu, H. R. Tavakoli, H. Afrabandpey, M. Hannuksela (Nokia)]

This contribution presents a Convolutional Neural Network (CNN) based content-adaptive post-processing filter. The filter is content-adaptive in the sense that it is trained offline on general video sequences, and later fine-tuned on the test video sequence. The work is based on the filter described in JVET-X0110, where only the bias terms of the CNN are fine-tuned resulting in a weight-update that is compressed with Lempel-Ziv-Markov chain algorithm 2 (LZMA2). In this study the weight-update is switched to use the MPEG Neural Network compression and Representation (NNR) standard. The weight-update was compressed with NNR’s reference model, the Neural network Compression Test Model (NCTM). The proposed technique was implemented on top of the EE1 anchor software (VVC Test Model (VTM) 11.0 with Neural Network for Video Coding (NNVC) technology 1.0). It is reported that compressing the weight-update by using NCTM achieves overall BD-rate gains in Random Access (RA) configuration of -2.64% (Y), -5.94% (U), -3.97% (V). Furthermore, when compared against the weight-update compression with LZMA2 in RA, the average gains are -0.13% (Y), -0.08% (U), -0.10% (V).

On top of EE contribution X0110, uses NNR for compressing parameters (lossy coding).

The number of bits for parameters is reduced by approximately 60%.

[JVET-X0113](https://jvet-experts.org/doc_end_user/current_document.php?id=11106) AHG11: CNN-based Low Complexity Super Resolution [E. Yeo, J. Kang (Ewha W. University), D. Kim, K. Kim, J.-H. Son, J. S. Kwak (WILUS Inc.)]

This contribution proposes convolutional neural network-based super resolution using relatively low-complexity network structures. RPR (Reference Picture Resampling) is applied to down-sample the input video, then the proposed CNN-based super resolution post-filter is applied to up-sample the decoded video. Neural network used in this proposal is simple CNN-based network consisted of 8 layers, so the computational complexity is relatively low. For Class A1 sequences, average Luma BD-Rate gains of -5.82% and -2.42% with patch size of 64 and -7.38% and -4.82% with patch size of 128 for RA and AI configurations respectively were observed.Less complexity than other SR approaches investigated so far. Roughly 12K parameters, 20 kMac/pixel. No report on decoder runtime.

Configuration with patch size 128 performs better than 64.

Class A1 performs more favorably for than A2.

Chroma shows losses in all sequences PSNR-wise.

MSSIM generally better than PSNR based metrics.

Subjective quality?

At the current time, proponents are not ready yet to join EE.

### Other (6)

Contributions in this area were discussed in session 14 at 1410–1500 UTC on Monday 11 Oct. 2021 (chaired by JRO), and session 18 at 1415-1500 on Tuesday 12 Oct. 2021 (chaired by JRO).

[JVET-X0043](https://jvet-experts.org/doc_end_user/current_document.php?id=11036) [AHG11 & AHG6] DOVC: Deep Omnidirectional Video Compression [Q. Qin, C. Jung (Xidian Univ.), Z. Dan, M. Li (OPPO)]

This contribution presents an end-to-end deep omnidirectional video compression (DOVC) framework with convolutional neural networks (CNNs). The proposed DOVC framework includes a bidirectional offset prediction module based on deformable convolution to generate effective motion information and a bidirectional motion compensation (MC) network composed of ordinary convolution and residual blocks for motion compensated prediction. The proposed DOVC framework also reuses the motion information generated by the bidirectional offset prediction module to enhance residual in the quality enhancement (QE) module. Regarding omnidirectional video, the modules in the proposed DOVC framework are jointly optimized by weighted rate distortion to minimize the geometric distortion of the sphere-to-plane projection. Experimental results show that DOVC achieves average 21% reduction in BD-BR and average 0.6217dB gain in BD-WS\_PSNR over HM-16.16 (360Lib-5.0) under LDP configuration for encoding omnidirectional videos, and the average encoding time of DOVC is only 0.0234 times that of HM-16.16 (360Lib-5.0) and 0.0064 times that of VTM-11.0 (360Lib-12.0).

ERP and CMP as projection formats for pre/post processing

Roughly 20M parameters

I frames (distance 8) coded by BPG. Why not using end-to-end intra compression (or VVC intra) which should have better performance? Proponents did not think about. How is the bit rate for intra frame decided?

Performance worse than VVC (measured by WS-PSNR): 12% bit rate increase in CMP, 38% in ERP

Coding currently in RGB 4:4:4, while VVC and HEVC in YUV 4:2:0 (proponents expect better performance when also coding in YUV)

Encoding faster than VVC, decoding slower (computed on GPU)

Could be applied to common video (by adjusting loss function)

Further study welcome.

[JVET-X0060](https://jvet-experts.org/doc_end_user/current_document.php?id=11053) 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 1.47%/4.19%/3.85% bitrate savings on Y/U/V component respectively, following the JVET-NNVC CTC RA configuration.

In this v3 revision, the results of test1 are updated.

In this v4 revision, the results of test2 (without additional filter network) are updated. Slightly worse results.

Roughly 45M parameters, 1.7M MAC/pixel.

Encoding time more than doubled, likely due to testing the additional reference picture (and its gerenation). Which reference frame in RPL is replaced, or is it an additional reference frame? Only applied in highest three, so number is not increased.

Further study and contributions welcome.

[JVET-X0102](https://jvet-experts.org/doc_end_user/current_document.php?id=11095) AHG11: Deep neural network for inter bi-prediction [Y.-J. Choi, Y.-W. Lee, B.-G. Kim (SWU), J. Kim, S. Y. Jeong (ETRI)]

This contribution proposes a deep neural network to generate a bi-prediction block by blending two motion compensated blocks. The proposed method replaces the regular bi-prediction methods. Compared with VTM-11.0, the proposed method reportedly shows on average 0.98%, 0.96%, and 2.08% BD-rate reductions for Y component on B, C, and D classes under RA configurations, respectively.

Attention model to generate a residual that is added on top of the usual averaging in bi-prediction. Other bi-prediction modes (BCW, BDOF) are replaced.

Encoding time 42x, decoding time 7x increased, though run on GPU.

Further study and contributions welcome.

[JVET-X0125](https://jvet-experts.org/doc_end_user/current_document.php?id=11118) AHG11: Autoencoder-based intra prediction with auxiliary feature [L. Xu, Y. Yu, H. Yu, K. Sato, Z. Dai, Z. Xie, D. Wang (OPPO)]

This document describes an autoencoder-based intra prediction mode that consists of a pair of jointly trained encoder network and decoder network. For a given block, the encoder network extracts feature information from the original samples and sends the feature to the video encoder, and the decoder network inside a video decoder generates a prediction block by using the decoded feature and neighboring reference samples. This autoencoder is integrated into VTM-11.0 as an additional intra prediction mode (also called “autoencoder mode” in this contribution). For the all-intra configuration, the test results showed -1.11%, -0.89 %, and -0.91 % of Y BD-rate gain, U BD-rate gain, and V BD-rate gains, respectively. With this configuration, the relative encoding and decoding runtimes of VTM-11.0 with the proposed neural network-based intra prediction mode vs. VTM-11.0 are 338% and 559%, respectively. For the random-access configuration, the test results showed, in average, -0.53 %, -0.49 %, and -0.50 % of Y BD-rate gain, U BD-rate gain, and V BD-rate gain, respectively. With this configuration, the relative encoding and decoding runtimes of VTM-11.0 with the proposed neural network-based intra prediction mode vs. VTM-11.0 are 138% and 253%, respectively.

3 layers, fully connected

Slightly better performance at lower end of QP range.

Auxiliary feature is extracted and transmitted as side information (2 bins)

9 models for the different block sizes, total number of parameters is approx. 5 M at decoder

Proponents would like to perform further study before joining a possible EE.

[JVET-X0126](https://jvet-experts.org/doc_end_user/current_document.php?id=11119) AHG11: Neural Network-based Adaptive Model Selection for CNN In-Loop Filtering [Z. Dai, Y. Yu, H. Yu, K. Sato, L. Xu, Z. Xie, D. Wang (OPPO)]

This contribution presents a convolutional neural network-based in-loop filter (CNNLF) wherein a neural network-based, adaptive model selection (NNAMS) method is introduced. The proposed CNNLF+NNAMS is integrated in the code base of VTM-11.0 + new MCTF. Compared with VTM-11.0 + new MCTF, the proposed method demonstrates {-3.20%, -4.65%, -5.96%} and {-1.77%, -5.25%, -5.94%} BD-rate savings with the AI and RA configurations, respectively.

10 layers of residual blocks, 64 channels each

5 different models per color component, NNAMS predicts which one to use.

Total number of parameters: 739+41+39K (for CNNLF+LumaNNAMS+ChromaNNAMS)

In RA, NNAMS provides approx. 0.4% of the bit rate gain

Model switched at CTU level, NNAMS saves approx. 8 bits/CTU

It is pointed out that invoking NNAMS at CTU level may not be competitive with other sophisticated methods of model selection (an signaling) that have been shown.

No specific action at this point, further study encouraged.

[JVET-X0130](https://jvet-experts.org/doc_end_user/current_document.php?id=11123) AHG11: Cross-component prediction based on a neural network model [Y. Y. Lee, T. M. Bae, D. Ruiz Coll, K. Goswami, A. Filippov, V. Rufitskiy (Ofinno)] [late]

This contribution presents cross-component prediction based on a neural network model (CCNN). As an alternative mode of intra chroma prediction, the CCNN may further reduce the cross-component redundancy. Based on the AHG11 anchor (VTM11+V0056), the proposed CCNN method reportedly provides the compression performance improvements of -0.86%/0.82%/1.04% and -0.32%/0.35%/0.62% (Y/Cb/Cr) in AI and RA configurations, respectively. The additional test results reportedly present the compression performance improvements of ‑0.37%/‑2.93%/ ‑2.70% (Y/Cb/Cr) for AI configuration after redistributing the compression performance improvements between luma and chroma using lambda adjustment for the chroma components.

2.65M parameters, and same number of MAC/pixel

18 convolutional layers

Additional results are based on lambda adjustment, to shift bit rate from luma to chroma.

RDO tends to choose the model which then generates fewer residual bits with less quality, this is why the lambda adjustment was done as an add-on.

Relative high complexity versus low gain. Could not a similar effect be achieved with a much simpler structure? Would the gain still be retained e.g. in combination with the intra prediction from the last EE, or cross-color aspects of loop filters and super-resolution.

No specific action at this point, further study encouraged in particular for complexity reduction and performance improvement.

### NN related HLS signalling (0)

Contributions in this area were discussed in session x at XXXX–XXXX UTC on XXday X Oct. 2021 (chaired by XXX).

## AHG12: Enhanced compression beyond VVC capability (52)

### General (0)

Contributions in this area were discussed in session x at XXXX–XXXX UTC on XXday X Oct. 2021 (chaired by XXX).

### EE2 contributions: Enhanced compression beyond VVC capability (13)

Contributions in this area were discussed in session 3 at 1425–1510 and session 4 at 1530–1710 UTC on Wednesday 6 Oct. 2021 (chaired by JRO), in session 8 at 1530-1630 UTC on Thursday 7 Oct. 2021 (chaired by JRO), and ….

[JVET-X0024](https://jvet-experts.org/doc_end_user/current_document.php?id=11183) EE2: Summary Report on Enhanced Compression beyond VVC capability [V. Seregin, J. Chen, S. Esenlik, F. Le Léannec, L. Li, J. Ström, M. Winken, X. Xiu, K. Zhang]

It is noted that tests 4.7..4.9 investigating combinations of loop filters was added late to the EE document after no objections were raised on the reflector. It is pointed out that in our common working practices this would rather be considered as an “EE related” contribution, and that there are many cases of evidence from the past that such proposals (if asserted to be reasonable) could be adopted right away.

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-2.0, released at [https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tags/ECM-2.0.](https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tags/ECM-2.0.E) ECM-2.0 is used as an anchor in the tests.

Software for EE tests is released in the corresponding branches at <https://vcgit.hhi.fraunhofer.de/ecm/jvet-w-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-w-ee2/simulation-results/-/tree/master>.

1. **List of tests**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Tests** | **Tester** | **Cross-checker** |
| **1 Partitioning** | | | |
| 1.1 | Encoder optimization for ECM | ByteDance  [Kai Zhang](mailto:zhangkai.video@bytedance.com)  InterDigital  [Fabrice Le Léannec](mailto:fabrice.leleannec@interdigital.com)  [JVET-X0068](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0068-v1.zip) | Qualcomm  [Muhammed Coban](mailto:mcoban@qti.qualcomm.com)  [JVET-X0144](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0144-v2.zip)  Alibaba  Jie Chen |
| 1.2 | ABT | ByteDance  [Kai Zhang](mailto:zhangkai.video@bytedance.com)  InterDigital  [Fabrice Le Léannec](mailto:fabrice.leleannec@interdigital.com)  [JVET-X0068](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0068-v1.zip) | Alibaba  Jie Chen |
| 1.3 | UQT | ByteDance  [Kai Zhang](mailto:zhangkai.video@bytedance.com)  InterDigital  [Fabrice Le Léannec](mailto:fabrice.leleannec@interdigital.com)  [JVET-X0068](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0068-v1.zip) | Alibaba  Jie Chen |
| 1.4 | ABT + UQT | ByteDance  [Kai Zhang](mailto:zhangkai.video@bytedance.com)  InterDigital  [Fabrice Le Léannec](mailto:fabrice.leleannec@interdigital.com)  [JVET-X0068](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0068-v1.zip) | Alibaba  Jie Chen |
| **3 Inter prediction** | | | |
| 3.1a | Combination of CIIP and DIMD | Alibaba  [Xinwei Li](mailto:sid.lxw@alibaba-inc.com)  [JVET-X0098](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0098-v1.zip) | ByteDance  [Yang Wang](mailto:wangyang.cs@bytedance.com)  [JVET-X0177](https://jvet-experts.org/doc_end_user/current_document.php?id=11186) |
| 3.1b | Combination of CIIP and TIMD | Alibaba  [Xinwei Li](mailto:sid.lxw@alibaba-inc.com)  [JVET-X0098](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0098-v1.zip) | ByteDance  [Yang Wang](mailto:wangyang.cs@bytedance.com)  [JVET-X0177](https://jvet-experts.org/doc_end_user/current_document.php?id=11186) |
| 3.2a | GPM with inter and intra prediction – method A | KDDI  [Yoshitaka Kidani](mailto:yo-kidani@kddi.com)  [JVET-X0077](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0077-v1.zip) | Tencent  Ling Li |
| 3.2b | GPM with inter and intra prediction – method B | KDDI  [Yoshitaka Kidani](mailto:yo-kidani@kddi.com)  [JVET-X0077](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0077-v1.zip) | Tencent  Ling Li |
| 3.3a | Bilateral matching AMVP-merge mode | Qualcomm  [Zhi Zhang](mailto:zhizhang@qti.qualcomm.com)  [JVET-X0083](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0083-v1.zip) | Kwai  [Wei Chen](mailto:chenwei06@kwai.com) |
| 3.3b | Bilateral matching AMVP-merge mode with TM disabled | Qualcomm  [Zhi Zhang](mailto:zhizhang@qti.qualcomm.com)  [JVET-X0083](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0083-v1.zip) | Kwai  [Wei Chen](mailto:chenwei06@kwai.com) |
| 3.3c | Template matching AMVP-merge mode | Qualcomm  [Zhi Zhang](mailto:zhizhang@qti.qualcomm.com)  [JVET-X0083](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0083-v1.zip) | Kwai  [Wei Chen](mailto:chenwei06@kwai.com) |
| 3.4a | Adaptive decoder side motion vector refinement | Qualcomm  [Han Huang](mailto:hanhuang@qti.qualcomm.com)  [JVET-X0049](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0049-v1.zip) | Alibaba  [Ru-Ling Liao](mailto:ruling.lrl@alibaba-inc.com) |
| 3.4b | Adaptive decoder side motion vector refinement with TM disabled | Qualcomm  [Han Huang](mailto:hanhuang@qti.qualcomm.com)  [JVET-X0049](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0049-v1.zip) | Alibaba  [Ru-Ling Liao](mailto:ruling.lrl@alibaba-inc.com) |
| **4 In-loop filtering** | | | |
| 4.1 | Chroma bilateral filter | Bytedance  [Wenbin Yin](mailto:yinwenbin.hit@bytedance.com)  [JVET-X0067](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0067-v1.zip) | Alibaba  [Mohammed Golam Sarwer](mailto:m.sarwer@alibaba-inc.com) |
| 4.2a | CTB level filter shape selection of CCALF | Alibaba  [Mohammed Golam Sarwer](mailto:m.sarwer@alibaba-inc.com)  [JVET-X0045](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0045-v1.zip) | Qualcomm  [Nan Hu](mailto:nanh@qti.qualcomm.com) |
| 4.2b | CTB level filter shape selection of CCALF with removal of power of 2 constraint of filter coefficients | Alibaba  [Mohammed Golam Sarwer](mailto:m.sarwer@alibaba-inc.com)  [JVET-X0045](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0045-v1.zip) | Qualcomm  [Nan Hu](mailto:nanh@qti.qualcomm.com) |
| 4.3a | CCALF with larger filter size | Alibaba  [Mohammed Golam Sarwer](mailto:m.sarwer@alibaba-inc.com)  [JVET-X0045](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0045-v1.zip) | Qualcomm  [Nan Hu](mailto:nanh@qti.qualcomm.com) |
| 4.3b | CCALF with larger filter size with removal of power of 2 constraint of filter coefficients | Alibaba  [Mohammed Golam Sarwer](mailto:m.sarwer@alibaba-inc.com)  [JVET-X0045](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0045-v1.zip) | Qualcomm  [Nan Hu](mailto:nanh@qti.qualcomm.com) |
| 4.4 | Alternative 2x2 classifier for ALF | Qualcomm  [Nan Hu](mailto:nanh@qti.qualcomm.com)  [JVET-X0070](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0070-v1.zip) | Alibaba  [Mohammed Golam Sarwer](mailto:m.sarwer@alibaba-inc.com) |
| 4.5 | Alternative sample-based classifier for ALF | Qualcomm  [Nan Hu](mailto:nanh@qti.qualcomm.com)  [JVET-X0070](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0070-v1.zip) | Alibaba  [Mohammed Golam Sarwer](mailto:m.sarwer@alibaba-inc.com) |
| 4.6a | Combination of 4.2a and 4.4 (CTB level filter shape selection of CCALF and alternative 2x2 classifier for ALF) | Alibaba  [Mohammed Golam Sarwer](mailto:m.sarwer@alibaba-inc.com)  Qualcomm  [Nan Hu](mailto:nanh@qti.qualcomm.com)  [JVET-X0046](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0046-v1.zip) | Bytedance  [Wenbin Yin](mailto:yinwenbin.hit@bytedance.com) |
| 4.6b | Combination of 4.2a and 4.5 (CTB level filter shape selection of CCALF and alternative sample-based classifier for ALF) | Alibaba  [Mohammed Golam Sarwer](mailto:m.sarwer@alibaba-inc.com)  Qualcomm  [Nan Hu](mailto:nanh@qti.qualcomm.com)  [JVET-X0046](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0046-v1.zip) | Bytedance  [Wenbin Yin](mailto:yinwenbin.hit@bytedance.com) |
| 4.6c | Combination of 4.3a and 4.4 (CCALF with larger filter size and alternative 2x2 classifier for ALF) | Alibaba  [Mohammed Golam Sarwer](mailto:m.sarwer@alibaba-inc.com)  Qualcomm  [Nan Hu](mailto:nanh@qti.qualcomm.com)  [JVET-X0046](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0046-v1.zip) | Bytedance  [Wenbin Yin](mailto:yinwenbin.hit@bytedance.com) |
| 4.6d | Combination of 4.3a and 4.5 (CCALF with larger filter size and alternative sample-based classifier for ALF) | Alibaba  [Mohammed Golam Sarwer](mailto:m.sarwer@alibaba-inc.com)  Qualcomm  [Nan Hu](mailto:nanh@qti.qualcomm.com)  [JVET-X0046](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0046-v1.zip) | Bytedance  [Wenbin Yin](mailto:yinwenbin.hit@bytedance.com) |
| 4.7a | Combination of 4.2a and 4.1 (CTB level filter shape selection of CCALF and Chroma Bilateral) | Alibaba  [Mohammed Golam Sarwer](mailto:m.sarwer@alibaba-inc.com)  Bytedance  [Wenbin Yin](mailto:yinwenbin.hit@bytedance.com)  [JVET-X0047](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0047-v1.zip) |  |
| 4.7b | Combination of 4.3a and 4.1 (CCALF with larger filter size and Chroma Bilateral) | Alibaba  [Mohammed Golam Sarwer](mailto:m.sarwer@alibaba-inc.com)  Bytedance  [Wenbin Yin](mailto:yinwenbin.hit@bytedance.com)  [JVET-X0047](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0047-v1.zip) |  |
| 4.8a | Combination of 4.6a and 4.1 (Chroma Bilateral, CTB level filter shape selection of CCALF and alternative 2x2 classifier for ALF) | Alibaba  [Mohammed Golam Sarwer](mailto:m.sarwer@alibaba-inc.com)  Bytedance  [Wenbin Yin](mailto:yinwenbin.hit@bytedance.com)  Qualcomm  [Nan Hu](mailto:nanh@qti.qualcomm.com)  [JVET-X0071](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0071-v1.zip) |  |
| 4.8b | Combination of 4.6b and 4.1 (Chroma Bilateral, CTB level filter shape selection of CCALF and alternative sample-based classifier for ALF) | Alibaba  [Mohammed Golam Sarwer](mailto:m.sarwer@alibaba-inc.com)  Bytedance  [Wenbin Yin](mailto:yinwenbin.hit@bytedance.com)  Qualcomm  [Nan Hu](mailto:nanh@qti.qualcomm.com)  [JVET-X0071](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0071-v1.zip) |  |
| 4.8c | Combination of 4.6c and 4.1 (Chroma Bilateral, CCALF with larger filter size and alternative 2x2 classifier for ALF) | Alibaba  [Mohammed Golam Sarwer](mailto:m.sarwer@alibaba-inc.com)  Bytedance  [Wenbin Yin](mailto:yinwenbin.hit@bytedance.com)  Qualcomm  [Nan Hu](mailto:nanh@qti.qualcomm.com)  [JVET-X0071](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0071-v1.zip) |  |
| 4.8d | Combination of 4.6d and 4.1 (Chroma Bilateral, CCALF with larger filter size and alternative sample-based classifier for ALF) | Alibaba  [Mohammed Golam Sarwer](mailto:m.sarwer@alibaba-inc.com)  Bytedance  [Wenbin Yin](mailto:yinwenbin.hit@bytedance.com)  Qualcomm  [Nan Hu](mailto:nanh@qti.qualcomm.com)  [JVET-X0071](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0071-v1.zip) |  |
| 4.9a | Combination of 4.1 and 4.4 (Chroma Bilateral Filter and alternative 2x2 classifier for ALF) | Bytedance  [Wenbin Yin](mailto:yinwenbin.hit@bytedance.com)  Qualcomm  Nan Hu  [JVET-X0069](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0069-v1.zip) | Ericsson  [Jacob Strom](mailto:jacob.strom@ericsson.com) |
| 4.9b | Combination of 4.1 and 4.5 (Chroma Bilateral Filter and alternative sample-based classifier for ALF) | Bytedance  [Wenbin Yin](mailto:yinwenbin.hit@bytedance.com)  Qualcomm  Nan Hu  [JVET-X0069](https://jvet-experts.org/doc_end_user/documents/24_Teleconference/wg11/JVET-X0069-v1.zip) | Ericsson  [Jacob Strom](mailto:jacob.strom@ericsson.com) |

1. **Description of tests**

***Partitioning***

**Test 1.1: Encoder optimization for ECM**

In this test, maximum BT and TT depths are increased per class from 32 to up to 128, MTT hierarchical depth is increased to 4 in I slices and MTT hierarchical depth is temporal ID dependent for P- and B- slices varying between 2 and 3 as summarized in the following table.

|  |  |  |
| --- | --- | --- |
| Class A | Class B | Classes C and D |
| CTUSize: 256  MaxMTTHierarchyDepth: 3  MaxMTTHierarchyDepthI: 4  MaxMTTHierarchyDepthISliceL: 4  MaxMTTHierarchyDepthISliceC: 4  MaxBTLumaISlice: 128  MaxBTChromaISlice: 128  MaxBTNonISlice: 128  MaxTTLumaISlice: 128  MaxTTChromaISlice: 128  MaxTTNonISlice: 128  MaxABTLumaISlice: 128  MaxABTChromaISlice: 128  MaxABTNonISlice: 128  MaxMTTHierarchyDepthByTid: 333322 | CTUSize: 128  MaxMTTHierarchyDepth: 3  MaxMTTHierarchyDepthI: 4  MaxMTTHierarchyDepthISliceL: 4  MaxMTTHierarchyDepthISliceC: 4  MaxBTLumaISlice: 128  MaxBTChromaISlice: 128  MaxBTNonISlice: 128  MaxTTLumaISlice: 128  MaxTTChromaISlice: 128  MaxTTNonISlice: 128  MaxABTLumaISlice: 128  MaxABTChromaISlice: 128  MaxABTNonISlice: 128  MaxMTTHierarchyDepthByTid: 333322 | CTUSize: 128  MaxMTTHierarchyDepth: 3  MaxMTTHierarchyDepthI: 4  MaxMTTHierarchyDepthISliceL: 4  MaxMTTHierarchyDepthISliceC: 4  MaxBTLumaISlice: 64  MaxBTChromaISlice: 64  MaxBTNonISlice: 128  MaxTTLumaISlice: 32  MaxTTChromaISlice: 32  MaxTTNonISlice: 64  MaxABTLumaISlice: 64  MaxABTChromaISlice: 64  MaxABTNonISlice: 64  MaxMTTHierarchyDepthByTid: 333333 |

In cross-check report JVET-X0144, other variations of the configuration parameters are tested for classes A and B.

All provided tests result are summarized in the next tables. In trade-offs of JVET-X0068, normative MTT depth for temporal layers signalling in SPS is introduced with additional speed-up method is added, which allows to bypass certain mode checks at encoder.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Class A | | Class B | |
|  | MaxMTTHierarchyDepth | MaxMTTHierarchyDepthByTid (0-5) | MaxMTTHierarchyDepth | MaxMTTHierarchyDepthByTid (0-5) |
| Trade-off 1 | 3 | 3 3 3 3 2 2 | 3 | 3 3 3 3 2 2 |
| Trade-off 2 | 3 | 3 3 3 3 3 2 | 3 | 3 3 3 3 3 2 |
| Trade-off 3 | 3 | 3 3 3 3 3 3 | 3 | 3 3 3 3 3 3 |

In cross-check report JVET-X0144 the trade-offs listed in the next table of MTTDepth dependency on temporal ID encoder only feature alone (MAX\_MTT\_DEPTH\_BY\_TID macro) which is in the EE2-Test1.1-1.4 SW.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Class A | | Class B | |
|  | MaxMTTHierarchyDepth | MaxMTTHierarchyDepthByTid (0-5) | MaxMTTHierarchyDepth | MaxMTTHierarchyDepthByTid (0-5) |
| Trade-off 1 | 3 | 3 3 3 3 2 2 | 3 | 3 3 3 3 2 2 |
| Trade-off 2 | 4 | 3 3 3 3 2 2 | 4 | 3 3 3 3 2 2 |
| Trade-off 3 | 3 | 3 3 3 3 3 3 | 3 | 3 3 3 3 3 3 |

**Test 1.2: 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.

Encoder optimization of the test 1.1along with the new partition specific speedups are applied.

The listed below trade-off options are tested.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Class A | | Class B | |
|  | MaxMTTHierarchyDepth | MaxMTTHierarchyDepthByTid (0-5) | MaxMTTHierarchyDepth | MaxMTTHierarchyDepthByTid (0-5) |
| Trade-off 1 | 3 | 3 3 3 3 2 2 | 3 | 3 3 3 3 2 2 |
| Trade-off 2 | 3 | 3 3 3 3 3 2 | 3 | 3 3 3 3 3 2 |
| Trade-off 3 | 3 | 3 3 3 3 3 3 | 3 | 3 3 3 3 3 3 |

In addition, in trade-off 3, block size 24 is allowed, but in the two other trade-offs only the added block size 12 is allowed in addition to dyadic block sizes of ECM-2.0.

**Test 1.3: 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.

Encoder optimization of the test 1.1along with the new partition specific speedups are applied.

|  |  |
| --- | --- |
|  |  |
| UQT-H1 | UQT-H2 |
|  |  |
| UQT-V1 | UQT-V2 |

**Figure 2. UQT partitions**

**Test 1.4: Asymmetric binary tree (ABT) and Unsymmetric quad tree (UQT)**

This test is a combination of the tests 1.2 and 1.3.

From discussion in JVET:

{

Results on RA:



From the results, it is difficult to conclude clear benefits of introducing additional partitioning schemes. It highly depends on the amount of encoder optimization. For example, JVET-X0144 with trade-offs 1 and 2 would be better complexity/performance tradeoff than ABT with similar trade-offs, whereas for trade-off 3, ABT has less encoder run-time than X0068 with similar compression benefit (both significantly increasing encoder run-time, though).

There seem to be some differences in the implementation of the three trade-offs (e.g. MTT hierarchical depth depending on TID, or increased, or optimized per class).

For UQT which uses a similar optimization as trade-off 1, also better complexity/performance tradeoffs could be identified with encoder-only solutions.

Combination of ABT/UQT (1.4) is not completed yet. ABT with LD is not completed yet, but neither UQT nor encoder-only (X0144) show promising performance in case of LD.

Most of the gain seems to be due to various degrees of encoder optimization. A possible way would be to use one of the encoder-only methods (e.g., X0144) as an anchor, and test additional partitioning relative to that. X0144 trade-off 1 or 2 could also be attractive for next ECM (as encoder trick).

Revisit after the cross-check of X0144 is finalized, and results on 1.4 (combination UQT/ABT) and ABT LD are available.

}

***Inter prediction***

**Tests 3.1: Combination of CIIP and DIMD/TIMD**

In the tests, DIMD or TIMD methods are used to derive an intra prediction mode and replace the planar mode to generate the intra prediction signal in CIIP mode without extra signaling. The proposed method is only applied to coding blocks with an area less than or equal to 1024. And the propagated intra prediction mode of a CIIP coding block is always set to planar mode.

In the tests, the weights (wIntra, wInter) are modified if the derived intra prediction mode is an angular mode. For near-horizontal modes (2 <= angular mode index < 34), the current block is vertically divided as shown in Figure 3 (a); for near-vertical modes (34 <= angular mode index <= 66), the current block is horizontally divided as shown in Figure 3 (b).

**  
Figure 3. The block division for angular modes**

The (wIntra, wInter) for different sub-blocks are derived as follows.

|  |  |
| --- | --- |
| **The sub-block index** | **(wIntra, wInter)** |
| 0 | (6, 2) |
| 1 | (5, 3) |
| 2 | (3, 5) |
| 3 | (2, 6) |

In DIMD intra mode derivation, the intra prediction mode with the highest amplitude in the DIMD histogram is selected.

In TIMD intra mode derivation, the intra prediction mode with the smallest SATD value in the TIMD mode list is selected and mapped to one of the 67 regular intra prediction modes.

Test 3.1a: Combination of CIIP and DIMD

Test 3.1b: Combination of CIIP and TIMD

**Tests 3.2: GPM with inter and intra prediction**

In the test, one of the GPM partitions can be inter or intra, a flag is signaled to indicate whether intra or inter prediction is used in a partition.

For the inter partition, a uni-directional predictor is generated by MVs from the merge candidate list. For the intra partition, an intra predictor is generated from the neighboring samples of the CU, pre-defined intra prediction directions considering GPM partitioning edge are selected and intra mode index is signalled. Finally, these two predictors are blended in the same way as in the existed GPM.

Test 3.2a: GPM with inter and intra prediction – method A.

Pre-defined intra prediction modes are restricted to the parallel and perpendicular angular modes relative to GPM partitioning edge.

Test 3.2b: GPM with inter and intra prediction – method B.

Pre-defined intra prediction modes are restricted to the parallel and perpendicular angular modes relative to GPM partitioning edge, and planar mode.

**Tests 3.3: AMVP-merge mode**

AMVP-merge mode is tested where bi-directional predictor is composed of an AMVP predictor in one direction and a merge predictor in the other direction.

Uni directional AMVP predictor is signalled as in the regular AMVP mode. Merge index is not signalled, and merge predictor is selected from the candidate list with the smallest template or bilateral matching cost and both approaches are tested.

When the selected merge predictor and the AMVP predictor satisfy DMVR condition, which is there is at least one reference picture from the past and one reference picture from the future relatively to the current picture and the distances from two reference pictures to the current picture are the same, the bilateral matching MV refinement is applied for the merge MV candidate and AMVP MVP as a starting point. Otherwise, if template matching functionality is enabled, template matching MV refinement is applied to the merge predictor or the AMVP predictor which has a higher template matching cost.

The third pass which is 8x8 sub-PU BDOF refinement of the multi-pass DMVR is enabled to AMVP-merge mode coded block.

Additionally, the introduced AMVP-merge mode is tested when template matching is disabled in the anchor and the test.

Test 3.3a: Bilateral matching AMVP-merge mode

Test 3.3b: Bilateral matching AMVP-merge mode with TM disabled

Test 3.3c: Template matching AMVP-merge mode

**Tests 3.4: Adaptive decoder side motion vector refinement**

In ECM, multi-pass DMVR refines both directions (L0 and L1) in a symmetrical manner as shown on Figure 4.

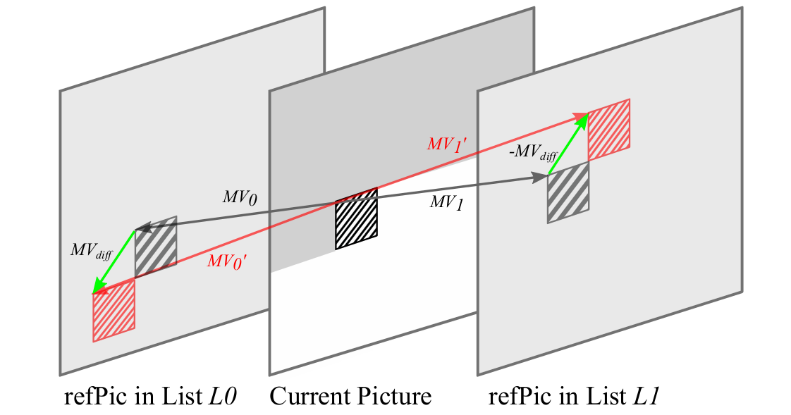


Figure 4. DMVR refinement

In the test, adaptive DMVR is introduced where it can refine only one either L0 or L1 direction of the bi-prediction. Those two additional modes are added to the existed DMVR and the choice of DMVR mode is signalled.

Like the regular merge mode, merge candidates for the proposed merge modes are derived from the spatial neighboring coded blocks, TMVPs, non-adjacent blocks, HMVPs, and pair-wise candidate. The difference is that only those meet DMVR conditions are added into the candidate list. The same merge candidate list is used by the two proposed merge modes and merge index is coded as in regular merge mode.

Test 3.4a: Adaptive decoder side motion vector refinement

Test 3.4b: Adaptive decoder side motion vector refinement with TM disabled

From discussion in JVET:

{

RA results (from X0024 Excel v2)

LB results (from X0024 Excel v2)

“TM disabled” means it is disabled also in the ECM2 anchor. This incurs a loss of approx. 2.5%.

3.1. shows mainly benefit for LB (in particular 3.1b)

Question on 3.1: How large is the gain due to subblock adaptation of weights? Answer: There is some gain, and complexity increase is marginal.

It is also mentioned that there are EE3.1 related contributions which improve performance. Revisit after review of those contributions.

3.2 shows no relevant benefit (except for LP, around 0.4% luma but with 5% encoder runtime increase), and there are EE related contributions showing more gain. 3.2 as from the EE are not showing sufficient benefit for taking action.

3.3: The version with template matching (3.3c) shows better performance than bilateral matching (3.3a), however it has a non-neglibible complexity impact on the construction of the merge list. Further, if template matching would be banned from ECM entirely, bilateral matching would give back some of the gain).

3.4 uses bilateral matching for adaptively deciding whether the MV in L0 or L1 should be refined (where ECM always refines both symmetrically by the same amount). This is signalled. It shows some additional benefit on top of TM, but even larger benefit if template matching would be banned from ECM (compensating partially for the loss that would occur then).

Note: The decrease in decoder run time in 3.3 and 3.4 is due to some software optimization (without change of results) which was mainly done to keep the encoder run time lower. If the optimization was likewise applied in ECM, the encoder run time would increase by another 2% for all 3.3 and 3.4 cases (results with a software optimized anchor are provided in the proposals X0083 and X0049).

At the current stage of exploration, removing TM from ECM is not desirable, as it would incur a significant loss, and implementability studies should rather be performed in the context of a real standard’s development. In the other hand, if it was removed at a later stage, neither 3.3a nor 3.4a would need to be modified. Though no combination tests were run, it is likely that the two tools are not in direct competition, as 3.3 is targeting reduction of merge signaling, and 3.4 is targeting improvement of the prediction.

Decision: Adopt 3.3a from X0083 and 3.4a from X0049. Also the software optimization for bilateral matching should be included (as well for other parts where bilateral matching is used in ECM).

}

***In-loop filtering***

**Test 4.1: Chroma bilateral filter**

Luma bilateral inloop filter (BIF) is extended to chroma component in parallel with SAO and CCSAO as shown on Figure 5.

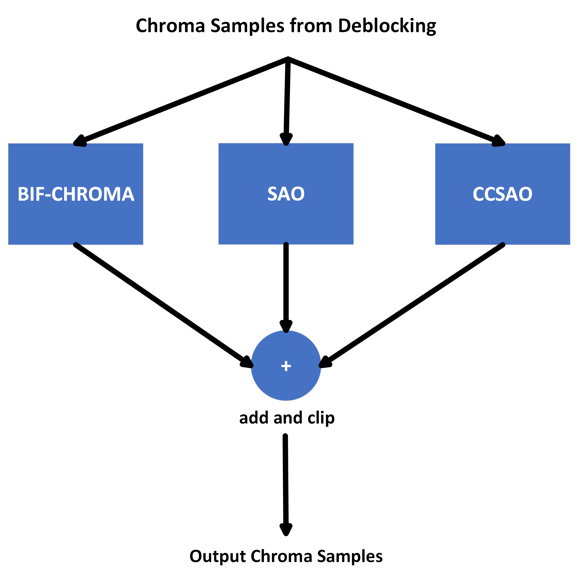


Figure 5. Chroma bilateral filter

Chroma BIF can be switched on/off at CTU and slice levels. The filtering process of luma BIF, 5×5 diamond shape filter is used for generating the filtering offset. The difference between the central sample and each surrounding sample is calculated first. The coefficient for each reference sample is extracted from a pre-defined look-up-table based on the calculated difference directly. The coefficients used for chroma components are retrained, different from those from BIF-luma.

In luma BIF, the block-level filtering strength parameter is determined based on luma TU size and CU mode. While in chroma BIF, the parameter for chroma components is determined based the chroma TU size and mode when dual-tree partitioning is enabled for the current slice, and based on the corresponding luma TU size and mode when dual-tree partitioning is disabled.

**Tests 4.2: CTB level filter shape selection of CCALF**

In ECM-2.0, CCALF uses 8-tap hexagon shaped filter illustrated in Figure 6.

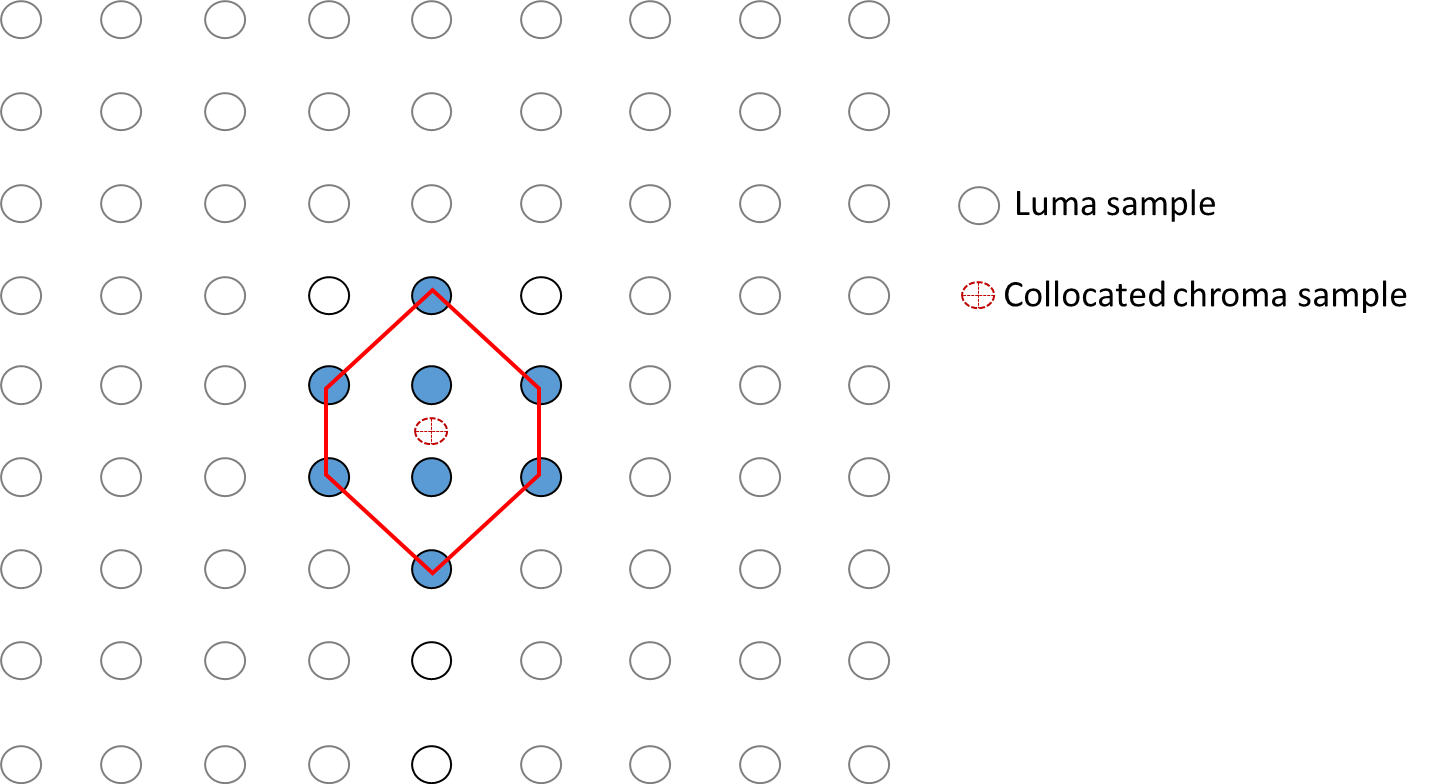


Figure 6. CCALF filter shape in ECM-2.0

In the EE, additional two filter shapes are tested as shown on Figure 7. Filter coefficients are carried in APS, up to 16 filters can be signalled, and for each filter shape index (shape\_idx) is added. For each CTB, an index is signalled to the decoder to specify which filter shape and coefficients are used for that CTB.

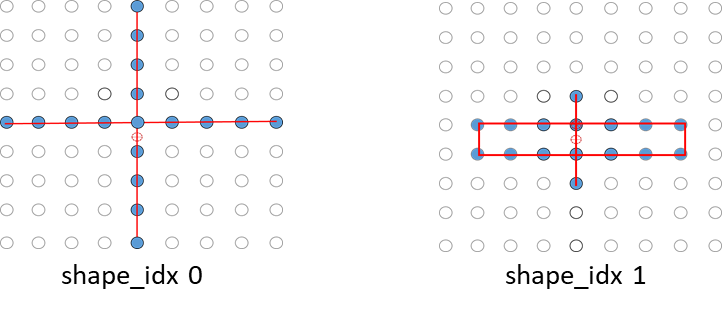


Figure 7. Additional CCALF filter shapes

In ECM-2.0, CCALF filter coefficients are restricted to having only values in the form of power of 2, from the following set: {-64, -32, -16, -8, -4, -2, -1, 0, 1, 2, 4, 8, 16, 32, 64}. Additionally, removal of power-of-2 constraint on CCALF filter coefficients is tested.

Test 4.2a: CTB level filter shape selection of CCALF

Test 4.2b: CTB level filter shape selection of CCALF with removal of power of 2 constraint of filter coefficients

**Tests 4.3: CCALF with larger filter size**

In this test, 25-tap large filter illustrated on Figure 7 is used replacing 8-tap hexagon filter and up to 16 filters are signalled in APS.

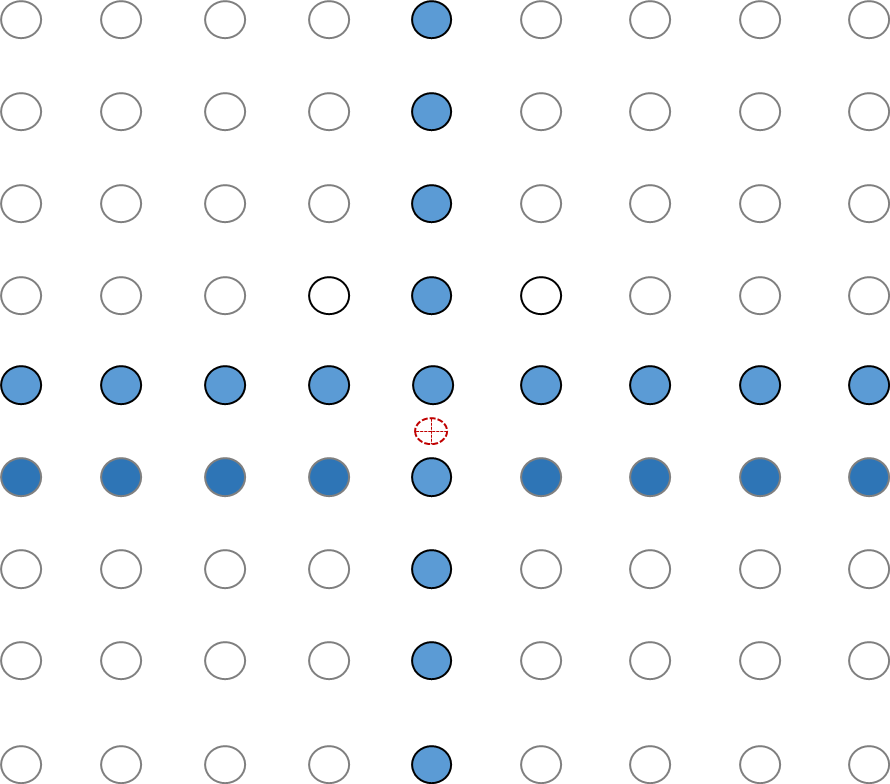


Figure 7. 25-tap CCALF filter shapes

Additionally, removal of power-of-2 constraint on CCALF filter coefficients is tested.

Test 4.3a: CCALF with larger filter size

Test 4.3b: CCALF with larger filter size with removal of power of 2 constraint of filter coefficients

**Tests 4.4 – 4.5: Alternative classifiers for ALF**

In ECM-2.0, ALF uses Laplacian based classifier applied to each 2x2 luma block in a CTB. For each class, a specific filter is applied among the different filters in a filter set carried in APS. Based on the classification, geometric transformation, such as 90-degree rotation, diagonal or vertical flip, of the coefficients within a diamond filter shape can also be applied.

In the test, another band-based classifier is introduced. A flag is added to a luma filter set in APS to indicate whether the classifier of ECM-2.0 or the one of the alternative classifiers is applied. Geometrical transformation is not applied with the tested band classifier. The number of classes is equal to 25 as in the existed ECM ALF classifier.

In the 2x2 block based band-based classifier, the sum of sample values of a 2x2 luma block is calculated at first. Then the class index is calculated as below,

class\_index = (sum \* 25) >> (sample bit depth + 2).

In the sample band-based classifier, the class index of a luma sample is calculated as

class\_index = (luma sample \* 25) >> (sample bit depth).

In test 4.5, the frequency of filter switch using band classifier in 2x2 luma blocks is summarized in the table below. The ratio is calculated by dividing the number of 2x2 blocks having different filters by the total number of 2x2 blocks using band classifier.

|  |  |  |  |
| --- | --- | --- | --- |
|  | AI | RA | LDB |
| Class A1 | 8.96% | 3.15% |  |
| Class A2 | 19.09% | 5.70% |  |
| Class B | 11.35% | 5.72% | 5.47% |
| Class C | 8.47% | 2.24% | 3.07% |
| Class E | 5.33% |  | 3.17% |
| **Overall** | 10.49% | 4.27% | 3.98% |
| Class D | 4.20% | 4.45% | 5.17% |
| Class F | 19.90% | 11.42% | 6.16% |

Test 4.4: Alternative 2x2 classifier for ALF

Test 4.5: Alternative sample-based classifier for ALF

**Tests 4.6 – 4.9: Combination tests**

The following combinations of the in-loop filtering methods are tested:

Test 4.6a: Combination of 4.2a and 4.4 (CTB level filter shape selection of CCALF and alternative 2x2 classifier for ALF)

Test 4.6b: Combination of 4.2a and 4.5 4 (CTB level filter shape selection of CCALF and alternative sample-based classifier for ALF)

Test 4.6c: Combination of 4.3a and 4.4 (CCALF with larger filter size and alternative 2x2 classifier for ALF)

Test 4.6d: Combination of 4.3a and 4.5 (CCALF with larger filter size and alternative sample-based classifier for ALF)

Test 4.7a: Combination of 4.2a and 4.1 (CTB level filter shape selection of CCALF and chroma bilateral filter)

Test 4.7b: Combination of 4.3a and 4.1 (CCALF with larger filter size and chroma bilateral filter)

Test 4.8a: Combination of 4.6a and 4.1 (CTB level filter shape selection of CCALF, alternative 2x2 classifier for ALF, and chroma bilateral filter)

Test 4.8b: Combination of 4.6b and 4.1 (CTB level filter shape selection of CCALF, alternative sample-based classifier for ALF, and chroma bilateral filter)

Test 4.8c: Combination of 4.6c and 4.1 (CCALF with larger filter size, alternative 2x2 classifier for ALF, and chroma bilateral filter)

Test 4.8d: Combination of 4.6d and 4.1 (CCALF with larger filter size, alternative sample-based classifiers for ALF, and chroma bilateral filter)

Test 4.9a: Combination of 4.1 and 4.4 (Chroma bilateral filter and alternative 2x2 classifier for ALF)

Test 4.9b: Combination of 4.1 and 4.5 (Chroma bilateral filter and alternative sample-based classifier for ALF)

Results AI:



Results RA:

Results LB:

4.1-4.3 target only gain for chroma

4.2. and 4.3 are mutually exclusive

4.4 and 4.5 are mutually exclusive

Combination tests show that the gains 4.1, 4.2|4.3, 4.4|4.5 are approximately additive.

In terms of results, the benefits of the mutually exclusive approaches are almost negligible.

4.4 is clearly less complex (in terms of possible frequent switching of filters) than 4.5 (2x2 vs. pixel based classification)

4.2 could have some advantage for software (less filter taps), but requires a decision for one of the two filters at the encoder (where already a fast algorithm is implemented).

4.3 is conceptually simpler and also simpler to understand.

Removal of the power-of-2 constraint shows only minor benefit in both 4.2. and 4.3 (b variants), but concern is raised about possible complexity.

Decision: Adopt combination 4.1/4.3a/4.4 (called 4.8.c above)

[JVET-X0045](https://jvet-experts.org/doc_end_user/current_document.php?id=11038) EE2-4.2, EE2-4.3: On CCALF filter [M. G. Sarwer, R.-L. Liao, J. Chen, Y. Ye, X. Li (Alibaba)]

[JVET-X0046](https://jvet-experts.org/doc_end_user/current_document.php?id=11039) EE2-4.6: Joint ALF and CCALF tests [M. G. Sarwer, R.-L. Liao, J. Chen, Y. Ye, X. Li (Alibaba), N. Hu, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-X0047](https://jvet-experts.org/doc_end_user/current_document.php?id=11040) EE2-4.7: Joint test on CCALF and chroma-bilateral filter [M. G. Sarwer, R.-L. Liao, J. Chen, Y. Ye, X. Li (Alibaba), W. Yin, K. Zhang, L. Zhang (Bytedance)]

[JVET-X0049](https://jvet-experts.org/doc_end_user/current_document.php?id=11042) EE2: Adaptive decoder side motion vector refinement (test 3.4) [H. Huang, Z. Zhang, V. Seregin, W.-J. Chien, C.-C. Chen, M. Karczewicz (Qualcomm)]

[JVET-X0067](https://jvet-experts.org/doc_end_user/current_document.php?id=11060) EE2-4.1: Bilateral Inloop Filter on Chroma [W. Yin, K. Zhang, L. Zhang (Bytedance)]

[JVET-X0068](https://jvet-experts.org/doc_end_user/current_document.php?id=11061) EE2-1.1~EE2-1.4: Tests on unsymmetric partitioning methods [K. Zhang, L. Zhang, Z. Deng, N. Zhang, Y. Wang (Bytedance), F. Le Léannec, K. Naser, T. Dumas, A. Robert, F. Galpin, E. François (InterDigital)]

[JVET-X0069](https://jvet-experts.org/doc_end_user/current_document.php?id=11062) EE2-4.9: A combining test of EE2-4.1 and EE2-4.4/4.5 [W. Yin, K. Zhang, L. Zhang (Bytedance), N. Hu, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-X0070](https://jvet-experts.org/doc_end_user/current_document.php?id=11063) EE2: Alternative classifiers for ALF (tests 4.4 and 4.5) [N. Hu, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-X0071](https://jvet-experts.org/doc_end_user/current_document.php?id=11064) EE2-4.8: Joint tests of chroma BIF, ALF and CCALF [N. Hu, V. Seregin, M. Karczewicz (Qualcomm), M. G. Mohammed, R.-L. Liao, J. Chen, Y. Ye, X. Li (Alibaba), W. Yin, K. Zhang, L. Zhang (Bytedance)]

[JVET-X0077](https://jvet-experts.org/doc_end_user/current_document.php?id=11070) EE2-3.2: GPM with inter and intra prediction (JVET-W0110) [Y. Kidani, H. Kato, K. Kawamura (KDDI)]

[JVET-X0083](https://jvet-experts.org/doc_end_user/current_document.php?id=11076) EE2: Bilateral and template matching AMVP-merge mode (test 3.3) [Z. Zhang, H. Huang, C.-C. Chen, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-X0098](https://jvet-experts.org/doc_end_user/current_document.php?id=11091) EE2-3.1: Combination of CIIP and DIMD/TIMD [X. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba)]

[JVET-X0177](https://jvet-experts.org/doc_end_user/current_document.php?id=11186) Crosscheck of JVET-X0098 (EE2-3.1: Combination of CIIP and DIMD/TIMD) [[Y. Wang (Bytedance)](mailto:wangyang.cs@bytedance.com)] [late]

[JVET-X0144](https://jvet-experts.org/doc_end_user/current_document.php?id=11137) EE2: Encoder partitioning optimization for ECM and crosscheck of EE2-1.1 [M. Coban, V. Seregin, M. Karczewicz (Qualcomm), K. Zhang, L. Zhang, Z. Deng, N. Zhang, Y. Wang (Bytedance), F. Le Léannec, K. Naser, T. Dumas, A. Robert, F. Galpin, E. François (InterDigital)]

TBP

[JVET-X0196](https://jvet-experts.org/doc_end_user/current_document.php?id=11206) Crosscheck of JVET-X0144 (EE2: Encoder partitioning optimization for ECM and crosscheck of EE2-1.1) [T. Nguyen (HHI)] [late]

### EE2 related contributions (15)

Contributions in this area were discussed in session 8 at 1640–1725 UTC on Thursday 7 Oct. 2021 (chaired by JRO), session 12 1520-1720 UTC on Friday 8 Oct. 2021 (chaired by JRO), and session 15 1520-1615 UTC on Monday 11 Oct. 2021 (chaired by JRO).

[JVET-X0056](https://jvet-experts.org/doc_end_user/current_document.php?id=11049) EE2-Related: Complexity reduction for decoder side motion derivation [H. Huang, V. Seregin, Z. Zhang, C.-C. Chen, M. Karczewicz (Qualcomm)]

This contribution presents complexity reduction methods for the decoder side motion derivation tools in ECM-2.0. In the first test, early termination is applied to the DMVR and TM process. It reports 0.02%, 0.11%, 0. 11% BD rate for luma and chroma components, respectively, with 96% encoder runtime and 90% decoder runtime in random access configurations of common test condition. In the second test, on top of first test, the maximum number of iterations in square search process of DMVR is reduced and DMVR is disabled if a merge candidate has inherited additional hypothesis. It reports 0.xx%, 0.xx%, 0. xx% BD rate for luma and chroma components, respectively, with xx% encoder runtime and xx% decoder runtime.

Good tradeoff in runtime vs. performance. Test 2 is expected to provide more encoder run time reduction.

Revisit after results are complete and cross-checked.

[JVET-X0193](https://jvet-experts.org/doc_end_user/current_document.php?id=11203) Crosscheck of JVET-X0056 (EE2-Related: Complexity reduction for decoder side motion derivation) [R.-L. Liao (Alibaba)] [late] [miss]

[JVET-X0072](https://jvet-experts.org/doc_end_user/current_document.php?id=11065) EE2-related: PDPC-skip scheme for angular intra modes [J. Zhang, C. Zhou, Z. Lv (vivo)]

In current ECM-2.0 PDPC is applied to calculate the prediction samples for the angular prediction modes smaller than or equal to mode 18 (Horizontal), and the modes greater than mode 55 (Vertical). This contribution presents a PDPC skip scheme, which makes use of the results of texture analysis in DIMD method to decide whether skip PDPC or not. In this way, some of the PDPC process will be skipped. Thus, the computational complexity can be reduced at both encoder and decoder sides. For method1 the skipping rate comes up to 26.68% in our simulation, with the average BD-rate loss Y 0.02%, U 0.06% and V 0.06% on average. For method2 the skipping rate comes up to 10.24% in our simulation, the complexity of 99% (AI) at the decoder, with the average BD-rate loss Y 0.02%, U 0.01% and V 0.05% on average. For method3 the skipping rate comes up to 38.03% in our simulation. On top of the ECM-2.0 [1], the simulation results show the complexity of 99% (AI) at the decoder, with the average BD-rate loss Y 0.06%, U 0.12% and V 0.14% on average.

PDPC is not of concern in terms of complexity. In fact, the proposal would even slightly increase the worst-case complexity due to the additional process to decide about PDPC usage. Benefit in terms of compression only for screen content, whereas minor loss is observed for camera captured content, while the reduction in decoder runtime is marginal.

No action.

[JVET-X0180](https://jvet-experts.org/doc_end_user/current_document.php?id=11190) Crosscheck of JVET-X0072 (EE2-related: PDPC-skip scheme for angular intra modes) [[Z. Deng (Bytedance)](mailto:zhipin.deng@bytedance.com)] [late]

[JVET-X0078](https://jvet-experts.org/doc_end_user/current_document.php?id=11071) EE2-related: Modified GPM with inter and intra prediction [Y. Kidani, H. Kato, K. Kawamura (KDDI)]

Include abstract

No separate presentation, see notes under X0166

[JVET-X0086](https://jvet-experts.org/doc_end_user/current_document.php?id=11079) EE2-related: Adaptive Filter Shape Selection for ALF [M. G. Sarwer, R.-L. Liao., J. Chen., Y. Ye., X. Li. (Alibaba), W. Yin, K. Zhang, L. Zhang (Bytedance), N. Hu, V. Seregin, M. Karczewicz (Qualcomm)]

In the current ALF design, the online-trained filters only have one fixed shape. In this contribution, a CTB-level adaptive filter shape selection (AFSS) method for ALF is proposed. With AFSS, a filter shape index that is associated with luma filters is signalled for each APS. Each CTB could select the best luma filter shape from the available APSs. In EE2-4.8, BIF-Chroma, filter shape selection for CCALF and alternative classifier for ALF are combined. In this contribution, proposed AFSS is further combined with EE2-4.8 to seek an additive coding gain.

On top of ECM-2.0, simulation results (BD-rate changes, encoding time and decoding time) are reported as below:

Combining with EE2-Test4.8a:

AI: -0.18%, -2.55%, -3.31%, 105%, 101%.

RA: -0.47%, -2.12%, -2.26%, 105%, 102%.

LB: -0.71%, -3.58%, -3.76%, 107%, 102%.

Combining with EE2-Test4.8b:

AI: -0.23%, -2.53%, -3.29%, 104%, 100%.

RA: -0.50%, -2.13%, -2.20%, 105%, 102%.

LB: -0.82%, -3.54%, -3.90%, 106%, 100%.

Combining with EE2-Test4.8c:

AI: -0.16%, -2.66%, -3.47%, 105%, 102%.

RA: -0.48%, -2.02%, -2.19%, 105%, 102%.

LB: -0.78%, -3.25%, -3.48%, 106%, 101%.

Combining with EE2-Test4.8d:

AI: -0.22%, -2.64%, -3.45%, 105%, 101%.

RA: -0.52%, -2.00%, -2.18%, 105%, 101%.

LB: -0.82%, -3.60%, -3.41%, 107%, 100%.

Relative to the adoption from EE (4.8c), the additional gain is 0.05% in AI, and 0.12% in RA. Encoding time increases by approx. 4% for both cases.

Not good tradeoff from encoding point of view.

Investigate in EE.

[JVET-X0175](https://jvet-experts.org/doc_end_user/current_document.php?id=11184) Cross-check of JVET-X0086 "EE2-related: Adaptive Filter Shape Selection for ALF" [[F. Le Léannec (InterDigital)](mailto:fabrice.leleannec@interdigital.com)] [late]

[JVET-X0100](https://jvet-experts.org/doc_end_user/current_document.php?id=11093) EE2-related: On propagating intra prediction mode for IBC [L.-F. Chen, X. Li, L. Li, S. Liu (Tencent)]

In order to improve the accuracy of the construction of MPM list, intra prediction mode propagation for inter CU was adopted in JVET-W meeting and ECM-2.0 reference software. However, the proposed intra mode propagation is applied on inter CU only. In this proposal, we propose to simply extend the idea to IBC blocks. The overall performance impact of the proposed method on top of ECM-2.0 for SCC content is reported as below.   
For Class-F test sequence:

* AI: -0.06%/-0.04%/-0.02% with 100% EncT and 100% DecT
* RA:-0.02%/-0.08%/-0.07% with 101% EncT and 101% DecT

For Class-TGM test sequence:

* AI: -0.05%/-0.08%/-0.01% with 101% EncT and 101% DecT
* RA: 0.00%/-0.06%/-0.04% with 101% EncT and 101% DecT

Only tiny gain (very low for the case of screen content). However, the method is straightforward, does not introduce any impact ob IBC latency, and just uses the same approach to fetch an intra mode from the IBC reference block (using the current block’s vector) as it is otherwise done using the motion vector in the case of inter CUs. This could at most be asserted to make the design more consistent.

No action.

[JVET-X0176](https://jvet-experts.org/doc_end_user/current_document.php?id=11185) Crosscheck of JVET-X0100 (EE2-related: On propagating intra prediction mode for IBC) [[W. Chen (Kwai)](mailto:chenwei06@kwai.com)] [late]

[JVET-X0114](https://jvet-experts.org/doc_end_user/current_document.php?id=11107) EE2-related: Fix on issues of TIMD mode [C. Zhou, Z. Lv, J. Zhang (vivo)]

This contribution reports two issues of TIMD mode in ECM2.0. The first issue is that it is possible for the two intra modes to be the same. The second issue is that the condition for deciding whether the fusion is applied is not set correctly.

On top of ECM-2.0, the performance is reported:

AI: -x.xx% (Y) x.xx% (U) -x.xx% (V), runtimes 10x% (enc) 10x% (dec)  
RA: -x.xx% (Y) x.xx% (U) -x.xx% (V), runtimes 10x% (enc) 10x% (dec)

The first issue does not appear to be a problem, as the performance is not significantly affected (initial results indicate gains in the range of 0.01%). The second issue may be a problem (formula and implementation in software not equivalent when unsigned integer is used). However, according to the crosschecker, a more appropriate solution (rather than swapping) would be direct implementation of the formula (costB<2\*costA).

Decision(SW/BF): A bug fix is needed, but not as suggested in the proposal, but by direct implementation of the “step 3” formula. It should be submitted as a software issue.

[JVET-X0155](https://jvet-experts.org/doc_end_user/current_document.php?id=11162) EE2-related: Crosscheck of JVET-X0114 (Fix on issues of TIMD mode) [K. Cao (Qualcomm)] [late]

[JVET-X0162](https://jvet-experts.org/doc_end_user/current_document.php?id=11169) Crosscheck of JVET-X0114 (EE2-related: Fix on issues of TIMD mode) [[X. Li (Alibaba)](mailto:sid.lxw@alibaba-inc.com)] [late]

[JVET-X0115](https://jvet-experts.org/doc_end_user/current_document.php?id=11108) EE2-related: Optimization on the second mode derivation of DIMD blending mode [C. Zhou, Z. Lv, J. Zhang (vivo)]

In this contribution, a method to optimize the second mode in the DIMD blending mode is proposed. The second mode is derived from the second and third largest histogram amplitude values by template-based method as in TIMD.

On top of ECM-2.0, the performance is reported:

AI: -0.02% (Y) 0.00% (U) 0.00% (V), runtimes 101% (enc) 102% (dec)  
RA: -0.02% (Y) 0.01% (U) -0.01% (V), runtimes 10x% (enc) 10x% (dec)

Small gain does not justify action.

[JVET-X0163](https://jvet-experts.org/doc_end_user/current_document.php?id=11170) Crosscheck of JVET-X0115 (EE2-related: Optimization on the second mode derivation of DIMD blending mode) [X. Li (Alibaba)] [late]

[JVET-X0121](https://jvet-experts.org/doc_end_user/current_document.php?id=11114) EE2-related: bug fixes for enabling RPR in ECM [P. Bordes, F. Galpin, F. Le Léannec, E. François (InterDigital)]

The ECM-2.0 software includes RPR functions and signaling as implemented in VVC. However, the RPR mode is broken. This contribution proposes some simple bug fixing to enable RPR in ECM-2.0 in same way as in VTM. The modified ECM-2.0 code is provided with macro RPR\_ENABLE.

Some tools that were implemented in ECM were not compatible with RPR – this is fixed in this contribution. RPR benefit on top of ECM 2 is comparable to that on top of VTM (as per contribution X0117 which provides RPR results as comparison point for NN based superresolution).Decision(SW/BF): Adopt JVET-X0121

One expert mentions that also other VVC functionalities such as subpictures are not working properly in ECM.

[JVET-X0132](https://jvet-experts.org/doc_end_user/current_document.php?id=11125) EE2-Related: On MVD sign prediction [Y. Zhang, B. Ray, H. Huang, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution presents the results of motion vector difference sign prediction, which is initially proposed in JVET-K0067, to further include affine mode, SMVD mode, MMVD mode and affine MMVD mode. The proposed method was implemented on top of ECM-2.0, and it reports -0.11%, -0.11%, -0.10% BD rate saving for luma and chroma components, respectively, with 100% encoder and 99% decoder run time in random access configurations and -0.xx%, -0.xx%, -0.xx% BD rate saving for luma and chroma components, respectively, with xxx% encoder and xxx% decoder run time in low delay B configurations of common test condition.

The original contribution K0067 had roughly 0.5% gain. How much is the gain due to extending it to affine etc.? Roughly half. Probably some overlap with other MV tools.

TM is used to derive the sign prediction (as was the case in the original contribution), which is at another stage here than other TM. Could this be considered to be avoided, e.g. by replacing it by bilateral matching?Context coding is considering magnitude of motion vectors (4 contexts).

Investigate in EE.

[JVET-X0200](https://jvet-experts.org/doc_end_user/current_document.php?id=11210) Crosscheck of JVET-X0132 (Non-EE2: On MVD sign prediction) ]M. Salehifar (Bytedance)] [late]

[JVET-X0133](https://jvet-experts.org/doc_end_user/current_document.php?id=11126) EE2-related: MV candidate type-based ARMC [Y.-J. Chang, H. Huang, V. Seregin, C.-C. Chen, M. Karczewicz (Qualcomm), R.-L. Liao, J. Chen, Y. Ye, X. Li (Alibaba)]

In this contribution, an MV candidate type-based ARMC is proposed to reorder the merge candidates in a candidate type, e.g., TMVP, NA-MVP, etc., based on the TM cost values. It is operated before the merge candidate list is constructed. Compared to ECM-2.0 anchor, simulation results of the proposed method are summarized as follows:

Setting 1:

RA: -0.19% (Y), -0.18% (U), -0.14% (V), 99% (EncT), 107% (DecT)

LB: -0.15% (Y), -0.15% (U), -0.36% (V), 102% (EncT), 105% (DecT)

Setting 2:

RA: -0.xx% (Y), -0.xx% (U), -0.xx% (V), 100% (EncT), 10x% (DecT)

LB: -0.xx% (Y), -0.xx% (U), -0.xx% (V), 100% (EncT), 10x% (DecT)

Initial results in setting 2 indicates slightly higher gains for RA, similar for LBInvestigate in EE

[JVET-X0190](https://jvet-experts.org/doc_end_user/current_document.php?id=11200) Cross-check of JVET-X0133 on MV candidate type-based ARMC [P. Onno (Canon)] [late]

[JVET-X0134](https://jvet-experts.org/doc_end_user/current_document.php?id=11127) EE2-related: On the number of TM merge candidates [Y.-J. Chang, V. Seregin, M. Karczewicz (Qualcomm)]

In this contribution, the number of TM merge candidates is changed from 4 to 7. Compared to ECM-2.0 anchor, simulation results are summarized as follows:

RA: -0.11% (Y), -0.17% (U), -0.09% (V), 103% (EncT), 100% (DecT)

LB: -0.14% (Y), -0.32% (U), -0.27% (V), 102% (EncT), 101% (DecT)

The increased number of TM operations is increasing the encoding time, because all merge candidates need to be checked. This makes the tradeoff suboptimum. Should be studied if that can be decreased.Only first 5 candidates are re-ordered.

Investigate in EE (also interactions with the other tools that are put in EE in the area of MV coding).

[JVET-X0170](https://jvet-experts.org/doc_end_user/current_document.php?id=11178) Crosscheck of JVET-X0134 (Non-EE2: On the number of TM merge candidates) [R.-L. Liao (Alibaba)] [late]

[JVET-X0141](https://jvet-experts.org/doc_end_user/current_document.php?id=11134) EE2-3.1-related: CIIP with template matching [Z. Deng, K. Zhang, L. Zhang (Bytedance), X. Li, R.-L. Liao, J. Chen, Y. Ye (Alibaba), Y.-J. Chang, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes to refine the motion vector of CIIP inter part with a template matching based method. The proposed method was implemented on top of EE2-3.1b, in which the intra-prediction mode of CIIP is implicitly derived by TIMD. Simulation results based on ECM-2.0 are reported as below (tool-off):

RA: -0.11%, -0.11%, -0.00%, 101%, 100%; LB: -0.22%, -0.27%, -0.28%, 101%, 100%.

For RA, performance improves from 0.07% -> 0.11% relative to EE2-3.1b. For LB, no change.

Replaces planar by TIMD candidate (put in merge list of CIIP). TIMD candidate is already used in normal inter prediction.

Decision: Adopt JVET-X0141

[JVET-X0189](https://jvet-experts.org/doc_end_user/current_document.php?id=11199) Cross-check of JVET-X0141 on CIP with template matching [P. Onno (Canon)] [late]

[JVET-X0145](https://jvet-experts.org/doc_end_user/current_document.php?id=11138) EE2-related: Template matching CIIP on top of EE2-3.1 [Y.-J. Chang, V. Seregin, M. Karczewicz (Qualcomm)] [late]

In this contribution, a template matching CIIP (TM-CIIP) is proposed to refine the MV of inter part of CIIP. Compared to ECM-2.0 anchor, simulation results of the proposed method combined with EE2-3.1 are summarized as follows:

(1) TM-CIIP combined with EE2-3.1a

RA: -0.xx% (Y), -0.xx% (U), -0.xx% (V), 100% (EncT), 10x% (DecT)

(2) TM-CIIP combined with EE2-3.1b

RA: -0.xx% (Y), -0.xx% (U), -0.xx% (V), 100% (EncT), 10x% (DecT)

Basically included in X0141 (but different signaling in that contribution). Results not complete yet, and the difference in signaling is minor according to proponents.

[JVET-X0147](https://jvet-experts.org/doc_end_user/current_document.php?id=11140) EE2-related: intra mode derivation based on TIMD for GPM inter/intra [H. Jang, S. Kim, J. Lim (LGE), Y. Kidani, H. Kato, K. Kawamura (KDDI)] [late]

This contribution proposes TIMD based intra mode derivation and its application without signaling to indicate which intra mode is used on geometry intra-region for the geometry partitioning mode with inter and intra(GPM inter/intra) which was proposed by JVET-W0110[1]. It also restricts that intra with intra combination is applied. With the proposed method, GPM inter/intra do not need pre-defined IPM candidates which is adaptively selected along with GPM angle. Intra mode is derived by reusing conventional TIMD method. The suggested GPM with inter/intra method show further against EE2-3.2 a/b test with similar encoding complexity.

Simulation result of the proposed method over ECM-2.0(Tool-off test) is summarized as follow.

RA : -0.06%, -0.15%, -0.11% / EncT : 101%, DecT : 100%

LDB : -0.13%, -0.54%, -0.59% / EncT : 102%, DecT : 100%

LDP : -0.48%, -1.64%, -1.19% / EncT : 106%, DecT : 99%

For an additional information, GPM storage modification which is introduced in JVET-X0166 option5 also implemented and simulated for this propose method. Simulation result are summarized as follow:

RA : TBD / TBD/ TBD

LDB : TBD/ TBD / TBD

LDP : -0.54%, -1.60%, -1.35% / EncT : 106%, DecT : 98%

In v1, only intermediate results are attached in the simulation results section.

In v2, the simulation results are added.

In v3, the partial supplemental result are attached

In v4, presentation deck are included and simulation result are fulfilled

Included in X0166, no separate review necessary.

[JVET-X0206](https://jvet-experts.org/doc_end_user/current_document.php?id=11216) Crosscheck of JVET-X0147 (EE2-related: intra mode derivation based on TIMD for GPM inter/intra) [W. Lim (ETRI)] [late]

[JVET-X0166](https://jvet-experts.org/doc_end_user/current_document.php?id=11174) EE2-related: Combination of JVET-X0078 (Test 7/8), JVET-X0147 (Proposal-2), and GPM direct motion storage [Y. Kidani, H. Kato, K. Kawamura (KDDI), H. Jang, S. Kim, J. Lim (LGE), Z. Deng, K. Zhang, L. Zhang (Bytedance)] [late]

Presentation deck to be provided. The presentation given includes substantially more results than the currently uploaded word file.

Was presented in session 8 Thu 7 Oct. 1700 UTC

This contribution proposes the combination test results of Combination of JVET-X0078 (Test 7/8), JVET-X0147 (Proposal-2), and a direct motion vector and intra mode storage method on geometry partition mode (GPM) with intra and inter prediction to further enhance the coding performance beyond VVC.

Simulation results over all classes of the best coding performance and the best trade-off conditions over ECM-2.0 are shown as follows:

* The best coding performance condition over ECM-2.0: (To be updated)
  + RA: BD-rate YUV: %, %, %; EncT: %; DecT: %
  + LB: BD-rate YUV: %, %, %; EncT: %; DecT: %
  + LP: BD-rate YUV: %, %, %; EncT: %; DecT: %
* The best trade-off condition over ECM-2.0: (To be updated)
  + RA: BD-rate YUV: %, %, %; EncT: %; DecT: %
  + LB: BD-rate YUV: %, %, %; EncT: %; DecT: %
  + LP: BD-rate YUV: %, %, %; EncT: %; DecT: %

It is commented that introducing TM might not be desirable unless justified by substantial compression gain.

It is commented that a combination intra/intra would require duplication of intra logic at block level

Performance is significantly higher in LP than in RA and LB. It is however commented that adding complicated modes to LP might not be desirable, as this is mainly foreseen for low-latency (which imposes real-time and somewhat low-comlexity) encoding.

Was reviewed again during session 15 after results were complete.

Simulation results over all classes of the best coding performance and the best trade-off conditions over ECM-2.0 are shown as follows:

* The best coding performance condition over ECM-2.0:
  + RA: BD-rate YUV: -0.17%, -0.42%, -0.36%; EncT: 102%; DecT: 100%
  + LB: BD-rate YUV: -0.35%, -0.96%, -0.98%; EncT: 102%; DecT: 100%
  + LP: BD-rate YUV: -0.77%, -1.86%, -1.72%; EncT: 122%; DecT: 100%
* The best trade-off condition over ECM-2.0:
  + RA: BD-rate YUV: -0.13%, -0.34%, -0.30%; EncT: 102%; DecT: 100%
  + LB: BD-rate YUV: -0.26%, -0.71%, -0.65%; EncT: 101%; DecT: 100%
  + LP: BD-rate YUV: -0.69%, -1.84%, -1.61%; EncT: 107%; DecT: 100%

Relative to EE, restricting intra/intra is favorable (called “solution 2” in the contribution). The additional TIMD/DIMD candidates (called “solution 3” in the contribution) add complexity over the EE. This contributes approx. 0.05% of the overall gain. Direct MV&IPM storage (called “solution 4” in the contribution) is favorable. “Solution 1” adds TM to the GPM-MMVD (which is different from “normal” MMVD.

In LP, GEO is only enabled for intra/inter combination.

Results on adding only “solutions 2 and 4” on top of EE are currently not available. No cross-check available.

Further investigation in EE about the benefits of the different elements of the proposal

### ECM modifications beyond EE2 (24)

Contributions in this area were discussed in session 16 at 0600–0710 UTC and session 18 at 1300-1405 on Tuesday 12 Oct. 2021 (chaired by JRO), and session 19 at 0620-0920 on Wednesday 13 Oct. 2021 (chaired by JRO).

[JVET-X0058](https://jvet-experts.org/doc_end_user/current_document.php?id=11051) AHG12: Bilateral matching SMVD mode [D. Kim, H. Kwon, J. Kim, W. Lim, S. Y. Jeong (ETRI)]

In this contribution, bilateral matching Symmetric MVD mode (BM-SMVD) is proposed as a sub-mode of SMVD. In the proposed method, only MVP indices for list0 and list1 are signaled and MVs are refined by minimizing bilateral matching error. Performance of the proposed method on top of VTM-13.0 and ECM-2.0 are reported as follows.

* Test 1. BM-SMVD on VTM-13.0  
   : -0.08% (Y), -0.03% (U), - 0.08% (V), 101% (EncT), 100% (DecT)
* Test 2. BM-SMVD on ECM-2.0 (TM-off)  
   : -0.xx% (Y), -0.xx% (U), - 0.xx% (V), xxx% (EncT), xxx% (DecT)

Results of test 2 only shown for classes B,C,D; around 0.2% gain. However, a multi-stage version of DMVR is used which also increases encoding time more (3%). Another change is about using the refined MV for spatial and temporal MV propagation.

With TM enabled, gain might be lower

Would the gain be retained after the new ECM adoptions? Not known.

Further study (not EE yet).

[JVET-X0179](https://jvet-experts.org/doc_end_user/current_document.php?id=11189) Crosscheck of JVET-X0058 (AHG12: Bilateral matching SMVD mode) [[J. Nam (LGE)](mailto:junghak.nam@lge.com)] [late]

[JVET-X0085](https://jvet-experts.org/doc_end_user/current_document.php?id=11078) Non-EE2: Template Matching-based Reordering for Extended MMVD Design [M. Salehifar, Y. He, K. Zhang, N. Zhang, L. Zhang (Bytedance)]

A template matching based reordering method for extended MMVD is proposed, in which the MMVD refinement positions are reordered based on template costs. In the extended MMVD, additional positions are added along k×π/8 diagonal angles, meanwhile half of distance offsets in the current MMVD design are removed in test 1 (for test 2, all the 8 distance offsets are kept). After the reordering process, the top 1/8 positions with the smallest SAD costs are kept as available positions, consequently for MMVD index coding.

For test 3 similar concept is extended to Affine MMVD.

Simulation results are reported as (tool-off tests, ECM-2.0 as the anchor):

Test 1: RA: -0.12%, -0.17%, -0.15%, 101%, 102%; LB: -0.18%, - 0.25%, -0.17%, 100%, 105%

Test 2: RA: -0.17%, -0.27%, -0.28%, 10x%, 10x%; LB: -0.xx%, - 0.xx%, -0.xx%, xxx%, xxx%

Test 3: RA: -0.xx%, - 0.xx%, -0.xx%, xxx%, xxx%; LB: -0.xx%, - 0.xx%, -0.xx%, xxx%, xxx%

Test 2/3 results not complete.

It is mentioned that the runtime could be reduced by using bilateral filters (results not complete yet).

Was it also tested with TM off? No

Would the gain be retained after the new ECM adoptions? Not known.

Further study in EE

[JVET-X0208](https://jvet-experts.org/doc_end_user/current_document.php?id=11218) Crosscheck for X0085 (Non-EE2: Template Matching-based Reordering for Extended MMVD Design) [J. Zhang (??)]

[JVET-X0087](https://jvet-experts.org/doc_end_user/current_document.php?id=11080) Non-EE2: Template Matching Based Merge Candidate List Construction (TM-MCLC) [L. Zhao, K. Zhang, N. Zhang, L. Zhang (Bytedance)]

In this contribution, a method of template matching based merge candidate list construction (TM-MCLC) is proposed. Instead of constructing the merge candidate list based on a predefined traversing order, TM-MCLC builds the merge candidate list by selecting more efficient candidates based on template matching costs. On top of ECM-2.0, simulation results of the proposed method are reported as below:

Tool-off test (ECM-2.0 as the anchor): RA: -0.27%, 100%, 106%; LB: -0.32%, 101%, 110%.

Additional 32 spatial and 12 temporal positions for constructing the candidate list.

Integer position is used for TM.

Not cross-checked.

Would the gain be retained after the new ECM adoptions?

Significant decoding time increase, but proponents believe this could be reduced.

How much gain comes by introducing TM, and how much from additional candidates? Without TM, the additional candidates don’t show benefit.

Could the same be applied with less additional candidates? The re-ordering could be highly complex, when all are available (some method of pruning is used).

Further study in EE.

[JVET-X0088](https://jvet-experts.org/doc_end_user/current_document.php?id=11081) Non-EE2: History-based Affine Model Inheritance [K. Zhang, L. Zhang, Z. Deng, N. Zhang, Y. Wang (Bytedance)]

This contribution presents a method of history-parameter-based affine model inheritance (HAMI). With HAMI, affine parameters of previously affine-coded block have been stored in a History-Parameter Table (HPT). New affine merge and AMVP candidates as well as new regular merge candidates can be constructed with affine parameters fetched from HPT and base MVs fetched from neighbouring blocks, and put into the sub-block-based merge candidate list, the affine AMVP candidate list and the regular merge candidate list, respectively. On top of ECM-2.0, simulation results of the proposed method are reported as below:

RA: -0.15%, 101%, 100%; LB: -0.12%, 101%, 100%.

Not cross-checked.

Subblock merge list increased from 5 to 9

Would the gain be larger with TM off? Not known, likely.

Is pruning applied? Yes, but not described in contribution.

Further study in EE.

[JVET-X0089](https://jvet-experts.org/doc_end_user/current_document.php?id=11082) Non-EE2: Modifications of IBC Merge/AMVP List Construction [N. Zhang, K. Zhang, L. Zhang, J. Xu (Bytedance)]

In this contribution, the IBC merge/AMVP list construction process is modified with three aspects :1) Only if a block vector (BV) candidate is valid, it can be inserted into the IBC merge/AMVP candidate list; 2) More BV candidates, including those from above-right, bottom-left, and above-left spatial neighboring blocks and one pairwise average candidate, can be added into the IBC merge/AMVP list; 3) ARMC-TM is extended to IBC merge list.

On top of the ECM2.0, simulation results of the proposed method are reported as below:

AI: Class F -0.43%, 101%, 98%; Class TGM -1.10%, 101%, 94%

RA: Class F -0.18% , 100%, 100%; Class TGM -0.89%, 100%, 98%

LB: Class F -0.18%, 99%, 99%; Class TGM -0.77%, 100%, 101%

Not cross-checked.

ARMC-TM is extended to be used in IBC merge list construction.

Number of candidates not increased.

Extends existing ARMC-TM to IBC.

Study in EE.

[JVET-X0090](https://jvet-experts.org/doc_end_user/current_document.php?id=11083) Non-EE2: On combination of CIIP, OBMC and LMCS [R.-L. Liao, X. Li, J. Chen, Y. Ye (Alibaba)]

In ECM-2.0, when CIIP, OBMC and LMCS are all applied to a CU, the samples generated by CIIP mode in the mapped domain are weighed with the samples generated by OBMC mode in the original domain. In this contribution, it is proposed to weigh the samples generated by CIIP and OBMC modes in the same domain. When performing weighting in the mapped domain, it is reported that on top of ECM-2.0, the overall coding performance impact for {Y, U, V, EncT, DecT} is {-0.01%, -0.02%, 0.01%, 100%, 100%} for RA and {-0.05%, 0.05%, -0.18%, 101%, 101%} for LDB. When performing weighting in the original domain, it is reported that on top of ECM-2.0, the overall coding performance impact for {Y, U, V, EncT, DecT} is {-0.01%, 0.09%, 0.05%, 100%, 100%} for RA and {-0.04%, 0.06%, -0.13%} for LDB. When performing OBMC before performing weighting with intra predicted samples, it is reported that on top of ECM-2.0, the overall coding performance impact for {Y, U, V, EncT, DecT} is {-0.01%, -0.02%, 0.02%, 100%, 100%} for RA and {-0.03%, -0.14%, -0.18%, 101%, 100%} for LDB.

Several experts supported the proposal, pointing out that this is removing an inconsistency in ECM. Solution 3 (weighting inter predicted samples together in a first step before weigthing intra) is asserted to be the mst appropriate solution, as it avoids an additional mapping step.

It is also pointed out that in case of HDR the effect would likely be more noticeable in terms of compression gain. However, HDR is currently not in CTC of EE2

Decision: Adopt JVET-X0090 solution 3

[JVET-X0192](https://jvet-experts.org/doc_end_user/current_document.php?id=11202) Crosscheck of JVET-X0090: Non-EE2: On combination of CIIP, OBMC and LMCS [X. Xiu (Kwai)] [late] [miss]

[JVET-X0197](https://jvet-experts.org/doc_end_user/current_document.php?id=11207) Cross-check of JVET-X0090 (Non-EE2: On combination of CIIP, OBMC and LMCS) method 1 [H. Huang (Qualcomm)] [late]

[JVET-X0091](https://jvet-experts.org/doc_end_user/current_document.php?id=11084) Non-EE2: On TMVP improvement [R.-L. Liao, J. Chen, Y. Ye, X. Li (Alibaba)]

This contribution proposes to modify the derivation process of TMVP merge candidate. In the method 1, the reference picture of the TMVP merge candidate is selected from any one of reference pictures in the reference picture list instead of being fixed to the picture with index 0. The selected reference picture is the one whose scaling factor is closest to 1. In the method 2, when constructing the merge candidate list of non-subblock merge modes, the TMVP is conditionally added to the list according to the motion information of neighboring spatial blocks. Moreover, after the list is constructed, the TMVP is refined using the template matching method. The template is used to decide the scaling factor and inter prediction direction of the TMVP. It is reported that on top of ECM-2.0, the overall coding performance impact for {Y, U, V, EncT, DecT} is:

Method 1: RA{-0.04%, -0.01%, 0.00%, 100%, 101%}, LDB{-0.28%, -0.64%, -0.56%, 102%, 102%}

Method 1 + 2: RA {-0.15%, -0.12%, -0.18%, 99%, 103%}, LDB{-0.34%, -0.44%, -0.50%, 102%, 103%}

It is pointed out that in method 2 information has to be stored whether the neighboring blocks use TMVP (one flag per 4x4 block in a CTU row).

6 additional condition checks are necessary in method.

Study in EE

[JVET-X0181](https://jvet-experts.org/doc_end_user/current_document.php?id=11191) Crosscheck of JVET-X0091 (Non-EE2: On TMVP improvement) [[Z. Deng (Bytedance)](mailto:zhipin.deng@bytedance.com)] [late]

[JVET-X0099](https://jvet-experts.org/doc_end_user/current_document.php?id=11092) Non-EE2: Extension of TIMD to intra chroma coding [Y. Wang, K. Zhang, L. Zhang, H. Liu (Bytedance)]

This contribution presents to extend template-based intra mode derivation to intra chroma coding, known as TIMD-Chroma. Compared to ECM-2.0, simulation results of the proposed method are reported as below:

AI: {-0.09%, -0.16%, -0.14%; 101%, 104%};

RA: {-0.05%, -0.13%, -0.04%; 101%, 100%}.

No good tradeoff complexity vs. compression benefit. No action at this point.

[JVET-X0105](https://jvet-experts.org/doc_end_user/current_document.php?id=11098) AHG12: Edge Classifier for Cross-component Sample Adaptive Offset (CCSAO) [A. M. Kotra, N. Hu, V. Seregin, M. Karczewicz (Qualcomm)] [late]

The cross-component sample adaptive offset (CCSAO) in ECM-2.0 [2] only uses the band classifier to classify a given sample. This contribution extends the CCSAO design by adding the edge-based classifier.

The objective results, over ECM-2.0 Anchor for CTC configuration are as follows:

When the edge-based classifier is applied only to “Chroma” components:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Config. | Y | U | V | EncT | DecT |
| AI | 0.06% | -1.14% | -0.88% | 101% | 102% |
| RA | 0.04% | -0.88% | -0.67% | 99% | 99% |
| LDB | -0.02% | -0.93% | -0.53% | 98% | 98% |

When the edge-based classifier is applied only to both “Luma” and “Chroma” components:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Config. | Y | U | V | EncT | DecT |
| AI | ?? | ?? | ?? | ?? | ?? |
| RA | ?? | ?? | ?? | ?? | ?? |
| LDB | ?? | ?? | ?? | ?? | ?? |

First method (only chroma) does not show sufficient benefit, as the additional signalling makes luma performance worse, not justified by the still relative low chroma gain.

Second method seems to provide some more benefit, but results are incomplete still.

X0152 proposes a similar concept as the second method.

Investigate in EE (see further notes under X0152).

[JVET-X0191](https://jvet-experts.org/doc_end_user/current_document.php?id=11201) Cross-check of AHG12: Edge Classifier for Cross-component Sample Adaptive Offset (CCSAO) [[J. Chen](mailto:jiechen.cj@alibaba-inc.com), X. Li (Alibaba)] [late]

[JVET-X0119](https://jvet-experts.org/doc_end_user/current_document.php?id=11112) Non-EE2: On pairwise merge candidate [G. Laroche, P. Onno, R. Bellessort (Canon)]

This contribution presents a modification of the merge candidates derivation. It is proposed to derive the pairwise candidate during the reordering process of the merge candidates list (AMCR-TM, JVET-W0090). In addition, each merge candidate, in the non-reordered subgroup, is replaced by a pairwise between the first candidate and this candidate.

On top of the ECM-2.0, simulation results of the proposed methods are reported as follows:

RA: -0.13%, 101%, 101%

LB: -0.10%, 99%, 100%

Study in EE.

[JVET-X0153](https://jvet-experts.org/doc_end_user/current_document.php?id=11160) Crosscheck of JVET-X0119 (Non-EE2: On pairwise merge candidate) [N. Zhang (Bytedance)] [late]

[JVET-X0120](https://jvet-experts.org/doc_end_user/current_document.php?id=11113) AHG12: On sign prediction [M. G. Sarwer, Y. Ye, J. Chen, R.-L. Liao (Alibaba)]

This contribution proposes a sorting-based approach to identify the signs to be predicted. In the proposed method, after decoding absolute values of the levels, the transform block (TB) is sorted based on the qIdx values and the first maxNumPredSigns signs of the sorted TB are predicted and rest of the signs are EP coded. This contribution also extended the sign prediction area of a TB from 4x4 block to maximum 32x32 block.

Following results are reported:

* AI: -0.11% (Y), -0.10%(U), -0.08%(V), 105%(EncTime), 105% (DecTime)
* RA: -0.14% (Y), -0.05 %(U), -0.05 %(V), 102%(EncTime), 107% (DecTime)
* LB: -0.16% (Y), 0.16 %(U), 0.09 %(V), 102%(EncTime), 106% (DecTime)

Current sign prediction in ECM gives approx. 0.8%, increases encoder run time by approx. 9% relative to VTM. The new proposal has a worse tradeoff of complexity vs. compression benefit.

Study in EE along with X0150; also in terms of the complexity tradeoff

[JVET-X0195](https://jvet-experts.org/doc_end_user/current_document.php?id=11205) Crosscheck of JVET-X0120 (AHG12: On sign prediction) [B. Ray (Qualcomm)] [late]

[JVET-X0122](https://jvet-experts.org/doc_end_user/current_document.php?id=11115) Non-EE2: Unification of negative modes processing in TIMD [A. Filippov, V. Rufitskiy, K. Goswami, D. Ruiz Coll, Y. Y. Lee, T. M. Bae, E. Dinan (Ofinno)]

This contribution presents a modification to the inverse angle processing that is performed when TIMD mode is selected. Reportedly, the proposed modification provides the following overall results: ‑0.09% (Y) /‑0.13% (U) /-0.14% (V) and x.xx% (Y) / x.xx% (U) / x.xx% (V) in AI and RA configurations, respectively. The coding performance improvement is achieved at no complexity and runtime increase for both encoder and decoder.

This is rather fixing a wrong factor in mode derivation. This issue was already submitted as a bug fix merge request #16 by Alibaba, who confirm the results, and also report that gain in RA/LB luma is 0.05%.

Decision(SW/BF): Adopt JVET-X0122, and perform merge request #16

[JVET-X0165](https://jvet-experts.org/doc_end_user/current_document.php?id=11172) Crosscheck of JVET-X0122 (Non-EE2: Unification of negative modes processing in TIMD) [K. Kondo (Sony)] [late]

[JVET-X0124](https://jvet-experts.org/doc_end_user/current_document.php?id=11117) AHG12: On signalling of intra template matching [Z. Xie, Y. Yu, H. Yu, L. Xu, F. Wang, D. Wang (OPPO)]

In this contribution, harmonization of CU-level flags of decoder-side intra mode derivation (DIMD) and intra template matching prediction (TMP) is proposed, which aims to remove redundant signalling bits in the bitstream. In the ECM-2.0 software, one dedicated flag at CU level is signalled to indicate if TMP is executed. It is proposed to skip TMP signalling flag when DIMD is in use at the current CU. This modification has been implemented in the ECM-2.0 software and the test results of class F showed, on average, coding gain of {-0.05%, -0.17%, -0.08%} for AI, {-0.03%, 0.02%, -0.37%} for RA, and {-0.18%, -0.36%, -0.81%} for LDB.

This is asserted as a straightforward syntax cleanup.

Decision: Adopt JVET-X0124.

[JVET-X0164](https://jvet-experts.org/doc_end_user/current_document.php?id=11171) Crosscheck of JVET-X0124 (AHG12: On signalling of intra template matching) [X. Li (Alibaba)] [late]

[JVET-X0131](https://jvet-experts.org/doc_end_user/current_document.php?id=11124) Non-EE2: Low-Complexity Improvements of Intra Coding for Screen Content [D. Ruiz Coll, T. M. Bae, A. Filippov, V. Rufitskiy, Y. Y. Lee, K. Goswami (Ofinno)] [late]

This contribution proposes (1) a sign prediction of the IBC block vector difference method and (2) an MRLP modification for screen content coding. On top of ECM-2.0, the test results for the 1st technique reportedly show the class F compression performance improvements of -0.09%, -0.08%, -0.08% (Y/U/V) and -0.21%, -0.20%, -0.20% (Y/U/V) for class TGM. On top of ECM-2.0, the test results for the 2nd technique reportedly show the class F compression performance improvements of -0.06%, -0.04%, -0.14% (Y/U/V). The joint test results for both techniques reportedly show the class F compression performance improvements ‑0.15%, -0.13%, -0.22 % (Y/U/V) and -0.24%, -0.23%, -0.20%, (Y/U/V) for class TGM. Changes of encoding and decoding runtime for all the tests are reported to be negligible.

The gain is comparably low (in terms of screen content where tools mostly tend to show more gain), and additional processing is necessary. The sign prediction also introduces a parsing dependency on IBC displacemet vector reconstruction. It is also asked how the MRL modification would work at CTU boundary. There is a boundary constraint implemented.

No action on this proposal.

[JVET-X0135](https://jvet-experts.org/doc_end_user/current_document.php?id=11128) Non-EE2: Adaptive intra MTS [B. Ray, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes to use variable number of MTS candidates for intra coded blocks depending on the position of last significant coefficient. Simulation results on top of ECM-2.0 are:

AI (Y/U/V): -0.10%/-0.06%/-0.07%, EncT 100%, DecT 99%

RA (Y/U/V): -0.03%/-0.02%/0.00%, EncT 95%, DecT 99%

LDB (Y/U/V): -0.10%/0.13%/-0.05%, EncT 101%, DecT 98%

Current ECM uses only DCT2 when only DC coefficient is non-zero, otherwise all options. The proposal is about using less MTS candidates when the number of other (AC) non-zero coefficients depending on last coeff. Position (1, 4 or 6 are used). Reduction in run time for RA is achieved by some additional encoder optimization possible by this.

Have other configurations of variable number been tested? Yes, some other options, not extensively, was the best tradeoff found.

Would the same encoder optimization be possible in ECM anchor? Yes, but would give loss without the change.

Is it simple to extract the last coeff. position? It is parsed independent and ahead of the MTS signaling.

Transform kernels are the same as currently in MTS of ECM.

Study in EE, also investigate encoder-only option, dependency on the thresholds of number of candidates, and possibly dependency of threshold on QP (e.g. signalling it).

[JVET-X0194](https://jvet-experts.org/doc_end_user/current_document.php?id=11204) Crosscheck of JVET-X0135 (Non-EE2: Adaptive intra MTS) [M.G. Sarwer (Alibaba)] [late] [miss]

[JVET-X0139](https://jvet-experts.org/doc_end_user/current_document.php?id=11132) AHG12: removing a discontinuity in the discrete angle comparison in DIMD [T. Dumas, P. Bordes, F. Galpin, F. Le Léannec (InterDigital)]

This contribution proposes to remove a discontinuity in the discrete angle comparison in Decoder Side Intra Mode Derivation (DIMD). Indeed, at a given decoded reference around the current luminance CB, this discontinuity can make the inferred intra prediction mode inconsistent with the local gradient.

Mean BD-rate reductions of -0.01% -0.02% -0.01% and -0.01% 0.05% 0.08% in AI and RA configurations respectively are reported when comparing ECM-2.0 including the removed discontinuity in the discrete angle comparison in DIMD with respect to ECM-2.0.

JVET-X0156 is targeting the same issue by a different software bug fix, and leads to identical results (see further notes there).[JVET-X0142](https://jvet-experts.org/doc_end_user/current_document.php?id=11135) Non-EE2: Extended MRL candidate list [K. Cao, Y.-J. Chang, B. Ray, V. Seregin, M. Karczewicz (Qualcomm)]

This contribution proposes to extend the number of multiple reference lines (MRL) for intra prediction. This method was implemented on top ECM-2.0 software, it reportedly provides Y, U, V-BD rate reduction with encoder and decoder runtime as follows:

AI: {-0.10%, -0.05%, -0.01%} runtime: {101%, 98%}

RA: {-0.05%, -0.01%, 0.06%} runtime: {101%, 100%}

The farthest line/column used is 12 away from the boundary

At CTU boundary, MRL is not used from row above (only one line buffer). Also, when the block is close to the horizontal boundary, less candidates are allowed (and signaling is modified).

Proponents report that no encoder optimization was used to decide about the additional lines and columns.

Investigate in EE. Also test with different number of candidates, i.e. if really all five are needed, and how often the farther away candidates are really used.

[JVET-X0172](https://jvet-experts.org/doc_end_user/current_document.php?id=11180) Crosscheck of JVET-X0142 (Non-EE2: Extended MRL candidate list) [[H.-J. Jhu (Kwai)](mailto:jhuhong-jheng@kwai.com)] [late]

[JVET-X0146](https://jvet-experts.org/doc_end_user/current_document.php?id=11139) Non-EE2: Decoder side motion derivation using sample's spatial correlation [H. Huang, V. Seregin, C.-C. Chen, M. Karczewicz (Qualcomm)]

This contribution presents a decoder side motion derivation method that minimizes the boundary discontinuity across the current prediction block and its neighboring blocks. It reports -0.11%, -0.10 %, -0.12 % BD rate for luma and chroma components, respectively, with 103% encoder and 99% decoder run time in random access configurations of common test condition.

A new mode of template matching is introduced, and it is signaled by a flag at block level if it is used

Complexity of the old and new method may be similar, but it is a substantially different way of processing (including the current block from current picture).

Is the new method still performing well at high QP?

Not applied to GPM+TM.

Tradeoff complexity vs. compression benefit not favorable.

Study in EE[JVET-X0205](https://jvet-experts.org/doc_end_user/current_document.php?id=11215) Crosscheck of JVET-X0146 (Non-EE2: Decoder side motion derivation using sample's spatial correlation) [H.-J. Jhu (Kwai)] [late]

[JVET-X0148](https://jvet-experts.org/doc_end_user/current_document.php?id=11141) AHG12: On the PDPC handling in DIMD and TIMD [H.-J. Jhu, X. Xiu, Y.-W. Chen, W. Chen, C.-W. Kuo, N. Yan, X. Wang (Kwai)] [late]

In ECM-2.0, two decoder-side intra coding techniques, namely, decoder-side intra mode derivation (DIMD) and template-based intra mode derivation (TIMD), are applied to reduce the overhead of intra mode signaling. In either DIMD or TIMD, prediction samples generated from multiple intra modes could be combined together to from the final prediction samples of one intra CU where position dependent intra prediction combination (PDPC) may be applied.

It is reported that in the current design, the handling of the PDPC in the DIMD and the TIMD is not unified. Specifically, in the TIMD, the PDPC can be performed either before or after the fusion process according to the derived intra mode. On the other hand, in the DIMD, the PDPC is always performed before the fusion. It is asserted that such non-unified design not only complicates software/hardware design logics but also causes latency issue in the decoding pipeline of intra blocks.

In this proposal, one unification scheme is proposed to resolve the latency issue by performing all the PDPC operations before the fusion process of the DIMD and the TIMD. The proposed method was implemented on top of ECM-2.0, and it reports the average Y BD-rate changes of -0.02% and 0.00% for AI and RA configurations, respectively, without encoding/decoding impacts.

TIMD change is asserted as a clear simplification

As in current ECM, three intra threads are needed in DIMD, however an additional PDPC is needed in the proposal after planar mode.

A better solution would be to keep DIMD as is, and modify TIMD as suggested.

Decision: Adopt JVET-X0148 (only the TIMD pipeline modification)

[JVET-X0160](https://jvet-experts.org/doc_end_user/current_document.php?id=11167) AHG12: Crosscheck of JVET-X0148 (On the PDPC handling in DIMD and TIMD) [K. Cao (Qualcomm)] [late]

[JVET-X0149](https://jvet-experts.org/doc_end_user/current_document.php?id=11142) AHG12: Removal of floating operations in DIMD and TIMD [X. Xiu, J.-H. Jhu, W. Chen, C.-W. Kuo, N. Yan, Y.-W. Chen, X. Wang (Kwai)] [late]

In ECM-2.0, DIMD and TIMD are two decoder-side coding technologies which reduces the intra coding overhead by deriving the intra prediction mode of intra CU at decoder. In the existing DIMD/TIMD design, multiple floating-point operations are involved during the derivation of the optimal intra modes, which are expensive for hardware and software implementations. In this contribution, it is proposed to replace all the floating-point operations in the DMD/TIMD by the existing lookup table used by CCLM, such that only integer additions/multiplications are needed to calculate the DIMD/TIMD parameters. Simulation results reportedly show the BD-rate changes of 0.00% and -0.01% for AI and RA configurations, respectively, without noticeable enc/dec impact.

Also divisions are removed.

Confirmed by cross-check

This is asserted as a desirable and straightforward simplification

Decision: Adopt JVET-X0149.

[JVET-X0178](https://jvet-experts.org/doc_end_user/current_document.php?id=11187) Crosscheck of JVET-X0149 (AHG12: Removal of floating operations in DMD and TIMD) [[Y. Wang (Bytedance)](mailto:wangyang.cs@bytedance.com)] [late]

[JVET-X0150](https://jvet-experts.org/doc_end_user/current_document.php?id=11143) AHG12: Enhanced sign prediction [X. Xiu, Y.-W. Chen, N. Yan, C.-W. Kuo, H.-J. Jhu, W. Chen, X. Wang (Kwai)] [late]

In this contribution, two enhancements are proposed to further improve the performance of the sign prediction tool in ECM-2.0, including 1) extension of the sign prediction to LFNST mode; 2) adaptive selection of transform coefficient positions for the sign prediction by maximizing their impacts on reconstructed border samples. Compared to the ECM-2.0 anchors, simulation results reportedly show that the proposed enhancements provide average {Y, U, V}BD-rate changes of {-0.31%, -0.31%, -0.41%} and {-0.23%, -0.22%, -0.11%} and {-0.16%, -0.17%, -0.27%} for AI and RA and LDB configurations, respectively. The corresponding enc/dec run-time are 107%/107% for AI and 101%/102% for RA and 100%/101% for LDB.

Encoding/decoding time increase mainly due to LFNST (therefore higher in AI), might have no impact on hardware complexity.

How much gain comes from each of the parts 1 and 2? Roughly half, slightly less for LFNST.

Adaptive selection seems to target a similar benefit as re-ordering in X0120.

Study in EE along with X0120, also study possible combination e.g. extending the region to 32x32 in X0150

[JVET-X0198](https://jvet-experts.org/doc_end_user/current_document.php?id=11208) Cross Check of JVET-X0150 (AHG12: Enhanced sign prediction) [L.-F. Chen (Tencent)] [late] [miss]

[JVET-X0151](https://jvet-experts.org/doc_end_user/current_document.php?id=11144) AHG12: Non-adjacent spatial neighbors for affine merge mode [W. Chen, X. Xiu, Y.-W. Chen, H.-J. Jhu, C.-W. Kup, N. Yan, X. Wang (Kwai)] [late]

This contribution proposes to use non-adjacent spatial neighbors for deriving candidates in affine merge mode. In the proposed scheme, the motion information of non-adjacent spatial neighbors, which is obtained in a fixed pattern, is utilized to derive additional inherited and constructed affine merge candidates. The derived candidates are inserted into the existing candidate list of affine merge mode. The proposed method was implemented on top of ECM-2.0, and it reports the average {Y, U, V} BD-rate savings of {-0.24%, -0.14%, -0.12%} and {-0.28%, -0.04%, -0.16%} for RA and LD configurations, respectively. The corresponding encoding and decoding time are 102% and 102% for RA, and 102% and 103% for LD, respectively.

Maximum number of candidates increased to 15.

How much gain from inherited candidates? Approximately half of the gain.

Is pruning applied? Yes.

Maximum number of candidates that could be derived? For inherited case, 8. For constructed candidates, could be up to 3. Maximum distance is 4 PUs.

Study in EE, also in combination with JVET-X0088 (history based affine).

[JVET-X0199](https://jvet-experts.org/doc_end_user/current_document.php?id=11209) Cross Check of JVET-X0151 (AHG12: Non-adjacent spatial neighbors for affine merge mode) [L.-F. Chen (Tencent)] [late] [miss]

[JVET-X0152](https://jvet-experts.org/doc_end_user/current_document.php?id=11145) AHG12: CCSAO classification with edge information [C.-W. Kuo, X. Xiu, Y.-W. Chen, H.-J. Jhu, W. Chen, N. Yan, X. Wang (Kwai)] [late]

In this contribution, one classifier considering both edge and band information is proposed to improve the CCSAO in ECM-2.0, which only uses band information from three component samples to classify the current component sample. Specifically, for a given luma/chroma sample, luma EO and BO are combined to classify the given sample into different categories. The original three component band classifier and the proposed classifier are switched at CTB level.

Compared to ECM-2.0 anchors, the BD-rate impact is summarized as below.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | Y | U | V | EncT | DecT |
| Combined EO/BO classifier | RA | -0.11% | -0.61% | -0.42% | 110% | 104% |
| LB | -0.31% | -0.68% | -0.23% | 112% | 105% |

Similar as X0105 method 2, but different encoder and is different in that the number of bands is reduced when edge classifier is used which is not the case in the other proposal. Switching done at CTU level

Encoding time larger than in the other proposal, the encoder is less optimized. Performance in chroma seems to be somewhat lower (though the other proposal did not have complete result yet)

Investigate in EE along with X0105 method 2, also analyse in detail the contributions of the different elements of the two proposals, and possible combinations.

[JVET-X0168](https://jvet-experts.org/doc_end_user/current_document.php?id=11176) Crosscheck of JVET-X0152: AHG12: CCSAO classification with edge information [J. Chen, X. Li (Alibaba)] [late]

[JVET-X0156](https://jvet-experts.org/doc_end_user/current_document.php?id=11163) Non-EE2: Fix for histogram of gradients derivation in DIMD mode [K. Cao, V. Seregin, M. Karczewicz (Qualcomm)] [late]

This contribution proposes a fix for deriving an index for the diagonal intra mode direction while generating histogram of gradients in DIMD mode. It was implemented on top of ECM-2.0 software, it reportedly provides Y, U, V-BD rate reduction with encoder and decoder runtime as follows:

AI: {-0.01%, -0.02%, -0.01%} runtime: {100%, 101%}

The bug fix suggested does not require a dedicated check of certain modes (as X0139 suggests), but solves it by changing the loop and conditions therein, whih seems to be more straightforward to implement.

Decision (SW/BF): Adopt JVET-X0156

V. Seregin to coordinate the setup of next EE description. Proponents of contributions selected for EE to get in contact.

# High-level syntax (HLS) proposals (6)

## AHG9: SEI message studies and proposals (4)

Contributions in this area were discussed in session 10 at 0800–0915 UTC on Friday 8 Oct. 2021 (chaired by JRO), and at 0720–0805 UTC on Tuesday 12 Oct. 2021 (chaired by JRO).

[JVET-X0092](https://jvet-experts.org/doc_end_user/current_document.php?id=11085) AHG9: Down-sample phase indication (SEI message) [J. Samuelsson, A. Tourapis, D. Podborski, K. Rapaka, D. Singer (Apple)]

This input contribution relates to video streams that have been produced by a down-sampling operation applied to a video of higher resolution. The contribution contains a proposal for indication of the phase that has been used when creating the down-sampled video. It is claimed that having access to such information would make it more feasible for receivers and players to appropriately present the decoded video. One example use-case mentioned in the contribution is an MPEG-DASH client that switches between representations of different resolution and where seamless switching is highly desirable. It is proposed that the suggested phase indication is included in a new SEI message added to Versatile supplemental enhancement information messages for coded video bitstreams (ITU-T H.274 | ISO/IEC 23002-7).

Both luma and chroma could have a phase shift in downsampling, e.g. depending on using an even-tap or odd-tap symmetric filter in case of 2x downsampling. Chroma location type would also have impact on the correct downsampling filter phase to be used for chroma, and it is assumed in the contribution that it has been done correctly.

If the wrong position is used in upsampling, a content shift may become visible in case of resolution switching, and it is the purpose of this contribution to provide a solution on this. The intent is keeping it simple for the most common case of downsampling in stream switching.

Is the syntax “co-sited/centered” sufficient only for a 2:1 downsampling? E.g., for a 3:2 downsampling, it might be necessary to indicate which pixel positions are co-sited. Further, it may be necessary to allow different positions for horizontal and vertical direction.

Semantics is not sufficient to understand what kind of processing.

Would it be necessary to indicate the original chroma position? The proposal is mainly meant for the correct luma position. It is necessary to take information about chroma type from VUI to arrive at correct chroma position again. It is asked if this would be possible with all chroma types, or if it would be necessary to signal for luma and chroma separately.

It is asserted that this SEI message is useful. Proponents are asked to further study the aspects mentioned above, and provide an improved text.

[JVET-X0096](https://jvet-experts.org/doc_end_user/current_document.php?id=11089) AHG2/AHG9: On Multiview View Position (MVP) SEI message [B. Choi, A. Hinds, X. Zhang, S. Wenger, S. Liu (Tencent)]

Item 1 of this document belongs to this category.

Two modifications on multiview view position SEI message, which was adopted into VSEI, are proposed.

Proposal-1: Extend the dimension of view positions from 1-D array to 2-D matrix.

Proposal-2: Add a constraint that a multiview view position SEI message shall not be contained in a scalable nesting SEI message.

A syntax extension is proposed which would not be backward compatible with the existing SEI message that was defined in AVC and HEVC.

It is noted that, when the SEI message was defined in JCT-3V, the purpose of 2D camera arrays was resolved with other SEI message like camera parameters, Multiview acquisition information (MAI).

This is likely not appropriate for the purpose of light fields. Further study is recommended regarding the requirements for encoding light fields with video codecs, covering also other aspects such as packing into tiles, lightfield acquisition devices and display related aspects, rather than re-defining an existing SEI message for one of the aspects (number of views). This should also be done in coordination with other MPEG WGs working in that area.

[JVET-X0101](https://jvet-experts.org/doc_end_user/current_document.php?id=11094) AHG9: On the CREI SEI message [R. Skupin, C. Bartnik, A. Wieckowski, K. Sühring, Y. Sanchez, B. Bross, T. Schierl (HHI)] [late]

At the last JVET meeting, the Constrained RASL Encoding Indication (CREI) SEI message was adopted into the VVC v2 Draft Specification that indicates that RASL pictures within a CVS are encoded with encoder-side constraints. The constraints enable bitstream resolution switching with efficient open GOP coding structures (CRA+RASL pictures) without severe artifacts in RASL pictures.

Such bitstream resolution switching involves post-encoder usage of RPR that disables wrap-around motion compensation, even if it has been used on encoder side. In order to avoid mismatches in such situations, wrap-around motion compensation needs to be disabled on encoder side. While this is possible at PPS level, the same issues than for the other constraints indicated by the CREI SEI message arise from systems perspective, e.g. packaging such a set of bitstreams would require thorough parsing to understand the situation at each switching point in the bitstreams.

This document proposes to add a constraint to the CREI SEI message requiring disabling of wrap-around motion compensation in the PPSs referred to by the RASL pictures of a CVS. Thereby, the remaining non-RASL inter-pictures may use wrap-around motion compensation and avoid an unnecessary coding efficiency penalty while post-encoder usage of RPR by RASL pictures does not lead to severe artifacts.

The problem had not been detected before, as CREI had not been tested with 360 video.

It is agreed that the problem exists for resolution switching, but would not occur for quality switching. The benefit of disabling only for resolution switching may however be marginal.

Decision: Adopt JVET-X0101 (for VSEI v2 draft)

[JVET-X0112](https://jvet-experts.org/doc_end_user/current_document.php?id=11105) AHG9: On post-filter SEI [M. M. Hannuksela, E. B. Aksu, F. Cricri, H. R. Tavakoli, M. Santamaria (Nokia)]

This contribution proposes the following SEI messages into a working draft of a VSEI amendment:

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.

Additionally, this contribution proposes constraints to be included in VVC regarding the use of the proposed SEI messages in VVC bitstreams.

The proponents suggest the following observations based on their and others’ related previous contributions:

1. Post-filtering can be performed for multiple purposes, including but not limited to those mentioned in JVET-U0091 and JVET-T0076.
2. Different types of post-filters (e.g. ALF or a neural-network-based filter) can be used for the same purpose, such as generic visual quality improvement.
3. Content-adaptive NN-based post-processing for quality enhancement and super-resolution has been shown to provide a BD-rate reduction compared to pre-trained NN-based post-processing with the same neural network topology.
4. Coding neural network parameters with MPEG NNR has been shown to be a viable approach for content-adaptive NN-based post-processing for quality enhancement and super-resolution.

Questions/comments:

* Purpose of PfP 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)
* 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.
* 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?
* How would block-wise processing, block-overlapping, and boundary processing be managed?
* 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.
* What does URI mean, and what does it refer to?
* Extensibility aspect might require more consideration. In the past, extensibility was not often used, even if it had been foreseen in an SEI message.

It is generally agreed that the approach goes into a right direction, but it is not clear yet if it is timely to start a draft (including the current status of NNR adoption). Further study recommended as possible candidate of next VSEI extension.

## CE on Film Grain Synthesis (2)

Contributions in this area were discussed in session x at XXXX–XXXX UTC on XXday X Oct. 2021 (chaired by XXX).

[JVET-X0022](https://jvet-experts.org/doc_end_user/current_document.php?id=11188) CE: Summary Report on Film Grain Synthesis [S. McCarthy, M. Radosavljević, J. Shingala]

This contribution provides a summary report for the Core Experiment on film grain synthesis technology proposed at the W meeting of JVET. A total of 2 tests were completed within the CE between the JVET W and X meetings. In CE1, the use of the core 64x64 DCT-2 specified in VVC instead of the 64x64 DCT-2 defined in SMPTE RDD 5 specification was studied. In CE2, resolution-adaptive grain block size instead of fixed 8x8 grain block size was studied. CE1 results indicate that the selection of transform type does not affect subjective quality and does not change film grain processing time. CE1 also highlights additional implementation benefits. CE2 results indicate that resolution-adaptive grain block size does not affect subjective quality and reduces film grain processing time by approximately 20% for 4K content and 40-50% for 8K content. The software basis for the CE is VTM-13.2. JVET SDR-CTC Class A test sequences were use for complexity testing within the CE. Six test sequences from JVET FTP were used for subjective testing within the CE. All tests were conducted following CTC random access (RA) test conditions. Crosscheck results for the performed tests are submitted as separate documents and their summaries are integrated in this contribution.

1. **Methodology**

The film grain synthesis (FGS) processes studied in this CE are based on SMPTE RDD 5 [1] with two independent modifications to test: CE1, DCT-2 transform type; and CE2, grain block size.

No change in the syntax of the film grain characteristics (FGC) SEI message is required.

The CE is described in JVET-W2022 [2].

***Software***

CE test software is available on Git repo: <https://vcgit.hhi.fraunhofer.de/jvet-w-ce/ce-fgs/-/tree/CE-FGS>

CE test software is based on VTM-13.2 with extensions to be tested in CE proposed in JVET-V0093 [3] and JVET-W0095 [4].

CE software supports additional functionality as follows:

1. Film grain analysis supporting multiple intensity intervals as proposed in JVET-W0072 [5]
2. FGC SEI rewriter (described in section 1.1.1) enabling FGC SEI parameter values to be modified in existing bitstreams without re-encoding

**FGC SEI rewriter tool**

In order to facilitate the FGS CE experiments, an FGC SEI rewriter tool based on VTM13.2 is developed and is released with the CE software. This tool allows FGC SEI parameter values to be modified in existing bitstreams without re-encoding. It also supports inserting FGC SEI messages in a VVC bitstream originally encoded without FGC SEI message. The bitstream with the inserted FGC SEI has identical MD5sum value to the bitstream having the FGC SEI enabled during encoding. The tool supports three operational modes, controlled by SEIFilmGrainOption: 0-no change; 1-FGC SEI removal; 2-FGC SEI insertion; 3-FGC SEI rewriter.

1. FGC SEI removal (applied on bitstreams with FGC SEI)
2. FGC SEI insertion (applied on bitstreams without FGC SEI)
3. FGC SEI modification (applied on bitstreams with FGC SEI)

The two typical use cases and examples are given below:

1. A bitstream (test\_fg1.bit) is encoded with a pre-configured FGC SEI (fg1.cfg), but the film grain parameters are not producing satisfactory film grain effects. One can use the tool to modify the values (fg2.cfg) of the existing FGC SEI. Example command line:   
   ./SEIFilmGrainAppStatic -c fg2.cfg -b test\_fg1.bit -o test\_fg2.bit --SEIFilmGrainOption=3
2. A bitstream (test.bit) is encoded without FGC SEI so it cannot take advantage of benefit of the film grain synthesis. One can use the tool to easily insert the FGC SEI (fg.cfg) into the bitstream. Example command line:

./SEIFilmGrainAppStatic -c fg.cfg -b test.bit -o test\_fg.bit --SEIFilmGrainOption=2

The tool manipulates the VVC bitstream at NALU level, so the processing time is extremely fast (e.g., to rewrite FGC SEI in a 4K/8K bitstream, the entire processing time is less than 0.1 sec). The basic workflow is illustrated in Figure 1.



Figure 1 Basic workflow of FGC SEI rewriter

***Complexity Tests***

**Test sequences and conditions**

CE complexity tests using 4K content were conducted with the JVET SDR-CTC [6] Class A test sequences. For tests using 8K content, JVET SDR-CTC Class A test sequences were upsampled to 8K using HDRTools-v0.19.1. All tests were conducted following CTC random access (RA) test conditions [6].

CE tests were conducted for both SIMD enabled and SIMD disabled.

The three combinations of FGC SEI parameter values listed in Table 1 were tested to investigate the impact of number of intensity intervals and number of colour components on processing times for grain blending.

Table 1 Film grain configuration settings for complexity test

|  |  |  |
| --- | --- | --- |
| Film grain configuration | # of intensity intervals | Components |
| FGC\_SEI\_I10C3 | 10 | Luma & Chroma |
| FGC\_SEI\_I10C1 | 10 | Luma only |
| FGC\_SEI\_I1C1 | 1 | Luma only |

**Calculation of film grain processing times**

Absolute FGS processing time was measured with YUV output disabled. For each decoded bitstream, FGS processing time was calculated over all frames independently from decoding time. The reduction in FGS processing time is determined as follows:

Reduction in FGS processing time = (FT\_CEx.y – FT\_CEAnc ) / FT\_CEAnc

where FT\_CEAnc represents the FGS processing time for the anchor and FT\_CEx.y represents the FGS processing time for CE1.1 (x = 1, y = 1), CE2.1 (x = 2, y = 1) and CE2.2 (x = 2, y = 2).

***Subjective Tests***

For subjective tests, 6 test sequences with 1080p or 4K resolution (listed in Table 2) were selected from JVET FTP: <ftp://ftp.ient.rwth-aachen.de>. 4K test sequences were upsampled to 8K using HDRTools-v0.19.1. Random access test conditions were used with QP32 and QP37. Film grain configurations were adjusted for each clip. For CE2, decoded YUV files with and without FGS were viewed on a BenQ pro-level SW271C 27” 4K monitor.

Table 2 Test sequences (4K and 1080p) used in subjective test

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| SeqName | Category | Resolution | FPS | Frames | BitDepth | MD5 of Source YUV |
| MountainBay2 | Clean | 3840x2160 | 30 | 300 | 10 | f27b6b70244fb083baac546958fcf696 |
| OberbaumSpree | Clean | 3840x2160 | 60 | 600 | 10 | 6975b81c9e63c92b3bf4223796102da1 |
| Tango2 | Clean | 3840x2160 | 60 | 294 | 10 | 0471a59c423b7059c5c6c8b395e864a9 |
| RaceNight | Noisy | 3840x2160 | 50 | 600 | 10 | c06ed896df6e7dd90b4164df39652d50 |
| TrafficFlow | Noisy | 3840x2160 | 30 | 300 | 10 | 63135fc6fd7c73675f4ec1211b11dc93 |
| Crowdrun | Noisy | 1920x1080 | 50 | 500 | 8 | da34812b5b2c316d40481c7b6c841e41 |

***Crosscheck instructions***

Detailed crosscheck instructions were provided to JVET before the beginning of the current meeting. They are available at: <https://vcgit.hhi.fraunhofer.de/jvet-w-ce/ce-fgs/-/tree/CE-FGS> within the “results” folder (see *CEx\_FGS\_xCheck\_Instruction.pptx*). Bitstreams and additional supporting documents (per-sequence configurations, film grain configurations, md5sums) are also provided at the same location.

**Tests** related **to DCT2 transform type (CE1)**

The goal of CE1.1 was to study the use of VVC DCT-2 in the context of film grain synthesis, compared to previously used SMPTE RDD 5 DCT-2 (anchor).

The DCT-2 transform is used to pre-compute a grain pattern database from which grain blocks are selected during the grain blending process.

Fixed 8x8 grain block size was used for both test and anchor in CE1.1.

***Results for FGS processing times***

FGS processing times are provided for CE2.1 in Table 3

Table 3: Results of the complexity test

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | SIMD OFF | | | SIMD ON | | |
| Configure | CEAnc FT (sec) | CE1.1 FT (sec) | **Time Saving (%)** | CEAnc FT (sec) | CE1.1 FT (sec) | **Time Saving (%)** |
| FGC\_SEI\_I10C3 | 8.57 | 8.67 | **1.21%** | 3.99 | 3.87 | **-2.96%** |
| FGC\_SEI\_I10C1 | 5.43 | 5.53 | **1.89%** | 2.50 | 2.49 | **-0.21%** |
| FGC\_SEI\_I1C1 | 5.93 | 6.01 | **1.29%** | 2.66 | 2.59 | **-2.96%** |
| **Average** | 6.51 | 6.61 | **1.46%** | 2.98 | 2.92 | **-2.05%** |

Time difference between anchor and CE1.1 test is shown to be negligible. Small difference is the result of testing error (the way we utilize our cluster and the fact that the time between two executions can slightly vary). Note that the absolute difference is relatively small. The observed performance was expected, since both versions are implemented in the same way by using matrix multiplication. Unification with VVC core transform leads to additional implementation cost savings and implementation simplifications, see Sec. 4.1.

***Subjective*** test ***results***

Subjectively, there is no difference between anchor and CE1.1. In addition, the average difference per pixel is 0.5905 over all sequences and QPs (mean squared error), computed on subjective test set. Such a small difference in pixel intensity is not visible to human eye.

It is to note that legacy SEI messages are not impacted by the proposed modifications, since the results are subjectively the same.

1. **Tests related to grain block size (CE2)**

The goal of CE2 was to study resolution-adaptive grain block size. The anchor of CE2 was SMPTE RDD 5 with fixed 8x8 grain block size for all resolutions (HD, 4K and 8K). The test was use of resolution-adaptive BlockSize x BlockSize grain block size, where the variable BlockSize is determined as:

PicSizeInLumaSamples = PicHeightInLumaSamples \* PicWidthInLumaSamples

if PicSizeInLumaSamples <= (1920 \* 1080)  
 BlockSize = 8  
else if PicSizeInLumaSamples <= (3840 \* 2160)  
 BlockSize = 16  
else  
 BlockSize = 32

In addition, two sub-CEs tested different ways of pre-computing the grain database from which grain blocks are selected during grain blending. CE2.1 tested use of the transform specified in SMPTE RDD 5 to pre-compute the grain database. CE2.2 tested use of the DCT2 transform specified in VVC to pre-compute the grain database.

***Results for FGS processing times***

FGS processing times are provided for CE2.1 in Table 4 and for CE2.2 in Table 5

Table 4 Average reduction in FGC processing time for CE2.1

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | SIMD\_OFF | | | SIMD\_ON | | |
| Content | Configure | CEAnc FT (sec) | CE2.1 FT (sec) | **Reduction in FGS processing time (%)** | CEAnc FT (sec) | CE2.1 FT (sec) | **Reduction in FGS processing time (%)** |
| CTC4K | FGC\_SEI\_I10C3 | 9.99 | 7.97 | **-20.16%** | 3.45 | 2.70 | **-21.77%** |
|  | FGC\_SEI\_I10C1 | 6.46 | 5.13 | **-20.54%** | 2.21 | 1.70 | **-22.92%** |
|  | FGC\_SEI\_I1C1 | 6.86 | 5.47 | **-20.25%** | 2.31 | 1.78 | **-22.92%** |
|  | **Average** | 7.62 | 6.07 | **-20.32%** | 2.60 | 2.01 | **-22.54%** |
|  |  |  |  |  |  |  |  |
|  |  | SIMD\_OFF | | | SIMD\_ON | | |
| Content | Configure | CEAnc FT (sec) | CE2.1 FT (sec) | **Reduction in FGS processing time (%)** | CEAnc FT (sec) | CE2.1 FT (sec) | **Reduction in FGS processing time (%)** |
| CTC8K | FGC\_SEI\_I10C3 | 39.78 | 17.41 | **-56.25%** | 13.33 | 7.47 | **-43.98%** |
|  | FGC\_SEI\_I10C1 | 25.83 | 12.88 | **-50.14%** | 8.61 | 5.17 | **-40.04%** |
|  | FGC\_SEI\_I1C1 | 27.39 | 13.15 | **-52.01%** | 9.09 | 5.14 | **-43.43%** |
|  | **Average** | 30.42 | 14.34 | **-52.87%** | 10.14 | 5.83 | **-42.51%** |

Table 5 Average reduction in FGC processing time for CE2.2

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | SIMD\_OFF | | | SIMD\_ON | | |
| Content | Configure | CEAnc FT (sec) | CE2.2 FT (sec) | **Reduction in FGS processing time (%)** | CEAnc FT (sec) | CE2.2 FT (sec) | **Reduction in FGS processing time (%)** |
| CTC4K | FGC\_SEI\_I10C3 | 9.99 | 7.94 | **-20.46%** | 3.45 | 2.68 | **-22.15%** |
|  | FGC\_SEI\_I10C1 | 6.46 | 5.14 | **-20.45%** | 2.21 | 1.70 | **-22.93%** |
|  | FGC\_SEI\_I1C1 | 6.86 | 5.47 | **-20.26%** | 2.31 | 1.76 | **-24.01%** |
|  | **Average** | 7.62 | 6.07 | **-20.39%** | 2.60 | 2.00 | **-23.04%** |
|  |  |  |  |  |  |  |  |
|  |  | SIMD\_OFF | | | SIMD\_ON | | |
| Content | Configure | CEAnc FT (sec) | CE2.2 FT (sec) | **Reduction in FGS processing time (%)** | CEAnc FT (sec) | CE2.2 FT (sec) | **Reduction in FGS processing time (%)** |
| CTC8K | FGC\_SEI\_I10C3 | 39.78 | 17.27 | **-56.60%** | 13.33 | 7.38 | **-44.62%** |
|  | FGC\_SEI\_I10C1 | 25.83 | 13.10 | **-49.29%** | 8.61 | 5.27 | **-38.87%** |
|  | FGC\_SEI\_I1C1 | 27.39 | 13.14 | **-52.05%** | 9.09 | 5.15 | **-43.37%** |
|  | **Average** | 30.42 | 14.38 | **-52.74%** | 10.14 | 5.85 | **-42.34%** |

The data in Table 4 and Table 5 indicate that resolution-adaptive grain block size results in a significant reduction in FGS processing time compared with the fixed 8x8 grain block size CE-Anchor. The average time reduction is around 20% for 4K clips (with 16x16 grain block size) and around 40~50% for 8K clips (with 32x32 grain block size).

It is noted that the FGC processing time is affected mainly by the number of colour components processed. The number of intensity intervals signalled in the FGC SEI message does not significantly affect FGS processing time. (Note, the number of intensity intervals does not affect how grain blocks are selected from the pre-computed grain pattern database. Thus, no effect on FGC processing time is expected.)

The average reduction in FGS processing time, measured as a percentage, is not strongly dependent on SIMD. However, the absolute FGS processing time for SIMD enabled tests is approximately 2.5 to 3 times faster than the FGS processing time tests in which SIMD was disabled.

The average reduction in FGS processing time is not affected by which transform is used to pre-compute the grain database from which grain blocks are selected during the FGS process, i.e., the FGS processing times for CE2.1 (SMPTE RDD 5 transform) and CE2.2 (VVC DCT-2) are very similar.

***Subjective test results***

For subjective tests, similar visual quality was observed for CE2.1 and CE2.2 compared with the anchor for selected 1080p and 4K test sequences. For test sequences with grain or noise in the source, decoded pictures with grain blending appeared more similar to the source than decoded pictures without grain blending. For test sequences without grain or noise in the source, grain blending tended to mask compression artefacts (texture loss, basis patterns etc.) when coding at low bitrate.

1. **Complexity and memory requirements analysis**

The FGS processes studied in this CE are based on SMPTE RDD 5 with two independent modifications to test: CE1, DCT-2 transform type; and CE2, grain block size.

The FGS process described in SMPTE RDD 5 consists of the following steps:

1. Pre-compute a grain pattern database
2. Calculate an average luma value for each BlockSize x BlockSize block in the decoded picture to determine the intensity interval to which the block belongs
3. Calculate pseudo-random offsets to be used to select grain blocks from the pre-computed grain pattern database
4. Deblock vertical edges across adjacent grain blocks
5. Add grain blocks to the decoded picture

***Analysis of transform type (CE1)***

The VVC DCT-2 transform is well studied, documented, and designed to ensure efficient implementation in both software and hardware [7,8]. The VVC transform may also be already available on decoders and may thus be considered implementation friendly, reducing the need for one additional transform. Supporting multiple transforms has several consequences related to complexity and implementation costs, e.g., additional burden to hardware implementation. It requires additional memory for storing the multiple transform cores. Additional kernel leads to increased area since hardware sharing is not possible. Thus, by omitting one additional transform the cost of implementation is reduced, e.g., on-chip implementation, memory requirement reduction. Even efficient software implementation such as partial butterfly implementation is simplified since it can be reused from the core transform. In other words, the transform implementation is simplified and unified by means of using only DCT-2 from VVC.

Historically, there was much effort put in the research how to reduce the memory requirements for transform cores, excluding some transforms from consideration, reducing the number of bits per transform coefficient, reducing total number of unique values, etc. This effort was mainly summarized in Ljubljana and Macao JVET meetings (K and L meetings). For example, in VVC it is proposed to use 8-bit primary transform cores, optimizing the memory for storing the primary transform cores and relaxing additional implementation cost [7].

For example, additional memory requirement for RDD 5 transform is: 44 unique coefficient (no sign counted) that can be stored with 6-bits precision (RDD5 design is a straightforward way of deriving transform cores using scaling and rounding – scaling by 2^8). In total there are 4096 coefficients requiring 6-bits for storing transform core. From the analysis of this CE, it can be considered that this leads to potentially unnecessary resource allocation, and that all process can be implemented with VVC DCT-2 without any loss of quality.

***Analysis of grain block size (CE2)***

Estimates of theoretical computational complexity for the FGS process for the luma component are enumerated in Table 6. Memory requirements are listed in Table 7. The complexity and memory requirements for chroma components can be estimated by scaling the estimate for luma based on the chroma format idc.

Following conclusions can be drawn from the analysis of normalized operations per sample listed in Table 6:

1. FGS deblocking filter complexity is reduced to half for 16x16 block size compared to 8x8 block size. The filter complexity is reduced to a quarter for 32x32 block size.
2. The bit masking and bit extract operations are reduced by a factor of 6x for 16x16 and 24x for 32x32 compared to 8x8 block size
3. The modulo(%) operations are reduced by a factor of 4x for 32x32 block size, while they remain the same for 16x16 block size
4. Rest of the processing has similar computational complexity for all the block sizes but implementation specific benefits such as effective SIMD utilization and reduced process (function) overhead for larger blocks bring in additional benefits as explained in the additional section on software (SW) benefits below

Table 6 Theoretical computational complexity for different grain block sizes

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Operation** | **Total Operations** | | | **Operations Per Sample** | | |
| **Block size** | | | **Block size** | | |
| **8x8** | **16x16** | **32x32** | **8x8** | **16x16** | **32x32** |
| Block Average (for intensity interval selection) | | | | | | |
| Load | 64 | 256 | 1024 | 1 | 1 | 1 |
| Addition | 64 | 256 | 1024 | 1 | 1 | 1 |
| Shift | 1 | 1 | 1 | 1/64 | 1/256 | 1/1024 |
| Grain Block (fg\_block) creation using Pseudo Random Number Generated offsets x(r,ec) | | | | | | |
| Modulo (%) | 2 (per 16x16) | 2 | 2 | 2/256 | 2/256 | 2/1024 |
| Load/Store | 64\*2 | 256\*2 | 1024\*2 | 2 | 2 | 2 |
| Multiply | 64 | 256 | 1024 | 1 | 1 | 1 |
| BitMask / BitExtract | 6 | 4 | 4 | 6/64 | 4/256 | 4/1024 |
| 3-Tap Deblocking Filter for vertical edges of adjacent grain blocks | | | | | | |
| Load/Store | 32+16 | 64+32 | 128+64 | 3/4 | 3/8 | 3/16 |
| Additions | 32 | 64 | 128 | 1/2 | 1/4 | 1/8 |
| Shift | 32 | 64 | 128 | 1/2 | 1/4 | 1/8 |
| Grain Blending | | | | | | |
| Load/Store | 64\*3 | 256\*3 | 1024\*3 | 3 | 3 | 3 |
| Additions | 64 | 256 | 1024 | 1 | 1 | 1 |
| Clip | 64 | 256 | 1024 | 1 | 1 | 1 |

Table 7 Memory requirements for different grain block sizes

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Buffer** | **Total Memory** | | | **Memory per sample** | | |
| **Block Size** | | | **Block Size** | | |
| **8x8** | **16x16** | **32x32** | **8x8** | **16x16** | **32x32** |
| FG Database | X \* 4096 | X \* 4096 | X \* 4096 | X \* 4096/64 | X \* 4096/256 | X \* 4096/1024 |
| Recon buffer\* | 64 \* 2 | 256 \* 2 | 1024 \* 2 | 2 | 2 | 2 |
| FG Block\* | 64 \* 2 | 256 \* 2 | 1024 \* 2 | 2 | 2 | 2 |

**Notes:**

1. X can vary based on implementation:
   1. If all the 13x13 combinations of cut-off frequencies allowed in SMPTE RDD 5 are used for the grain pattern database, X = 169
   2. Otherwise, X is the number of component models with different cut-off frequencies as signalled in the SEI message of current access unit (e.g., max of 10 for SMPTE-RDD 5)
2. Recon buffer and FG block buffer sizes are multiplied by a factor of 2 because one current block and one previous block are required for vertical edge deblocking
3. The complexity for pre-computing the grain pattern database is negligible because the grain pattern database can be performed offline and stored as lookup.

Larger grain block sizes have additional benefits in SW due to following factors:

1. **SIMD benefits**: Larger grain block sizes can utilize the SIMD processing more efficiently compared to smaller grain block sizes. For instance, 8x8 grain blocks use 128-bit SIMD parallelism which provides only 8x parallelism while 32x32 grain block sizes can be efficiently optimized for 256-bit (AVX2) SIMD width which provides 16x parallel processing for 16-bit input data. Even if 8x8 grain block sizes are optimized for 256-bit AVX2 SIMD, it will still need to load and store 128-bit data from different regions of memory which will increase the loads and stores by a factor of 2x compared to larger block sizes. It will also require additional overhead of scatter gather instructions and significant restructuring of code to process adjacent 8x8 grain blocks.
2. **Function overheads**: For a given resolution, larger grain block sizes have fewer function calls compared to smaller block sizes because they process much larger data in one function call. ‘C’ calling function overheads are typically 15 cycles – 20 cycles for most x86 architectures. These function calls can be a considerable percentage of total cycles for simple operations related to block averaging and grain blending. Compared to 8x8 grain block sizes, function calling overhead is reduced by a factor of 4x for 16x16 grain block sizes and a factor of 16x for 32x32 grain block sizes which gives additional benefits beyond SIMD processing.

The benefits of efficient SIMD processing and reduced function calling overheads are reflected in the percentage FGS processing time reduction for 4k (16x16 block size) and 8k (32x32 block sizes) CE tests as shown in Table 3, Table 4 and Table 5.

Crosschecks

Crosscheck reports are summarized in Table 8.

Table 8 Summary of crosscheck reports

|  |  |  |
| --- | --- | --- |
| **Test** | **Crosschecker** | **Crosscheck report** |
| *CE-1.1* | Dolby, JVET-X0154 | **results:**    **code verification:**  .  **draft text verification:**  .  **further comments:** |
| *CE-2.1* | Alibaba, JVET-X0169 | **results:**    **code verification:**  .  **draft text verification:**  .  **further comments:** |
| *CE-2.2* | InterDigital, JVET-X01xx | **results:**    **code verification:**    **draft text verification:**  .  **further comments:** |

**Input** contributions **for CE**

1. Evaluation of VVC DCT-2 transform for Film Grain Synthesis (test CE1.1), M. Radosavljević, E. François (InterDigital), JVET-X0103.
2. CE: Film Grain Synthesis (test CE2.1 and CE2.2), F. Pu, S. McCarthy, W. Husak, P. Yin, T. Lu, A. Arora, T. Chen (Dolby), V. G R, J. N. Shingala, K. Patankar, S. K., A. Shyam (Ittiam), JVET-X0048.

Block size and DCT type are not signalled, left to implementation.

Questions/comments:

* Are the sequences representative enough for film grain? Only CrowdRun is shot with non-electronic camera? However, nowadays artificial generation of film grain might be more relevant to hide compression artifacts. As a follow-up comment, it is mentioned that this might not require sending parameters.
* Why does the adaptation of block size have such an effect on run time (in particular for the SIMD off case)? May be an implementation or compiler-setting issue.
* It is pointed out that bigger block size could require more line buffers, which might not be implementation friendly in hardware. Run time may not be the only aspect of complexity.
* If the purpose of the CE is to investigate technology for a possible TR, what would be its purpose? Would it only describe this technology, and as it is non-normative, would it discourage usage of other technology already common?
* It is commented that the usage of VVC DCT would avoid an additional transform, in which case it would no longer be bit-exact matching with SMPTE RDD-5 (not backward compatible, and therefore may not be allowing using the noise pattern database from RDD-5). If the purpose is enforcing its usage this however might better be fulfilled with a TS rather than a TR. It is further commented by proponents that it would not be required to invoke that specific synthesis if the SEI message appears in the bitstream.

Some concern is expressed about the option of making one method mandatory.

Could it be an option to issue a technical report that describes the method as one possibility of implementing analysis and synthesis related to the FG SEI message, while also pointing to other options of implementing the synthesis (e.g., RDD-5)? Also the option of including the software as an example in VTM is suggested.

Further discussed in session 21 1710 UTC

It was suggested from the proponents to develop a TR which includes 3 examples of implementation (autoregressive, frequency-based with two variants, the one from CE and RDD-5). In particular for the AR method, additional study is needed if it could be linked with the existing SEI message, or other messaging means would be required. This is agreed.

Decision(SW): Adopt the software from the CE, JVET-X0048 to VTM

New AHG to study and develop the draft report; AHG should also investigate whether the software could be ported to HM.

[JVET-X0048](https://jvet-experts.org/doc_end_user/current_document.php?id=11041) CE: Film Grain Synthesis (test CE2.1 and CE2.2) [F. Pu, S. McCarthy, W. Husak, P. Yin, T. Lu, A. Arora, T. Chen (Dolby), V. G R, J. Shingala, K. Patankar, S. Kadaramandalgi, Ajayshyam (Ittiam)]

[JVET-X0169](https://jvet-experts.org/doc_end_user/current_document.php?id=11177) Crosscheck of JVET-X0048 (CE: Film Grain Synthesis (test CE2.1 and CE2.2)) [R.-L. Liao (Alibaba)] [late]

[JVET-X0182](https://jvet-experts.org/doc_end_user/current_document.php?id=11192) Crosscheck of JVET-X0048 (CE: Film Grain Synthesis (test CE2.1 and CE2.2)) [[M. Radosavljevic (Interdigital)](mailto:milos.radosavljevic@interdigital.com)] [late]

[JVET-X0103](https://jvet-experts.org/doc_end_user/current_document.php?id=11096) [CE] Evaluation of VVC DCT-2 transform for Film Grain Synthesis (test CE1.1) [M. Radosavljevic, E. François (InterDigital)]

[JVET-X0154](https://jvet-experts.org/doc_end_user/current_document.php?id=11161) Crosscheck of JVET-X0103 ([CE] Evaluation of VVC DCT-2 transform for Film Grain Synthesis (test CE1.1)) [F. Pu (Dolby)] [late]

## Non-SEI HLS aspects (0)

Contributions in this area were discussed in session x at XXXX–XXXX UTC on XXday X Oct. 2021 (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 (especially those of Monday 12 July 0750–0920 and Wednesday 14 July 0820-0935).

Monday 11 October 0745–0920:

* Joint meetings:
  + Mon 16:20-17:20 with AG5, WG4, WG7 on guidelines for verification testing and remote experts viewing
  + Request JM with WG3 Wed/Thu: VDI, green metadata etc.?
* Planning of outputs: Software, Conformance, DoCRs?, EEs, CE?
  + Conformance v1 FDIS&DoCR, Conformance v2 CDAM&Request
  + BoG on Conformance (I. Moccagatta, D. Rusanovskyy) to address v1 ballot comments, and clarify remaining v2 issues (Tue 5-7, B. Bross creates meeting)
  + Software v1/v2 FDIS in January; draft DoCR (Gary)
* Scheduling, remaining doc reviews
* Whether to hold a hybrid meeting in April 2022
* Review docs from section 4.x

Wed. 14 Jul. 0820-0935:

* Report from NN BoG:
  + Further work on summary analysis (Excel sheet attached to report)
  + Updates on future contributions’ reporting requirements
  + EE1 planning (JVET-W2023): proposals identified for next round
  + More emphasis on cross-checking, training should be cross-checked as well
  + No need to meet again; offline activities on preparing EE and CTC documents
  + Recommendations of BoG (as per v3 of JVET-W0182, further editing cleanups in v4) were approved
  + Plan liaison statement to JPEG on NN activities (E. Alshina and A. Segall were asked to prepare text for that)
* Report from BoG on VVC V2 conformance (JVET-W0188) discussed at 0850
  + Identified 65 desired bitstreams
  + Plan to issue draft 1 of VVC V2 conformance
  + Clarification and agreements: 9 new profiles, new tools not enabled for Main 12, Main 12 Still Picture, Main 12 Intra
  + Add chairs to AHG5
  + Make a request for ISO amendment? (this did not seem necessary/appropriate yet)
* EE2 (JVET-W2024):
  + It was agreed to also investigate the previous 1.5a in the next round, together with 1.5b
  + EE document to be prepared offline (by previous coordinators)
* CE on high bit depth to be discontinued
* New CE on film grain: Investigate the method from JVET-W0095 versus the RDD-5, in terms of complexity and quality; also investigate the impact of the block sizes, and whether larger block sizes than 8x8 would be necessary (JVET-W2022); it is commented that it would be desirable having a precise description of the algorithms, as useful for a TR later.

## 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 11 Oct. 0500–0730, Wednesday 13 Oct. 0500–0630, and Friday 15 Oct. 2100–2300. The status of the work in the MPEG WGs was reviewed at these information sharing sessions.

## Joint meeting with Q6/16 (VCEG) and SC 29/WG 2 MPEG Requirements 0730–0845 Tuesday 13 July

The following topics were discussed in this joint session. See also the notes recorded on these topics in other sections of this document.

* VVC V2 Profiles (Tuesday 0730-0845 UTC with JVET and MPEG Requirements on VVC v2 profiles)
  + [JVET-W0136](https://jvet-experts.org/doc_end_user/current_document.php?id=10965) Suggested initial profile text for VVC operation range extension [T. Ikai (Sharp)] [late]
    - Monochrome profiles proposed
    - All-intra video profiles proposed (with loop filter support required)
    - 10, 12, and 16 b profiles proposed
    - Monochrome, 4:2:2, 4:4:4
    - Inter-predictive profiles proposed (unlike in HEVC)
    - Main 12 Still Picture, Main 12 4:4:4 Still Picture proposed (not in HEVC)
    - ~16 profiles were proposed in the contribution
    - Discussion:
      * Tiers and levels roughly per HEVC (OK for now – JVET-W0183?)
      * Do we need 4:2:2 profiles? (maybe not)
      * Do we need monochrome profiles? (maybe not)
      * Loop filter support requirement? (OK)
      * Palette mode and new coding tools proposed to be supported except in 10/12 b 4:2:0 & 10/12 b Mono profiles (OK)
        + (Some tools are already conditioned on b > 10)
* Film grain (optional mandatory, additional profile(s), or as an alternative VVC-based example)
  + JVET-W0095, proposed 8x8 blocks to be used for HD, 16x16 for 4K, and 32x32 for 8K
  + The possible publication of a technical report on the subject was discussed
  + Reference software can be provided
  + The SMPTE relationship was discussed, as SMPTE had previously published the RDD 5 scheme.
  + Discussion in VSEI vs VVC was discussed, as the proposed scheme uses VVC-based transforms.
  + The scheme was described as not necessarily for pre-filtering removal with post-process restoration of grain, but rather as a potential method for subjective improvement of low bit rate coded video
  + The current proposal is for inclusion in VSEI as an alternative example

## Joint meeting with MPEG WG3 0630–0700 Wednesday 14 July on Systems metadata

Topics of discussion:

* Proposal m57457 is to use a single code for any system metadata rather than multiple SEI messages as proposed in JVET-W0169. The MPEG Systems WG did not have a strong opinion on that.
* Should the specification of such an SEI message go in VSEI or in a video coding standard directly? Tentatively directly in the video coding standard.
* What should it reference? Should it reference multiple systems standards (initially two)? The tentative answer is yes.
* Should it be extensible? Yes.
* Timeline? VDI work might go to CD at this meeting, with green metadata following shortly after.
* It was agreed we are not planning to put such SEI message(s) into the VVC v2 DAM text; it should follow a bit later.

Conclusions:

* Codepoints for Green Metadata and VDI: Better to define separate ones in VVC, and refer the corresponding systems specs directly
* SC 29/WG 3 would have the primary responsibility on what is going inside such an SEI message
* JVET should start a VVC amendment in October or January to align with the progression in WG3 (expecting this could be final at DAM stage)

## Joint meeting of SC 29/AG 5 with WGs 4 (MPEG Video), 5 (JVET) and 7 (MPEG 3D Graphics Coding) 0720–0820 Wednesday 14 July

There was a presentation of JVET-W0185 (MPEG m57393) on remote testing coordination of proposals by J. Jung. See the summary of that document in section 4.4.

* Further work was expected to study and refine various aspects.
* It was commented that further information on the setup of testing sites and the level of training for test subjects should be further detailed.
* It was commented that information on suitable players for being used in the testing purposes would be very valuable. Initiatives from experts are solicited to contribute to the development of a suitable open source player to enable stable playout capabilities for these. It was commented that HDRtools provides means to pack raw video files and associated metadata into mp4 files which enlarges the number of players to handle this data.
* It was suggested to clarify the status of verifications tests in the standardization process in the guidelines document.
* The relationship with ITU subjective assessment Recommendations was discussed; nearly all aspects are well aligned – perhaps the forced-choice A-B comparison is less well established (5-grade comparison vs. 4-grade comparison)
* The scheme is currently not intended for formal subjective assessment, but rather for informal testing.
* It was asked how well verified it is that a remote viewing test can match the results of a lab test; this is a major question for study, e.g., in VQEG. However, since testing in labs is currently not feasible, such studies are not immediately planned.
* Also, expert viewing tests would not necessarily always match naïve subject viewing in any case.

There was also a presentation by M. Wien of draft verification test guidelines that had been submitted to MPEG as m57463.

* This document defines guidelines for MPEG verification tests to serve as a reference for ongoing and future verification test activities. It describes verification testing of standards which define visual media output. In order to assert the suitability of the standard for visual assessment by users, formal subjective evaluation of the visual media signal reconstructed or synthesized from the compressed bitstream is needed.
* MPEG verification tests has been since the beginning a fundamental step in the path to the release of a new standard. Nevertheless, the new organization of SC29 in Working Groups and Advisory Groups induces the need for a procedure defining the necessary communication and the exchange of test results between the relevant WGs and AG5. The document is supposed to reflect the common understanding of previous verification test activities in MPEG, and the transition of the best practice into the new organizational structure of MPEG.
* The description currently relies on the assumption of video sequences to be compressed and reconstructed. For visual media which is rendered (e.g. in case of 3D-video, or 3DoF, 3DoF+, and 6DoF), additional steps may be required and should be added in a revision of this document.
* There was discussion of types of encoder optimization, depending on what is intended to be tested and demonstrated.
  + Having similar types of optimization vs independent optimization of each encoder separately
  + Similarity of configurations - e.g., intra period and GOP structures – in some cases a different GOP structure is intended to be part of what is being tested, and in other cases it is not
* It was suggested that encoder implementations used in the verification test should include optimized rate allocation algorithm, potentially also rate control. It was commented that the level of optimization of encoder implementations of a new specification typically is lower than for a longer available, well-studied anchor scheme. At the same time, the test model implementations typically follow comparable encoder decision strategies, e.g. including fixed QP settings and rate-distortion based tool decisions. Further, these test model implementations and applicable configurations are typically well understood by the group such that a fair comparison is achievable. Nonetheless, additional implementations may be considered, such as the example of VVenC in the context of the VVC verification tests.
* It was further recommended to add a note to the reporting section on highlighting specific features or differences between the new scheme and the anchor in order to make readers of the public document aware of such aspects.
* A list was included of what prior verification tests have been conducted in recent years.

## BoGs (2)

The following break-out groups were established at this meeting to conduct discussion and develop recommendations on specific topics.

[JVET-X0209](https://jvet-experts.org/doc_end_user/current_document.php?id=11219) EE1-related: Report on results of JVET-X remote viewing session [M. Wien]

Presented in Session 21 1520 UTC

Remote expert viewing tests were conducted during the 24th JVET meeting for the exploration activity on DNN-based coding tools on UHD test sequences as defined in EE1. Calls for participation in the subjective viewing were issued on the JVET reflector and verbally during JVET sessions. Overall, 17 experts volunteered in participating in the tests. The testing procedure followed the draft SC29/AG5 guidelines for remote experts viewing, providing an A/B comparison of sequences under test.

The results reveal some performance improvements for the proposals or similar performance compared to the VTM anchor. In one case, a slight loss is observed. In the super-resolution category, the visual benefit of JVET-X0064 appears to be comparable to the visual benefit observed for the RPR configuration under test reported in JVET-X0117. The results for the loopfilter category suggest some slight visual benefit of JVET-X0066 over the VTM anchor.

Three sequences (Campfire, Parkrunning, Tango) were investigated at QPs 37 and 42 each for superresolution. Another three sequences (Catrobot, Daylightroad, Tango) were investigated at QPs 37 and 42 each for loop filter. At QP 42, differences/benefits seem easier to spot due to increased number of artifacts. Duration was 5 sec for all sequences. Parkrunning was displayed too fast (60 rather than 50 fps)

Could more sequences be used? Foodmarket might be problematic due to scene change.

Would other QP points (rather than CTC) be used for visual testing? Not high priority now. Test is successful in confirming some improvements and no serious artifacts are produced by NN technology.

[JVET-X0188](https://jvet-experts.org/doc_end_user/current_document.php?id=11198) BoG Report: EE1 Viewing Preparation and Neural Networks Video Coding Results Analysis [A. Segall]

First reported in session 7 Thu 1300 UTC. Visual testing being prepared for beginning of next week; update of summary for EE and EE related being prepared, the NN library was discussed and recommended for AHG study.

It is asked if the library would be linked to VTM/ECM. This would in principle be possible. It was mentioned that it should come with the usual software licensing conditions.

Final report given in session 21 1600 UTC.

The BoG held the following meetings during the 24th JVET meeting:

• October 7 – 05:00-8:30 UTC

• October 8 – 06:05-06:40 UTC

Recommendations from the BoG are summarized as:

 Visual Evaluation

• Perform visual evaluation using three categories:

o Loop filtering: JVET-X0066

o Super-resolution: JVET-X0064

o VVC RPR: JVET-X0117

• Regenerate data for JVET-X0064 to better match the bit-rate of the anchor

• Follow the naming and data generation described in Section 2.1.2

• Following the schedule described in Section 2.1.3 that will complete the visual evaluation by October 14.

 EE Results and analysis

• Use the total number of parameters and parameter precision in high level reporting summaries

• Encourage (but not require) the use of the JVET-W0181 library in future development

• Add mandate to AHG11 to support and study the JVET-W0181 library

1 Mandates

The BoG was established with the following mandates:

• Visual evaluation preparation

o Identify the sequences and rate points to use for visual evaluation

o Confirm the visual evaluation procedure

• EE results analysis

o Review and refine the performance summary, including the memory requirements reported by proponents. Perform further analysis and align reporting as necessary.

o Discuss how to improve our analysis of the impact of training. Refine the reporting conditions as necessary.

o Discuss the possibility of using the library proposed in JVET-W0181

2 Visual Evaluation Preparation

The following questions were proposed to be considered:

1. What content should be included in the visual evaluation?

a. Sequences

b. QP points

c. Proposals

2. How should the content be prepared and delivered?

3. What is the schedule?

The group agreed that these were the main questions to be answered.

2.1 Discussion

2.1.1 On content that should be included in the visual evaluation?

One participant suggested the group decide if we should conduct viewing at 4K or HD resolutions.

The group discussed that a 4K session and HD session could be performed.

The group then discussed what proposals to include in the evaluation.

The JVET-X0064 proposal was discussed, and it was reported that the bit-rate difference between the anchor and test was on the order of 20%. It was proposed to re-generate the results for JVET-X0064 so that the bit-rates were closer to the anchor bit-rates. The proponents of JVET-X0064 expressed a willingness to generate the needed data.

It was commented that having a more dense sampling of the anchor could be beneficial going forward.

It was commented from the previous visual evaluation that we should temporally crop the CampFire sequence to 5 seconds.

It was commented from the previous visual evaluation that the sequences should not include scene cuts.

The group recommended including the following proposals, sequences and QPs in the visual evaluation:

Category Proposals Sequences Anchor QPs

Loop Filtering/Post-Processing JVET-X0066 (9.8% BD-rate) DaylightRoad (13.3%)

CatRobot (11.5%)

Tango (11.1%) 37,42

Super Resolution JVET-X0064 (6.1% BD-rate A1/A2)

CampFire\* (19%)

Tango (9.4%)

ParkRunning (9.2%) 37,42

Super Resolution JVET-X0117 (2.4% BD-rate A1/A2) CampFire\*

Tango

ParkRunning 37,42

\* CampFire should be temporally cropped to 5 seconds.

2.1.2 On content should be prepared and delivered?

One participant suggested the following naming convention:

<proposal\_name>

<sequence\_name>\_qp<qp\_value>.mp4

<sequence\_name>\_qp<qp\_value>.mp4

…

<sequence\_name>\_qp<qp\_value>.mp4

where the <qp\_value> is the QP value that was used for coding. And, the files would be placed in a directory using the proposal name.

The group recommended using this naming convention.

The data above is needed for the anchor and proposals.

The sequences should then be converted to MP4 using the command line:

ffmpeg -s:v <pix>x<lin> -c:v rawvideo -pix\_fmt <input\_pix\_fmt> -r <fps> -start\_number <start\_number> -i <input.yuv> -c:v libx265 -crf 15 -tag:v hvc1 -pix\_fmt <output\_pix\_fmt> <output.mp4>

where

<pix> denotes the number of pixels per line,

<lin> denotes the number of lines per picture,

<input\_pix\_fmt> denotes the input pixel format, yuv420p10le for 10 bits, yuv420p for 8 bits,

<output\_pix\_fmt> denotes the output pixel format,

<fps> denotes the frame rate of the video.

<start\_number> should be 150 for CampFire and 0 for all other sequences.

One participant commented that -vsync passthrough may need to be added as the first command line option for the command above.

Owners:

Proposal Owner

JVET-X0066 Yue Li

JVET-X0064

Chaoyi Lin

JVET-X0117 Elena Alshina

Anchor Elena Alshina

Original Elena Alshina

2.1.3 On the schedule

The following schedule was proposed:

Regeneration of JVET-X0064: Saturday End of Day (October 10 - UTC)

Upload of sequence data: Sunday Afternoon (October 11 – UTC)

Sanity Checking: Sunday Evening (October 11 – UTC)

Test Preparation: Sunday Evening (October 11 – UTC)

Conducting tests: Monday (October 12 – UTC)

Reporting results: Wednesday (October 14 – UTC)

The group recommended using the schedule above.

3 EE Results Analysis

The following issues were proposed to be considered when analyzing the EE results:

1. Is the memory summary information reported in the EE (and outside the EE) consistent?

2. What should we do to improve our analysis of training?

3. Should we use the library proposed in JVET-W0181 going forward?

The group agreed that these were the main questions to be answered.

3.1 Discussion

3.1.1 On what memory data should we use for EE (and NNVC) reporting?

One participant suggested that it may not be clear on what memory was requested in the EE and CTC.

One participant commented that the CTC and EE request the (i) total number of parameters (millions) and (ii) and precision of the parameters.

One participant commented that previous summary results were graphically shown in MBs, which were calculated from the CTC and EE data requested.

It was suggested to use the following approach for future summaries:

1. Report the Total Number of Parameters (Millions)

2. Report the Precision of the Parameters

3. Calculate (and graph) the total memory size in MBs. (Calculated by multiplying the Total Number of Parameters by the number of bytes corresponding to the Reported Precision.)

It was commented that it would be beneficial to also capture the largest single model size.

It was also commented that it would be beneficial to also capture the largest memory size required to decode one slice and/or one frame.

It was generally agreed that reporting the total memory size at different granularities is useful for analysis. These granularities could include the total memory size for the algorithm, the total memory size for the frame, the total memory size for the slice, or the total memory size for a different granularity. It will be valuable for the group to continue to discuss and refine the approach. For now, it is recommended to use the total memory size for the high-level summary activities. Information at other granularities is very welcome and should be discussed in the context of each proposal.

It was suggested to make it visually clear in the template what parameters are mandatory. This was generally agreed.

Next steps: Offline activity to use the above approach to update and/or refine the EE1 summary. Additionally, include the non-EE1 contributions to this meeting in the summary.

In the second meeting of the BoG, the result of the offline activity was reviewed. The data was confirmed by the group. During the review, the plots were also refined.

It was recommended to provide the summary document as an official output of the BoG.

3.1.2 On what should we do to improve our analysis of training?

There was no proposal to improve the analysis of training.

It was commented that reproducing the inference results would be a higher priority than cross-checking the training procedure.

3.1.3 On should we use the library proposed in JVET-W0181 going forward?

One participant asked what constraints should be considered with the JVET-W0181 library.

A cross-check of JVET-X0118, which includes the JVET-W0181 library, was reported in JVET-X0123. The cross-checker reported that they were able to run it, and that the results matched between both Linux and Windows implementations. It was reported that the command line parameters that JVET-X0118 used to interface with the JVET-W0181 library were inconvenient in their cluster environment, and it was suggested to move the parameters into a configuration file. The cross-checker further reported that this cross-check was the first successful cross-check they had performed using that cluster.

One participant asked if there was any information about the run-time of the library compared to other solutions.

The cross-checker reported that the JVET-W0181 did not appear to be slower than TensorFlow and/or PyTorch.

One participant expressed concerns on if JVET-W0181 could be used for both deep and shallow networks designs.

The contributor of JVET-W0181 highlighted that one benefit of the library was the transparency of the implementations as compared to other, potential, solutions. This could be beneficial when measuring complexity, for example. It was further highlighted that the library could be useful for comparing and interacting with the AHG12 activity.

One participant expressed support for using a library that was transparent in its implementation.

One participant reported that they had used a Docker solution successfully in their cluster environment.

One participant asked what frameworks and version were currently supported.

The contributor reported that JVET-W0181 support TensorFlow 1 and 2 as well as PyTorch v1.4.

One participant asked how much support the contributor of JVET-W0181 could provide in adding new features and/or custom layers to the library.

It was proposed that the JVET-W0181 be encouraged for future EE activity. However, the library should be not required. This was recommended by the group.

It was recommended to add a mandate for the AHG to maintain and study the library.

From JVET discussion:

The BoG report contains Excel sheets which are augmenting the EE report by additional information

The template for the next EE will be developed with the analysis methods in the BoG

It is further confirmed that cross-checking shall include the inference stage in first place, and if a technology is considered for “adoption”, also the training should be cross-checked to confirm the performance, and that no over-fitting happened.

For combining several tools in a common software, the common library could be a central element.

E. Alshina to coordinate the drafting of next EE description, for first review in session 23 Thu 1300.

[JVET-X0207](https://jvet-experts.org/doc_end_user/current_document.php?id=11217) JVET BoG Report: VVC v1/v2 Conformance Testing [I. Moccagatta, D. Rusanovskyy]

This is a report of activities from the BoG on VVC v1/v2 Conformance Testing. The BoG call was held during the 24rd JVET meeting, on October 12, at 5:00 – 7:15 GMT. Meeting time and zoom link was announced over the JVET reflector and made available through JVET calendar.

As output of the meeting, the BoG recommended the following:

1. Approve the text of the document JVET-X0161-v2 (VVCv1 conformance testing).
2. Approve the text of the document [JVET-X0185](https://jvet-experts.org/doc_end_user/current_document.php?id=11195) (VVCv2 conformance testing).
3. Approve the collaborative procedure for VVCv2 conformance bitstreams cross-check.

Cross-check to be done offline, prior to the upload to the JVET FTP. Progress of the bitstream generation and cross-check to be tracked by the Excel document maintained by conformance testing coordinators.

1. Submit conformance bitstreams generated according to the JVET-W2026 specification to the VVCv2 CDAM ballot.
2. Extent the VVCv2 conformance (in next versions) by adding a new category on WPP (JVET-X0128).
3. Use independently designed VVCv2 decoder for the cross-check, is available. Experts having access to such decoders are encouraged to volunteer for the conformance testing.
4. Discuss in JVET if the list of AhG5 chairs need to be extended by adding more experts contributing to the VVCv2 development and cross-check.
5. **Agenda**

The BoG met with following agenda:

1. Review of the input documents on conformance testing:

* [JVET-X0161](https://jvet-experts.org/doc_end_user/current_document.php?id=11168) AHG5: Editors update on VVC conformance testing [J. Boyce, E. Alshina, F. Bossen, K. Kawamura, I. Moccagatta, W. Wan].
* [JVET-X0185](https://jvet-experts.org/doc_end_user/current_document.php?id=11195) AHG5: Editors update on conformance testing for VVC operation range extensions [D. Rusanovskyy, H.-J. Jhu, I. Moccagatta, M. Sarwer, Y. Yu, T. Zhou].

1. Discuss organization of the VVCv2 conformance testing:
   * Agree on cross-check procedure for VVCv2 and bitstream exchange.
   * Complete the list of volunteers for bitstream generation and cross-check.
   * Extension of the VVCv2 conformance test set.

1. **Input documents review:**

* **Document [JVET-X0161](https://jvet-experts.org/doc_end_user/current_document.php?id=11168) was presented by Iole.**

No objections to the content have been received on the text of the document. The following items have been discussed:

1. Editorial changes:
   1. Removal of informational section 6.5.4. No voices were voiced on preserving this section in the text.
   2. Adding of IBC\_E\_Tencent bitstream, section 6.6.2.14.5.
   3. Table numbering update.
2. Comments from Summary of voting on ISO IEC DIS23090-15 conformance (m57766):
   1. Comments suggestion from Gary, Iole, Jill and Frank have been received and reflected in the document.
   2. Further review of the document is encouraged.
3. No reports on corrupted bitstreams or missing coverages were received during the BoG.

BoG recommended to approve the document JVET-X0161-v2 as basis of FDIS text. This recommendation is agreed by the JVET plenary in session 21.

DoC document to be reviewed on Friday

* **Document [JVET-X0185](https://jvet-experts.org/doc_end_user/current_document.php?id=11168) was presented by Dmytro.**

No objections to the content have been received on the text of the document. The following items have been discussed:

1. Editorial section of the document, incl. Table 1 and Table 2.
   * Collaborative cross-check procedure for VVCv2 conformance testing has been presented and agreed in principle.
   * Excel document to maintain of the progress on VVCv2 conformance bitstreams generation was presented (X2026\_StatusTable). Document to record following stages of preparation:
     1. List of categories and bitstreams.
     2. Volunteers for bitstream generation and cross-check.
     3. Availability of the bitstream and results of the cross-check, incl. notes and software version.
   * It was suggested to include names of PoC in the conformance testing coordination documents.
2. Status of the bitstream generation:
   * 97 streams have been generated (with 85 bitstreams being submitted to the dropbox).
   * No cross-check was conducted yet due to FTP access issues.
   * Few categories have been identified as missing of volunteers:
     1. Sony and KKDI volunteered to generate the bitstreams for those categories.
   * 4 14 bitstreams have been identified as missing cross-checkers:
     1. Qualcomm to take 6 additional bitstreams.
     2. OPPO to take 4 additional bitstreams.
     3. 4+ bitstreams are still missing cross-checkers.
   * BoG recommended volunteers to step forward to cover missing 4.
3. Bitstreams description in JVET-X0161:
   * BoG recommended that editing of the conformance to add bitstream description be done collaboratively. Details to be discussed off line. Bitstream description should follow template provided in conformance v2.
4. VVCv2 conformance test set extension due to adoption of the JVET-X0128.

BoG recommended to approve the document JVET-X0185 as basis for next draft (and CDAM).

BoG recommended to extend the VVCv2 conformance by adding a new category on WPP (JVET-X0128).

These recommendations are agreed by the JVET plenary in session 21.

The CDAM request document needs to be available by Friday.

1. **Organization of the VVCv2 conformance testing:**

* **VVCv2 collaborative cross-check procedure**
  1. The procedure was agreed at the last BoG and implemented in output document JVET-W2026.

This process specifies that volunteers for the bitstream generation will also contribute to the cross-check of the submitted bitstreams from other volunteers. Bitstreams assignment, progress on the generation and cross-check as well as comments (reports) to be recorded in the Excel document maintained by the VVCv2 conformance test coordinators. Goal is to reduce amount of cross-checking efforts of the AHG5 chairs.

BoG recommended the cross-checking to be done collaboratively.

* 1. Two options to organize cross-check and were discussed:

Option 1: Streams will be uploaded to the JVET drop-box and made available for cross-check by moving from dropbox to under\_test.

Option 2: More efficient to have the cross-check done off-line before uploading to the dropbox.

BoG recommended Option 2.

* 1. After cross-checked streams are uploaded to dropbox coordinators will take over.
  2. Note: Iole to delete streams currently in dropbox as not cross-checked. New bitstreams to be uploaded after offline cross-check.

BoG recommended experts who have access to independently decodable decoder to step and help.

* **Extension of the conformance test set and SW version.**

The following items have been discussed:

* Potential adoption at this meeting:
  + JVET-X0128 low-level adoption impacts WPP use case, will not affect existing bitstreams if they generated with WFS disabled.
  + HLS adoptions have been assessed as not affect parsing of the existing bitstreams.
  + Current bitstreams have been generated by VTM-14.0, as specified in JVET-W2026 do not cover adoptions from this meeting.
  + BoG recommends submitting to CDAM existing streams, and the new adoptions to be reflected in the bitstream generated and cross-checked using the next version of the SW.

1. **Other items:**
   1. Tomonori Hashimoto will contribute to work on VVC v2 conformance testing and AhG5 as Sharp’s representative.

2. It was suggested to extend existing list of AhG5 chairs with larger number of VVCv2 contributing experts.

## Liaison communications

The JVET did not directly receive any liaison statements at its current meeting, although a liaison communication relevant to JVET work was conveyed at the parent level from ITU-T SG 16 to ISO/IEC JTC 1/SC 29 for review at the SC 29 meeting planned to be held during the week following the JVET meeting. One outgoing liaison letter was prepared by JVET, which was to ISO/IEC JTC 1/SC 29/WG 1 (JPEG) to inform them of the status of the JVET exploration experiment work on neural network based video coding, which has some relevance to recent JPEG work on its new JPEG AI image coding project. The outgoing liaison statement was issued as WG 5 output document N 78.

# Project planning

## Software timeline

ECM2 software (including all adoptions) was planned to be available 2 weeks after the meeting. F. Le Léannec and J. Chen were named as additional coordinators (in addition to V. Seregin and K. Zhang).

VTM14 software was planned to be available on 2021-08-09.

## Core experiment and exploration experiment planning

A CE on film grain synthesis was established, as recorded in output document JVET-W2022.

An EE on neural network-based video coding was established, as recorded in output document JVET-W2023.

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-W2024.

Initial versions of these documents were presented and approved in the plenary on Friday 16 July.

## Drafting of specification text, encoder algorithm descriptions, and software

The following agreement has been established: the editorial team has the discretion to not integrate recorded adoptions for which the available text is grossly inadequate (and cannot be fixed with a reasonable degree of effort), if such a situation hypothetically arises. In such an event, the text would record the intent expressed by the committee without including a full integration of the available inadequate text.

## Plans for improved efficiency and contribution consideration

The group considered it important to have the full design of proposals documented to enable proper study.

Adoptions need to be based on properly drafted working draft text (on normative elements) and HM/VTM encoder algorithm descriptions – relative to the existing drafts. Proposal contributions should also provide a software implementation (or at least such software should be made available for study and testing by other participants at the meeting, and software must be made available to cross-checkers in EEs).

Suggestions for future meetings included the following generally-supported principles:

* No review of normative contributions without draft specification text
* VTM algorithm description text is strongly encouraged for non-normative contributions
* Early upload deadline to enable substantial study prior to the meeting
* Using a clock timer to ensure efficient proposal presentations (5 min) and discussions

The document upload deadline for the next meeting was planned to be Tuesday 13 April 2021.

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 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 in session 25 on Wednesday 28 April 2021.

|  |  |  |
| --- | --- | --- |
| **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 studies. * Report on project status to JVET reflector. * Provide a report to the next meeting on project coordination status. | J.-R. Ohm, G. J. Sullivan (co-chairs) | 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-W2005 and JVET-W2006). * 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-W1004 errata output collection. * Produce and finalize JVET-W2002 VVC Test Model 14 (VTM 14) 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))   * Produce the verification test report JVET-W2020, and consider plans for additional verification testing of VVC capability. * 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-U2008 draft conformance testing specification and investigate the need for extensions. * Study the JVET-W2026 draft conformance testing for operation rage extensions and investigate the need for improvements. * Study the requirements of VVC, HEVC, and AVC conformance testing to ensure interoperability. * Maintain and update the conformance bitstream database. * Study additional testing methodologies to fulfil the needs for VVC conformance testing. | J. Boyce and W. Wan (co-chairs), E. Alshina, F. Bossen, I. Moccagatta, K. Kawamura, D. Rusanovskyy, K. Sühring, X. Xu, T. Zhou (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-2.0 software version and the reference configuration encodings according to JVET-W2017 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 |
| **Coding of HDR/WCG material (AHG7)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study and evaluate available HDR/WCG test content. * Study objective metrics for quality assessment of HDR/WCG material, including investigation of the correlation between subjective and objective results. * Compare the performance of the VTM and HM for HDR/WCG content. * Study the luma/chroma bit allocation in the HDR CTC, especially for HLG content. * Contribute to activities in merging HDR related CTC documents, in ccordination with AHG3. * Study additional aspects of coding HDR/WCG content. | A. Segall (chair), E. François, W. Husak, S. Iwamura, D. Rusanovskyy (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 conformance testing for operation range extensions in coordination with AHG5. * Identify suitable test material for testing of high bit depth, high bit rate, and high frame rate coding in coordination with AHG4 and AHG7. * Study VVC entropy decoding throughput and latency in the cases of high bit depth, high bit rate, and high frame rate coding. | A. Browne and 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. * 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. | J. Boyce, S. McCarthy (co-chairs), C. Fogg, P. de Lagrange, A. Luthra, 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 and 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 JVET-W2023. * Solicit input contributions on NN-based video coding technologies. * Refine the test conditions for NN-based video coding, 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. * Review the outcome of the expert viewing conducted at the meeting, refine the methodology, and prepare viewing for the next meeting. * Generate and distribute anchor encodings and develop improvements of the JVET-W2016 common test conditions for NNVC technology. * 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) | N |
| **Enhanced compression beyond VVC capability (AHG12)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Solicit and study non-neural-network video coding tools with enhanced compression capabilities beyond VVC. * Discuss and propose refinements to the ECM2 algorithm description JVET-W2025. * Study the performance and complexity tradeoff of these video coding tools. * Coordinate with AHG6 on ECM software development. * Refine test conditions in JVET-W2017, 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-W2024 in coordination with the EE coordinators. * Coordinate with AHG11 to study the interaction with neural network-based coding tools. | M. Karczewicz, Y. Ye and L. Zhang (co-chairs), B. Bross, X. Li, K. Naser, H. Yang (vice chairs) | N |

It was confirmed that the rules which can be found in document ISO/IEC JTC 1/‌SC 29/‌AG 2 N010 “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 allow 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 45](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/2/List%20of%20AHGs%20established%20at%20the%202nd%20WG%205%20meeting)) 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 45](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/2/List%20of%20AHGs%20established%20at%20the%202nd%20WG%205%20meeting), as noted in section 9.

[JVET-X1000](https://jvet-experts.org/doc_end_user/current_document.php?id=11024) Meeting Report of the 23rd JVET Meeting [G. J. Sullivan, J.-R. Ohm] [WG 5 N 64] (2021-08-13)

Initial versions of the meeting notes (d0 … d8) were made available on a daily basis during the meeting.

Remains valid – not updated: [JCTVC-H1001](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=5095) HEVC software guidelines [K. Sühring, D. Flynn, F. Bossen (software coordinators)]

Remains valid – not updated: [JVET-V1002](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 15 [C. Rosewarne (primary editor), K. Sharman, R. Sjöberg, G. J. Sullivan (co-editors)]

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-X1004](https://jvet-experts.org/doc_end_user/current_document.php?id=11025) 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] (2021-09-30, near next meeting)

New: X1005 Draft of new level for HEVC

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 content was included in a new edition issued as a DIS of a new edition [WG 5 N 66](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/5/Text%20of%20ISO-IEC%20DIS%2014496-10%3A202X%20Advanced%20Video%20Coding%20(10th%20edition)) (due 2021-07-30). Twin text was already consented in ITU-T from the last meeting

Editors of the DIS will be J. Boyce, S. McCarthy, G. Sullivan, and Y.-K. Wang.

A DoCR ([WG 5 N 65](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/5/Disposition%20of%20comments%20on%20ISO-IEC%2014496-10%3A2020%20CDAM%201%20Additional%20SEI%20messages)) of the NB comments received in [m57147](https://dms.mpeg.expert/doc_end_user/current_document.php?id=79339&id_meeting=187) from the ISO/IEC JTC 1/SC 29 CDAM ballot was reviewed Thursday 15 July in session 22.

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-T1016 through JVET-T1019

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]

Remains valid for HM – not updated: [JVET-U1100](https://jvet-experts.org/doc_end_user/current_document.php?id=10675) Common Test Conditions for HM Video Coding Experiments [K. Sühring, K. Sharman] (2021-02-01)

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-W2002](https://jvet-experts.org/doc_end_user/current_document.php?id=11026) Algorithm description for Versatile Video Coding and Test Model 14 (VTM 14) [A. Browne, J. Chen, Y. Ye, S. Kim] [WG 5 N 71] (2021-09-30, near next meeting)

Changes to encoding algorithms, include the tools for operation range extensions.

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-X2005](https://jvet-experts.org/doc_end_user/current_document.php?id=11027) VVC operation range extensions (Draft 5) [F. Bossen, B. Bross, T. Ikai, D. Rusanovskyy, G. J. Sullivan, Y.-K. Wang] (2021-07-30)

Was integrated into a DIS of a new edition issued as [WG 5 N 70](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/5/Text%20of%20ISO-IEC%20DIS%2023090-3%3A202X%20Versatile%20Video%20Coding%20(2nd%20edition)) (due 2021-07-30)

Editors are F. Bossen, B. Bross, T. Ikai, D. Rusanovskyy, G. Sullivan, Y.-K. Wang.

A DoCR ([WG 5 N 69](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/5/Disposition%20of%20comments%20on%20ISO-IEC%2023090-3%3A2021%20CDAM%201%20Operation%20range%20extensions)) of the NB comments received in [m57149](https://dms.mpeg.expert/doc_end_user/current_document.php?id=79341&id_meeting=187) from the ISO/IEC JTC 1/SC 29 CDAM ballot was reviewed Thursday 15 July in session 22.

[JVET-X2006](https://jvet-experts.org/doc_end_user/current_document.php?id=11028) Additional SEI messages for VSEI (Draft 5) [J. Boyce, G. J. Sullivan, Y.-K. Wang] (2021-07-30)

Was integrated into a DIS of a new edition issued as [WG 5 N 68](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/5/Text%20of%20ISO-IEC%20DIS%2023002-7%3A202X%20Versatile%20supplemental%20enhancement%20information%20messages%20for%20coded%20video%20bitstreams%20(2nd%20edition)) (due 2021-07-30)

Editors are J. Boyce, G. Sullivan, Y.-K. Wang.

A DoCR ([WG 5 N 67](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/5/Disposition%20of%20comments%20on%20ISO-IEC%2023002-7%3A2021%20CDAM%201%20Additional%20SEI%20messages)) of the NB comments received in [m57148](https://dms.mpeg.expert/doc_end_user/current_document.php?id=79340&id_meeting=187) from the ISO/IEC JTC 1/SC 29 CDAM ballot was reviewed Thursday 15 July in session 22.

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-U2008](https://jvet-experts.org/doc_end_user/current_document.php?id=10679) Conformance testing for versatile video coding (Draft 6) [J. Boyce, E. Alshina, F. Bossen, K. Kawamura, I. Moccagatta, W. Wan] (2021-03-31)

Remains valid – not updated: [JVET-U2009](https://jvet-experts.org/doc_end_user/current_document.php?id=10680) Reference software for versatile video coding (Draft 2) [F. Bossen, K. Sühring, X. Li] (2021-03-31)

Remains valid – not updated: [JVET-T2010](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, J. Boyce, X. Li, V. Seregin, K. Sühring]

Remains valid – not updated: [JVET-V2011](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]

Remains valid – not updated: [JVET-U2012](https://jvet-experts.org/doc_end_user/current_document.php?id=10681) JVET common test conditions and evaluation procedures for 360° video [Y. He, J. Boyce, K. Choi, J.-L. Lin] (2021-03-31)

Remains valid – not updated: [JVET-T2013](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10546) VTM common test conditions and software reference configurations for non-4:2:0 colour formats [Y.-H. Chao, Y.-C. Sun, J. Xu, X. Xu]

Remains valid – not updated: [JVET-Q2014](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9683) JVET common test conditions and software reference configurations for lossless, near lossless, and mixed lossy/lossless coding [T.-C. Ma, A. Nalci, T. Nguyen]

Remains valid – not updated: [JVET-Q2015](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9684) JVET functionality confirmation test conditions for reference picture resampling [J. Luo, V. Seregin]

[JVET-W2016](https://jvet-experts.org/doc_end_user/current_document.php?id=11029) Common Test Conditions and evaluation procedures for neural network-based video coding technology [S. Liu, A. Segall, E. Alshina, R.-L. Liao] (2021-07-30)

Updates to template, and clarifications

[JVET-W2017](https://jvet-experts.org/doc_end_user/current_document.php?id=11030) Common Test Conditions and evaluation procedures for enhanced compression tool testing [M. Karczewicz and Y. Ye] (2021-07-30)

This is to include clarifications and editorial refinements.

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-W2019](https://jvet-experts.org/doc_end_user/current_document.php?id=11031) White paper on Versatile Video Coding [B. Bross and G. Sullivan] [SC 29 AG 3 N 35] (2021-09-14)

[JVET-W2020](https://jvet-experts.org/doc_end_user/current_document.php?id=11032) VVC verification test report for high dynamic range video content [V. Baroncini, M. Wien] [WG 5 N 72] (2021-08-31)

It was pointed out that it might be useful to release an integrated report of SDR, HDR, and 360° test results in the future.

No output: JVET-W2021

Number retained for future purposes of planning possible additional verification testing.

[JVET-W2022](https://jvet-experts.org/doc_end_user/current_document.php?id=11021) Core Experiment on Film Grain Synthesis [S. McCarthy, M. Radosavljević, J. Shingala] [WG 5 N 74] (2021-07-30)

An initial draft was reviewed and approved. It is suggested to determine the runtime of the proposals relative to each other rather than taking VTM runtime as reference. VTM 13 is used, as there would be no difference in performance compared to VTM 14.

[JVET-W2023](https://jvet-experts.org/doc_end_user/current_document.php?id=11022) Exploration Experiment on Neural Network-based Video Coding (EE1) [E. Alshina, S. Liu, W. Chen, F. Galpin, Y. Li, Z. Ma, H. Wang] [WG 5 N 75] (2021-07-30)

An initial draft of this document was reviewed and approved.

Three categories of experiments were planned: Enhancement filters (loop and post), super resolution, and intra prediction.

It was suggested to use negative values for bit rate saving in the diagrams.

[JVET-W2024](https://jvet-experts.org/doc_end_user/current_document.php?id=11020) Exploration Experiment on Enhanced Compression beyond VVC capability (EE2) [ V. Seregin, J. Chen, S. Esenlik, F. Le Léannec, L. Li, J. Ström, M. Winken, X. Xiu, K. Zhang] [WG 5 N 76] (2021-08-06)

An initial draft was reviewed and approved. It was suggested to also test UQT standalone additionally to its combination with ABT.

[JVET-W2025](https://jvet-experts.org/doc_end_user/current_document.php?id=11033) Algorithm description of Enhanced Compression Model 2 (ECM 2) [M. Coban, F. Le Léannec, J. Ström] [WG 5 N 73] (2021-08-31)

See notes under JVET-W0102.

[JVET-W2026](https://jvet-experts.org/doc_end_user/current_document.php?id=11023) Conformance testing for VVC operation range extensions (draft 1) [D. Rusanovskyy, J.-H. Jhu, I. Moccagatta, M. Sarwer, Y. Yu] (2021-08-31)

An initial version of this document was reviewed and approved.

# 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. 6 – Fri. 15 October 2021, 24th meeting under ISO/IEC JTC 1/‌SC 29 auspices as a virtual meeting
* During January 2022, 25th meeting under ITU-T SG16 auspices, location t.b.d.
* Fri. 22 – Fri. 29 April 2022, 26th meeting under ISO/IEC JTC 1/‌SC 29 auspices in Alpbach, AT
* Fri. 15 – Fri. 22 July 2022, 27th meeting under ISO/IEC JTC 1/‌SC 29 auspices in Cologne, DE.
* During October 2022, 28th meeting under ITU-T SG16 auspices in Geneva, CH.
* During January 2023, 29th meeting under ISO/IEC JTC 1/‌SC 29 auspices, location t.b.d.
* During April 2023, 30th meeting under ISO/IEC JTC 1/‌SC 29 auspices, location t.b.d.
* During July 2023, 31st meeting under ITU-T SG16 auspices in Geneva, CH.

The agreed document deadline for the 24th JVET meeting was planned to be Wednesday 30 Sept. 2021. [Ed. (GJS): Mismatch: Either Wednesday 29 Sept. or Thursday 30 Sept.]

Giacomo Baroncini and Vittorio Baroncini were thanked for conducting, and Mathias Wien was thanked for coordinating the VVC verification test for HDR video. It is greatly appreciated that this testing was successfully completed ahead of time relative to the original plans, and despite requiring extraordinary efforts due the complications caused by the pandemic situation.

InterDigital was thanked for providing financial support for the VVC verification tests.

Mathias Wien was thanked for planning, organizing and conducting the remote expert viewing related to the exploration experiment on neural network-based video compression. It is emphasized that this effort was extremely helpful to better understand the benefits of this new technology in terms of subjective quality improvement.

The 23rd JVET meeting was closed at approximately 0028 hours UTC on Saturday 17 July 2021.

# Annex A to JVET report: List of documents

# Annex B to JVET report: List of meeting participants

The participants of the twenty-third meeting of the JVET, according to the participation records from the Zoom teleconferencing tool used for the meting sessions (approximately 1 people in total, not including those who attended only the joint sessions with other groups), were as follows:

1. XX YY (ZZZ)

# Annex C to JVET report: Recommendations of the 4th 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)