|  |  |
| --- | --- |
| **Joint Video Experts Team (JVET)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29**  20th Meeting, by teleconference, 7 – 16 October 2020 | Document: JVET-T\_Notes\_dA |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Meeting Report of the 20th Meeting of the Joint Video Experts Team (JVET), by teleconference, 7–16 October 2020** | | |
| *Status:* | Report document from the chairs of JVET | | |
| *Purpose:* | Report | | |
| *Author(s) or Contact(s):* | **Gary Sullivan** Microsoft Corp. 1 Microsoft Way Redmond, WA 98052 USA  **Jens-Rainer Ohm** Institute of Communication Engineering RWTH Aachen Melatener Straße 23 D-52074 Aachen | Tel: Email:  Tel: Email: | +1 425 703 5308 [garysull@microsoft.com](mailto:garysull@microsoft.com)  +49 241 80 27671 [ohm@ient.rwth-aachen.de](mailto:ohm@ient.rwth-aachen.de) |
| *Source:* | Chairs of JVET | | |

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

# Summary

The Joint Video Experts Team (JVET) of ITU-T WP3/16 and ISO/IEC JTC 1/ SC 29 held its twentieth meeting during 7–16 October 2020 as an online-only meeting. It had previously been planned to be in Rennes, FR, but this plan was changed due to the difficulties resulting from the COVID-19 pandemic. The JVET meeting was held under the chairmanship of Dr Gary Sullivan (Microsoft/USA) and 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 2.13 of this document. 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 a new standard. The name Versatile Video Coding (VVC) was chosen in April 2018 as the informal nickname for the new standard.

The JVET meeting began at approximately 0500 hours UTC on Wednesday 7 October 2020. Meeting sessions were held on all days except the weekend days of Saturday and Sunday 10 and 11 October 2020, until the meeting was closed at approximately XXXX hours UTC on Friday 16 October 2020. Approximately XXX people attended the JVET meeting, and approximately XXX input documents (not counting crosschecks), and 18 AHG reports were discussed. The latter included reports of 5 AHGs which had originally been established by the Joint Collaborative Team on Video Coding (JCT-VC), after the parent bodies had negotiated to continue those work items within JVET as a single joint team. The meeting took place in a collocated fashion with a meeting of various SG16 Working Groups – where WG 5 is representing the Joint Video Coding Team(s) and their activities from the SC 29 parent body. 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, software and conformance packages. As a primary goal, the JVET meeting reviewed the work that was performed in the interim period since the nineteenth JVET meeting in producing a tenth draft of the VVC standard (which was approved by the parent bodies as version 1 of the VVC standard), and the tenth version of the associated VVC test model (VTM). Further important goals were reviewing technical input on novel aspects of video coding technology, and plan next steps for investigation of candidate technology towards further standard development.

The JVET produced X output documents from the meeting (Update):

* JVET-S2001 Versatile Video Coding specification text (Draft 10)
* JVET-S2002 Algorithm description for Versatile Video Coding and Test Model 10 (VTM 10)
* JVET-S2004 Algorithm descriptions of projection format conversion and video quality metrics in 360Lib (Version 11)
* JVET-S2005 Methodology and reporting template for coding tool testing
* JVET-S2007 Supplemental enhancement information messages for coded video bitstreams (Draft 5)
* JVET-S2008 Conformance testing for versatile video coding (Draft 4)
* JVET-S2009 VVC verification test plan (Draft 3)
* JVET-S2011 JVET common test conditions and evaluation procedures for HDR/WCG video
* JVET-S2016 Summary information on BD-rate experiment evaluation practices.

For the organization and planning of its future work, the JVET established X “ad hoc groups” (AHGs) to progress the work on particular subject areas. At this meeting, no Core Experiments (CE) were defined. The next four JVET meetings were planned for 8–15 January 2021 under SC29 auspices in Capetown, ZA, during 21–28 April 2021 under ITU-T SG16 auspices in Geneva, CH, during 9–16 July 2021 under SC29 auspices in Prague, CZ, and during … .

The document distribution site <http://phenix.int-evry.fr/jvet/> was used for distribution of all documents.

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 twentieth meeting 7 – 16 October 2020 as an online-only meeting, using Zoom teleconferencing tools. The JVET meeting was held under the chairmanship of Dr Gary Sullivan (Microsoft/USA) and 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.

## 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/2020_10_T_Virtual/>.

## Primary goals

As a primary goal, the JVET meeting reviewed the work that was performed in the interim period since the nineteenth JVET meeting in producing a tenth draft of the VVC standard (which was approved by the parent bodies as version 1 of the VVC standard), and the tenth version of the associated VVC test model (VTM). Further important goals were reviewing 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 documents of the JVET meeting are listed in Annex A of this report. The documents can be found at <http://phenix.int-evry.fr/jvet/>.

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 local time at the meeting facility.

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 the VTM software but have no normative effect are marked by the string “Decision (SW):”.
* Decisions that fix a “bug” in the VTM description (an error, oversight, or messiness) or in the software are marked by the string “Decision (BF):”.
* Decisions that are merely editorial without effect on the technical content of the 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 chairs. 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.

### Late and incomplete document considerations

The formal deadline for registering and uploading non-administrative contributions had been announced as Wednesday, 30 September 2020. Any documents uploaded after 1159 hours Paris/Geneva time on Thursday 1 October 2020 were considered “officially late”, giving a grace period of 12 hours to accommodate those living in different time zones of the world. The deadline does not apply to AHG reports, and other such reports which can only be produced after the availability of other input documents.

All contribution documents with registration numbers higher than JVET-T0093 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. Also, all cross-check reports were uploaded late.

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-T0XXX (a proposal on …), uploaded XX-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. For example, some of them were proposing combinations or simplifications of other proposals.

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

* JVET-T0XXX (a document on …), uploaded XX-XX.
* …

All cross-verification reports at this meeting were registered late, and even those registered before the deadline were all 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 (X) contribution registrations were later cancelled, withdrawn, never provided, were cross-checks of a withdrawn contribution, or were registered in error: JVET-T0083, ...

“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 applied to the initial uploads of documents JVET-T0XXX, … .

Contributions that had significant problems with uploaded versions included the following:

* JVET-T0XXX (…)
* …

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 was 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-S2000, the Versatile Video Coding specification text (Draft 10) JVET-S2001, the Algorithm description for Versatile Video Coding and Test Model 10 (VTM 10) JVET-S2002, the Algorithm descriptions of projection format conversion and video quality metrics in 360Lib (Version 11) JVET-S2004, the Methodology and reporting template for coding tool testing JVET-S2005, the Supplemental enhancement information messages for coded video bitstreams (Draft 5) JVET-S2007, the Conformance testing for VVC (Draft 4) JVET-S2008, the VVC verification test plan (Draft 3) JVET-S2009, the JVET common test conditions and nd evaluation procedures for HDR/WCG video JVET-S2011, the Working practices using objective metrics for evaluation of video coding efficiency experiments (Draft 3) JVET-S2016, and the Technologies under consideration for VSEI JVET-S2017 had been completed and were approved. The software implementations of VTM (versions 10.0 and 10.1) and 360lib (version 11.0) was also approved. Furthermore, the former JCT-VC output documents from the 40th JCT-VC meeting, particularly the meeting report JCTVC-AN1000, the High Efficiency Video Coding (HEVC) Test Model 16 (HM 16) Encoder Description Update 14 JCTVC-AN1002, the Errata report items for HEVC, AVC, Video CICP, and CP usage TR JCTVC-AN1004, the Annotated regions and shutter interval information SEI messages for AVC (Draft 1) JCTVC-AN1006, and the Usage of video signal type code points (Draft 1 for version 3) JCTVC-AN1008, were approved.

The group was initially asked to review the meeting report(s) of the previous meeting for finalization. The meeting report was later approved with a minor modification of including a missing output document in a list.

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 11 (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 on the WG11 host’s website. Access to the teleconference sessions of the main JVET meeting was controlled with a password that is distributed to the registered participants; this should help overloading the teleconferencing tool.

The following rules were initially set up for the Zoom teleconference meeting:

o 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.

o Stay muted unless you have something to say. (people were muted by default when they join and would 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).

o Identify who you are and your affiliation when you begin speaking.

o Use your full name and company/organization affiliation in your joining information. We will use the participation list for attendance records.

o Turn on the chat window and watch for chair communication and side commentary there as well as by audio.

o Avoid overloading people’s internet connections, we do not plan to use video for the teleconferencing calls – only voice and screen sharing. Extensive use of screen sharing is encouraged.

## Agenda

The agenda for the meeting was as follows:

* Opening remarks and review of meeting logistics and communication practices
* ISO code of conduct reminder
* IPR policy reminder and declarations
* Contribution document allocation
* Review of results of the previous meeting
* Reports of *ad hoc* group (AHG) activities
* Consideration of contributions on high-level syntax
* Consideration of contributions and communications on project guidance
* Consideration of video coding technology contributions
* Consideration of information contributions
* Coordination of visual quality testing
* Coordination activities with other working groups
* Approval of output documents and associated editing periods
* Future planning: Determination of next steps, discussion of working methods, communication practices, establishment of coordinated experiments (if any), establishment of AHGs, meeting planning, other planning issues
* Other business as appropriate for consideration

The plans for the times of meeting sessions were established as follows, in UTC (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
* [“overday” break – nearly 10 hours]
* 1900-2100 1st “evening” session [break after 2 hours]
* 2120-2320 2nd “evening” session
  1. ***ISO Code of Conduct reminder***

Participants were reminded of the ISO Code of Conduct, found at

<https://www.iso.org/publication/PUB100397.html>.

This includes 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)
* <http://www.itscj.ipsj.or.jp/sc29/29w7proc.htm> (JTC 1/‌SC 29 Procedures)

It is noted that the ITU TSB director’s AHG on IPR had issued a clarification of the IPR reporting process for ITU-T standards, as follows, per SG 16 TD 327 (GEN/16):

“TSB has reported to the TSB Director’s IPR Ad Hoc Group that they are receiving Patent Statement and Licensing Declaration forms regarding technology submitted in Contributions that may not yet be incorporated in a draft new or revised Recommendation. The IPR Ad Hoc Group observes that, while disclosure of patent information is strongly encouraged as early as possible, the premature submission of Patent Statement and Licensing Declaration forms is not an appropriate tool for such purpose.

In cases where a contributor wishes to disclose patents related to technology in Contributions, this can be done in the Contributions themselves, or informed verbally or otherwise in written form to the technical group (e.g. a Rapporteur’s group), disclosure which should then be duly noted in the meeting report for future reference and record keeping.

It should be noted that the TSB may not be able to meaningfully classify Patent Statement and Licensing Declaration forms for technology in Contributions, since sometimes there are no means to identify the exact work item to which the disclosure applies, or there is no way to ascertain whether the proposal in a Contribution would be adopted into a draft Recommendation.

Therefore, patent holders should submit the Patent Statement and Licensing Declaration form at the time the patent holder believes that the patent is essential to the implementation of a draft or approved Recommendation.”

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 software implementation package uses 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.

## Communication practices

The documents for the meeting can be found at <http://phenix.int-evry.fr/jvet/>.

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 1245.

+JCT-VC (1294). Future discussions should be on the JVET reflector rather than the JCT-VC reflector (or JVT reflector).

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)
* **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.
* **DMVR**: Decoder-side 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).
* **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
* **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 – the software codebase for future 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 (WG 11, the parent body working group 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.
* **TMVP**: Temporal motion vector prediction.
* **TS**: Transform skip.
* **TSRC**: Transform skip residual coding.
* **TT**: Ternary tree.
* **UCBDS**: Unrestricted center-biased diamond search.
* **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 7 October at 0500 UTC (chaired by GJS and JRO) were as follows.

* Merger of JVET
* Timing and organization of online meetings, calendar
* VVC/VSEI approval and publication status:   
  "H.266" | ISO/IEC 23090-3 for VVC and H.274 | ISO/IEC 23002-7 for VSEI
  + H.266 and H.274 pre-published 2020-09-09
* WPOM approval and publication status
* AVC, HEVC and CICP status
* The meeting logistics, agenda, working practices, policies, and document allocation were reviewed.
  + The meeting is 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 were reviewed.
* There was somewhat less of a problem of late non-cross-check documents and no “placeholders” – (see section 2.4.2).
* The primary goals of the meeting were
  + Errata
  + Conformance and software for VVC & VSEI
    - Requests for these were already approved in ISO/IEC, targeting CD at this meeting
  + Verification test planning
  + Potential extensions of VVC
    - Neural network tools
    - High bit rate / high bit depth
  + Potential additional SEI messages for VSEI
* Less documents than recently, not necessary to conduct sessions in parallel
* Scheduling was discussed
* 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 “afternoon” session
* [“overday” break – nearly 10 hours]
* 1900-2100 1st “evening” session [break after 2 hours]
* 2120-2320 2nd “evening” session

All 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. 7 October, 1st day
  + Session 1:
    - 0500–0550 Opening remarks, review of practices, agenda, IPR reminder
    - 0550–0700 Reports of AHGs 1–3
  + Session 2
    - 0720–0920 Reports of AHGs 4–9
  + Session 3
    - 1900–2100 Errata
  + Session 4
    - 2120–2320 SEI messages
* Thu. 8 October, 2nd day
  + Session 5
    - 0500–0700 High bit depth (incl. AHG12)
  + Session 6
    - 0720–0920 Neural network coding tools (incl. AHG11)
* Fri. 9 October, 3rd day
  + Session 7
    - 0500–0635 AHGs 10 & 13, JCT-VC AHGs 1-5
    - 0635–0700 SEI messages
  + Session 8
    - 0720–0920 Verification testing
  + Session 9
    - 2100–2330 Neural network coding tools
* Mon. 12 October, 4th day
  + Parent body meetings
    - 0500–0900 MPEG information sharing (parent body matter)
    - 1730–1845 VCEG opening session (parent body matter)
  + Session 10
    - 1900–1950 Planning, prep for joint meeting on SEI messages
  + Joint meeting (JVET/VCEG/Req/Sys)
    - 2000–2100 SEI messages
  + Session 11
    - 2120–2320 High bit depth
* Tue. 13 October, 5th day
  + Joint meeting (JVET/VCEG/Req/Video)
    - 0500–0550 Future video standardization work
  + Session 12
    - 0600–0700 Neural network coding
  + Session 13
    - 0720–0840 Neural network coding
    - 0840–0920 SEI messages
  + BoG meetings
    - 0720–0920 BoG on high bit depth coding
  + Session 14
    - 1900–2020 SEI messages
  + Joint meeting (JVET/VCEG/Sys/Vid)
    - 2030–2100 DASH-IF conformance communication
  + Session 15
    - 2120–2320 Test conditions, conformance test development, software development
* Wed. 14 October, 6th day
  + Parent body meetings
    - 0500–0600 MPEG information sharing (parent body matter)
  + Session 16
    - 0630–0920 Project development,
  + Joint meeting (JVET/VCEG/Test)
    - 1900–1950 VVC verification testing and ITU-R communication
  + Session 17
    - 2000–2100 VVC verification testing, MCTF discussion
  + BoG meetings
    - 2140–2320 BoG on High Bit Depth
    - 2140–2320 BoG on Neural network video coding
* Thu. 15 October, 7th day
  + Session 18
    - 0500–0700 BoG on High Bit Depth reporting, section 4 remainders, SEI
  + Session 19
    - 0720–0920 Section 4 remainders, SEI, other remainders
  + Session 20
    - 1900–2100 Output, AHG, CE/EE planning
  + BoG meetings
    - 2120–2320 BoG on Neural network video coding
  + Session 21
    - 2120–2320 General remainders
* Fri. 16 October, 8th day
  + Session 22
    - 0500–0700 Closing wrap-up
  + Session 23
    - 0720–0920 Closing wrap-up
  + Session 24
    - 1900–2045 Closing wrap-up
  + Parent body meetings
    - 2100–1030 MPEG information sharing (parent body matter)
  + Closing session
    - XXX–XXXX JVET closing plenary final recommendations approval

## Contribution topic overview

The approximate subject categories and quantity of contributions per category for the meeting were summarized as follows (note that the noted document counts do not include crosschecks, and may not be completely accurate):

* AHG reports (13 JVET + 5 JCT-VC) (section 3)
* Project development (section 4)
  + General (0)
  + Text development and errata reporting (6)
  + Test conditions (1)
  + Verification test (7) (1 TBP)
  + Coding studies and tools on specific use cases (0)
  + Test Material (3) (3 TBP)
  + Conformance development (1)
  + Software development (3)
  + Implementation studies (3)
  + Complexity analysis (2) (2 TBP)
  + Encoder optimization (2) (2 TBP)
  + Profile/tier/level specification (1)
* Low-level tool technology proposals (section 5) with subtopics
  + AHG12: High bit depth coding (11) (section 5.1) (BoG T. Ikai)
  + AHG10: Neural network-based technology (16) (section 5.2) (5 TBP?) (BoG A. Segall)
  + Other (0) (section 5.3)
* High-level syntax (HLS) proposals (section 6) with subtopics
  + AHG8: Layered coding and resolution adaptation (2) (section 6.1)
  + AHG9: SEI message studies and proposals (14) (section 6.2) (3 TBP)
  + Other (1) (section 6.3) (1 TBP)
* Withdrawn (1) (section 7)
* Joint meetings, plenary discussions, BoG reports, Summary of actions (section 8)
* Project planning (section 9)
* Establishment of AHGs (section 10)
* Output documents (section 11)
* Future meeting plans and concluding remarks (section 12)

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

# AHG reports (13 JVET + 5 formerly JCT-VC)

These reports were discussed Wednesday 7 October 2020 during 0550-0700 and 0720–0920 UTC (chaired by GJS & JRO), except as otherwise noted.

[JVET-T0001](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10474) JVET AHG report: Project management (AHG1) [J.-R. Ohm, G. J. Sullivan]

This document reports on the work of the JVET ad hoc group on Project Management, including an overall status report on the VVC standardization project and the progress made during the interim period since the preceding meeting.

The work of the JVET overall had proceeded well in the interim period with a moderate number of input documents 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 available at the "Phenix" site (http://phenix.int-evry.fr/jvet/) or the ITU-based JVET site (<http://wftp3.itu.int/av-arch/jvet-site/2020_06_S_Virtual/>), particularly including the following:

* The meeting report (JVET-S2000) [Posted 2020-10-07]
* Versatile Video Coding (Draft 10) (JVET-S2001) [Posted 2020-06-23, last update 2020-09-04]
* Algorithm description for Versatile Video Coding and Test Model 10 (VTM 10) (JVET-S2002) [Posted 2020-08-12]
* Algorithm descriptions of projection format conversion and video quality metrics in 360Lib (Version 11) (JVET-S2004) [Posted 2020-10-02]
* Methodology and reporting template for coding tool testing (JVET-S2005) [Posted 2020-07-20]
* Supplemental enhancement information messages for coded video bitstreams (Draft 5) (JVET-S2007) [Posted 2020-06-29, last update 2020-09-04]
* Conformance testing for Versatile Video Coding (Draft 4) (JVET-S2008) [Posted 2020-10-06]
* Verification test plan (Draft 3) (JVET-S2009) [Posted 2020-07-11, last update 2020-07-31]
* JVET common test conditions and evaluation procedures for HDR/WCG video (JVET-S2011) [Posted 2020-07-23]
* Working practices using objective metrics for evaluation of video coding efficiency experiments (Draft 3) (JVET-S2016) [Posted 2020-07-10]
* Technologies under consideration for VSEI (JVET-S2017) [Posted 2020-08-05]

As a major achievement, JVET-S2001 was submitted to the parent bodies for publication as ITU-T Rec. H.266, and ISO/IEC FDIS 23090-3. Likewise, JVET-S2007 was submitted for publication as ITU-T Rec. H.-274, and ISO/IEC FDIS 23002-7.

The thirteen *ad hoc* groups had made progress, and reports from those activities had been submitted.

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

During the interim period, meetings of AHG4 (for preparing the verification tests) and AHG11 (for discussing aspect of neural network-based video coding) were held.

Software integration of VTM 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 60 input contributions to the current meeting (not counting the AHG summary reports and crosschecks) had been registered for consideration at the meeting. No CEs had been conducted.

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 reports,
* Maintenance and minor enhancement work on the *Advanced Video Coding* (AVC) standard, associated conformance test sets and reference software.

Consequently, reports of the 5 AHGs which had been established at the 40th JCT-VC meeting were also submitted to JVET.

A preliminary basis for the document subject allocation and meeting notes for the 20th meeting had been made publicly available on the ITU-hosted ftp site <http://wftp3.itu.int/av-arch/jvet-site/2020_10_T_Virtual/>.

[JVET-T0002](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10475) JVET AHG report: Draft text and test model algorithm description editing (AHG2) [B. Bross, J. Chen, J. Boyce, S. Kim, S. Liu, Y.-K. Wang, Y. Ye]

This document reports the work of the JVET ad hoc group on draft text and test model algorithm description editing (AHG2) between the 19th meeting by teleconference, (22 June – 1 July 2020) and the 20th meeting by teleconference, (7–16 October 2020).

At the 19th JVET meeting, it was decided to simplify, clean up and fix the existing features for intra picture-prediction, inter-picture prediction, transform, CABAC engine, in-loop filter, and high-level syntax functionalities in the tenth draft of Versatile Video Coding (VVC D10) and the VVC Test Model 10 (VTM 10) encoding as well as the high-level syntax in the fifth draft of the VSEI text specification (VSEI D5).

The bitstream format, conformance, and normative decoding process for Versatile Video Coding are specified in the VVC draft 10 text specification document. The VVC Test Model 10 (VTM 10) Algorithm and Encoder Description document provides an algorithm description as well as an encoder-side description of the VVC Test Model 10, which serves as a tutorial for the algorithm and encoding model implemented in the VTM10.x software. Video usability information parameters and some supplemental enhancement information messages for coded video bitstreams are specified in the VSEI draft 5 text.

An issue tracker (<https://jvet.hhi.fraunhofer.de/trac/vvc>) was used to facilitate the reporting of errata with the VVC documents.

*JVET-S2001 VVC text specification (Draft 10)*

Seventeen versions of JVET-S2001 were published by the Editing AHG between the 19th meeting by teleconference, (22 June – 1 July 2020) and the 20th meeting by teleconference, (7–16 October 2020).

*JVET-S2002 VVC test model 10 (VTM 10) algorithm and encoder description*

One version of JVET-S2002 was published by the Editing AHG between the 19th meeting by teleconference, (22 June – 1 July 2020) and the 20th meeting by teleconference, (7–16 October 2020).

JVET-S2002 has been established based on JVET-R2002. It provides the algorithm description for majority of coding tools in VVC. In this editing period, the overall description has been improved.

*JVET-S2007 VSEI text specification (Draft 5)*

Seven versions of JVET-S2007 were published by the Editing AHG between the 19th meeting by teleconference, (22 June – 1 July 2020) and the 20th meeting by teleconference, (7–16 October 2020).

Related contributions

* Y.-K. Wang (Bytedance), G. J. Sullivan (Microsoft), “On the syntax design for extending SEI messages and VUI for HEVC, VVC and VSEI,” document JVET-T0048, 20th JVET meeting: by teleconference, 7 – 16 October 2020.
* Y.-K. Wang, Z. Deng (Bytedance), “Some errata items for VVC,” document JVET-T0055, 20th JVET meeting: by teleconference, 7–16 October 2020.

The AHG recommended to:

* Approve the edited JVET-S2001, JVET-S2002 and JVET-S2007 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 and act on them if found to be necessary.

It was commented that there are still some open issues in the bug tracker.

The pre-published ITU-T text and the FDIS input and the JVET output were reported to be aligned.

Additional fixes (roughly 15-20 tickets and some obvious pure typos) had been prepared by the editors with the intent to reflect them in the final publications of ITU-T and ISO/IEC.

An input document reflecting these was requested.

It was remarked that Ticket #1372 was particularly important and should be reviewed. It was commented that the software has the correct behaviour; this is just a matter of making sure the text is correct.

[JVET-T0003](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10476) JVET AHG report: Test model software development (AHG3) [F. Bossen, X. Li, K. Sühring]

This report summarizes the activities of the AhG3 on Test model software development that has taken place between the 19th and 20th JVET meetings.

*VTM 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 is documented at:

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

The VTM software can be found at

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

VTM 9.2 was tagged on Jul. 6, 2020.

VTM 9.3 was tagged on Jul.6, 2020. Relative to VTM 9.2 only macros of the previous cycle are removed.

VTM 10.0 was tagged Aug. 13, 2020. Changes include:

VTM 10.1 release candidate 1 was tagged on Oct. 6, 2020. It is expected that version 10.1 will be tagged shortly thereafter.

CTC Performance

The following tables show **VTM 10.0** performance over **HM 16.22**:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **All Intra Main 10** |  |  |
|  |  |  | **Over HM 16.22** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | -29,04% | -32,17% | -34,07% | 1591% | 166% |
| Class A2 | -29,29% | -23,92% | -21,06% | 2567% | 172% |
| Class B | -21,73% | -26,96% | -30,76% | 2809% | 172% |
| Class C | -22,54% | -18,95% | -22,70% | 3951% | 188% |
| Class E | -25,76% | -25,91% | -24,46% | 2318% | 165% |
| **Overall** | -25,06% | -25,37% | -26,85% | 2630% | 173% |
| Class D | -18,47% | -13,31% | -13,42% | 4447% | 180% |
| Class F | -39,33% | -39,73% | -42,22% | 5088% | 174% |
|  |  |  |  |  |  |
|  |  |  | **Random access Main 10** |  |  |
|  |  |  | **Over HM 16.22** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | -38,76% | -37,00% | -44,06% | 739% | 155% |
| Class A2 | -43,20% | -39,82% | -38,41% | 833% | 172% |
| Class B | -35,02% | -46,82% | -44,61% | 803% | 153% |
| Class C | -29,98% | -30,67% | -32,62% | 1070% | 153% |
| Class E |  |  |  |  |  |
| **Overall** | -36,06% | -39,15% | -40,06% | 859% | 157% |
| Class D | -27,88% | -26,68% | -26,12% | 1175% | 165% |
| Class F | -41,52% | -44,62% | -45,73% | 607% | 133% |
|  |  |  |  |  |  |
|  |  |  | **Low delay B Main 10** |  |  |
|  |  |  | **Over HM 16.22** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -30,81% | -37,35% | -35,42% | 738% | 142% |
| Class C | -29,13% | -22,47% | -22,45% | 890% | 151% |
| Class E | -33,36% | -40,20% | -34,18% | 364% | 120% |
| **Overall** | -30,89% | -33,11% | -30,79% | 658% | 139% |
| Class D | -26,02% | -16,68% | -15,88% | 917% | 152% |
| Class F | -42,80% | -44,45% | -44,71% | 498% | 125% |
|  |  |  |  |  |  |
|  |  |  | **Low delay P Main 10** |  |  |
|  |  |  | **Over HM 16.22** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -35,15% | -39,91% | -37,83% | 690% | 158% |
| Class C | -30,83% | -22,55% | -22,66% | 818% | 158% |
| Class E | -36,05% | -43,41% | -37,33% | 365% | 128% |
| **Overall** | -33,93% | -35,00% | -32,65% | 623% | 150% |
| Class D | -27,47% | -15,71% | -14,93% | 856% | 160% |
| Class F | -42,31% | -43,56% | -44,09% | 536% | 133% |

The following tables show **VTM 10.0** performance compared to **VTM 9.0**:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **All Intra Main 10** |  |  |
|  |  |  | **Over VTM 9.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | -0,01% | 0,01% | 0,00% | 102% | 101% |
| Class A2 | -0,01% | 0,03% | -0,01% | 102% | 99% |
| Class B | -0,01% | -0,05% | -0,04% | 101% | 98% |
| Class C | -0,02% | 0,01% | 0,01% | 101% | 98% |
| Class E | -0,03% | 0,05% | -0,01% | 103% | 99% |
| **Overall** | -0,01% | 0,00% | -0,01% | 101% | 99% |
| Class D | -0,06% | -0,03% | -0,08% | 100% | 102% |
| Class F | -0,03% | -0,04% | -0,02% | 101% | 96% |
|  |  |  |  |  |  |
|  |  |  | **Random access Main 10** |  |  |
|  |  |  | **Over VTM 9.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 | -0,02% | -0,01% | -0,01% | 102% | 101% |
| Class A2 | -0,01% | -0,01% | -0,01% | 102% | 99% |
| Class B | -0,03% | -0,04% | -0,03% | 100% | 98% |
| Class C | -0,04% | -0,04% | -0,04% | 102% | 96% |
| Class E |  |  |  |  |  |
| **Overall** | -0,03% | -0,03% | -0,03% | 101% | 98% |
| Class D | -0,12% | -0,12% | -0,12% | 99% | 98% |
| Class F | -0,07% | -0,07% | -0,07% | 102% | 96% |
|  |  |  |  |  |  |
|  |  |  | **Low delay B Main 10** |  |  |
|  |  |  | **Over VTM 9.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -0,03% | -0,03% | -0,03% | 100% | 98% |
| Class C | -0,02% | -0,02% | -0,02% | 99% | 96% |
| Class E | -0,46% | -0,47% | -0,42% | 100% | 96% |
| **Overall** | -0,13% | -0,14% | -0,12% | 99% | 97% |
| Class D | -0,06% | -0,05% | -0,05% | 99% | 94% |
| Class F | -0,20% | -0,20% | -0,20% | 100% | 96% |
|  |  |  |  |  |  |
|  |  |  | **Low delay P Main 10** |  |  |
|  |  |  | **Over VTM 9.0** |  |  |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | -0,03% | -0,03% | -0,03% | 98% | 97% |
| Class C | -0,02% | -0,01% | -0,01% | 98% | 92% |
| Class E | -0,44% | -0,44% | -0,40% | 101% | 95% |
| **Overall** | -0,13% | -0,13% | -0,12% | 99% | 95% |
| Class D | -0,05% | -0,04% | -0,04% | 99% | 94% |
| Class F | -0,20% | -0,20% | -0,19% | 101% | 98% |

Full results were attached to the AHG report as Excel files.

Issues in VTM 10.0 affecting conformance

The following issues in VTM 10.0 affect conformance:

* A bug related to the combination for deblocking, 4:2:2/4:4:4 chroma formats and ISP (ticket #1353)
* Incorrect output order of pictures (ticket #1396). In this context it should be noted that the Output Picture Log produced by VTM 10.0 doesn’t include layer ids making such issues more difficult to identify and debug.
* Missing HLS features

These issues have been partially resolved in VTM 10.1rc1. In particular the OPL format was changed to include layers ids. Some HLS features are still missing (see sections below) and there are several tickets related to multilayer coding that are still open.

Status of implementation of proposals of previous JVET meetings

Actions of the Q, R, and S meetings that had not yet been marked as resolved were listed in the AHG report.

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>

Please file all issues related to the VVC reference software into the 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.

The AHG recommended to:

* Continue to develop the VTM reference software
* Improve documentation, especially the software manual
* Resolve any normative issues resulting from the large number of integrations in the most recent development cycle
* Encourage people to test VTM software more extensively outside of common test conditions.
* Encourage people to report all (potential) bugs that they are finding.
* Encourage people to submit bitstreams/test cases that trigger bugs in VTM.
* Encourage people to submit non-normative changes that reduce encoder run time without significantly sacrificing compression performance
* 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

It was commented that we need people to check the issues that are marked as open to confirm whether any action is really needed and to make the software coordinators aware of the result. The software coordinators also emphasized the need for those who have contributed patches to keep an eye out for email and make sure the issues of its integration have been fully resolved.

End of session #1.

[JVET-T0004](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10477) JVET AHG report: Test material and visual assessment (AHG4) [V. Baroncini, T. Suzuki, M. Wien, A. Norkin, A. Segall, Y. Ye]

This report summarizes the activities of the AhG4 on Test material and visual assessment that has taken place between the 18th and 19th JVET meetings. Two online AHG meetings were held related to the preparation of the VVC verification tests in the SDR HD and 360° video categories. The inputs to these AHG meetings and the AHG meeting reports are provided in the input documents JVET-T0043, JVET-T0044, JVET-T0045, JVET-T0046, respectively. The results of the visual tests in the SDR UHD category of the VVC verification tests are reported in input document JVET-T0097.

***Verification test***

For the SDR UHD Random Access category, formal visual tests have been conducted at GBTech, Rome, IT, and RWTH Aachen University, Aachen, DE. The testing procedure was adapted to the restrictions due to the COVID-19 pandemic and therefore took more time than initially planned. The results are provided in input contribution JVET-T0097.

An AHG meeting for SDR HD was held on 2020-09-03, reviewing the outcome of a dedicated experts viewing session as reported in JVET-T0043. The AHG meeting report is provided in JVET-T0045. As a major outcome of this meeting, it was suggested to include both, SDR HD Random Access (RA) and SDR HD Low Delay (LD) configurations in the next update of the Verification Test Plan. The identification and collection of SDR HD test sequences for both categories is in progress.

An AHG meeting for 360° video was held on 2020-09-04, reviewing the outcome of an experts viewing session for rate point selection as reported in JVET-T0044. The AHG meeting report is provided in JVET-T0046. Potential improvements to the 360Lib software were identified for the projection formats GCMP and PCMP.

For the HDR category, the Zoom sessions were held between the verification test coordinators and A. Segall on display calibration and process planning. The identified display settings are collected in JVET-T0106. Due to the general delays for the pandemic, the effort of rate point selection has not been completed by the beginning of this meeting.

***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).

**Related contributions**

The following related contributions were submitted.

* JVET-T0043 “Status Report on SDR HD Low Delay Verification Test Preparation” [M. Wien (RWTH), V. Baroncini (VABTech ltd)]
* JVET-T0044 “Status Report on 360 Video Verification Test Preparation” [M. Wien (RWTH), V. Baroncini (VABTech), T. Suzuki (Sony)]
* JVET-T0045 “AHG4: Agenda and report of the AHG meeting on the SDR HD Verification Tests on 2020-09-03” [M. Wien (RWTH), V. Baroncini (VABTech ltd)]
* JVET-T0046 “AHG4: Agenda and report of the AHG meeting on the 360 Video Verification Tests on 2020-09-04” [M. Wien (RWTH), V. Baroncini (VABTech ltd), Y. Ye (Alibaba)]
* JVET-T0060 “AHG4: new video conference sequences for HD verification testing” [S. Xu, R.-L. Liao, J. Chen, Y. Ye (Alibaba)]
* JVET-T0082 “New HLG sequences for high bit-depth video coding” [K. Kondo, M. Ikeda, A. Browne (Sony)]
* JVET-T0097 “Report on VVC compression performance verification testing in the SDR UHD Random Access Category” [M. Wien (RWTH), V. Baroncini (VABTech)]
* JVET-T0103 “Information on and analysis of the VVC encoders in the SDR UHD verification test” [C. Helmrich, B. Bross, J. Pfaff, H. Schwarz, D. Marpe, T. Wiegand (HHI)]
* JVET-T0106 “AHG4/AHG7: Report on Display Settings for HDR/WCG Verification Tests” [A. Segall, M. Wien, V. Baroncini]

These contributions should be reviewed.

The AHG recommended:

* To review the input contributions related to the verification test preparation at an early stage of the meeting.
* To continue to discuss and to update the non-finalized categories of the verification test plan
* To create a BoG to progress the update of the verification test plan.
* 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.

[JVET-T0005](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10478) JVET AHG report: Conformance testing (AHG5) [J. Boyce, W. Wan, E. Alshina, F. Bossen, I. Moccagatta, K. Kawamura, K. Sühring, X. Xu]

This document summarizes the activity of AHG5: “Conformance testing” between the 19th Meeting (teleconference, 22 June – 1 July 2020) and the 20th Meeting (teleconference, 7–16 Oct. 2020).

*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 v1, 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
* 23rd meeting July 2021: Final conformance specification

*Status on bitstream submission*

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

* 102 bitstream categories have been identified
* Volunteers have been identified for all categories
* 206 bitstreams in 77 bitstream categories have been provided for VTM 10.0
* 25 bitstream categories have no VTM 10.0 bitstreams
* 12 bitstream categories have no provided bitstreams (for any VTM version)

*Activities and Discussion*

The AHG activities are on schedule with the preliminary timeline shown in section 2. It is expected to output CD of the Conformance specification from this meeting.

Output document JVET-S2008 “Conformance testing for versatile video coding (Draft 4)” was published on 5 October 2020.

Because CD is planned to be issued from this October meeting, all bitstream volunteers had been requested to update their bitstreams using VTM 10.0. At least one bitstream has been provided in 77 of the 102 identified categories for VTM 10.0. Bitstream submitters were also requested to provide text descriptions of the bitstreams for inclusion in the conformance specification on a shared document, but relatively few have been updated.

A modification was made to the VTM master branch to include layer ID for each picture in the .opl files such as to distinguish between pictures within an access unit. A new version of the VTM is expected to include the opl change, as well as other changes to multilayer coding, missing HLS features, and a bug fix to 4:2:2/4:4:4 deblocking. Bitstreams that have already been provided in those categories may need to be updated, or the submitters are waiting for the availability of the new VTM to provide initial bitstreams.

The change to the .opl format affects bitstream packages that have already been submitted. It is expected that bitstream packages can be updated to the new format using a script and individual packages do not need to be resubmitted.

It was suggested to add bitstreams with 9-bit depth.

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.

There was one related contribution, discussing gaps in the existing conformance bitstream suite.

* [JVET-T0100](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10497), AHG5: On gaps in conformance bitstreams [F. Bossen (Sharp)]

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

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

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

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, using the Site Manager with the following configuration is suggested:



The AHG recommends the following:

* Consider if should separate the draft Conformance specification into two output documents: the Conformance specification and instructions for generating conformance bitstreams, or should move the instructions to an Annex
* Encourage volunteers of missing conformance bitstreams to provide them quickly
* Encourage conformance bitstream providers to provide text descriptions of provided bitstreams for the Conformance specification online at <http://mpeg.expert/live/nextcloud/index.php/f/37368>
* Encourage conformance bitstream providers to provide updated bitstreams with corrected naming for those bitstreams indicated in the JVET-S2008 document as needing name changes to be consistent with the naming convention
* Discuss and refine the list of conformance bitstreams and the conformance specification, especially consider adding 9-bit depth and solicit volunteers
* Review submitted bitstreams and consider if the flexibility of the tested tool is sufficiently exercised, including consideration of input contribution JVET-T0100
* Proceed with CD issuance as output of this meeting

There was discussion of whether the conformance testing spec really needs to be normative, and whether the dataset really needs to be part of the approved standard, or could be something external to it that is referenced informatively. (The tests are far from complete in any case.)

It was asked whether to issue a TR rather than a standard (removing all “shall” and “should”).

This was initially agreed – with a long editing period – and a request to change to use a TR.

The same approach was discussed for the software.

The fact that we have called prior reference software and conformance test data normative causes a problem with basically normative dependencies on non-standard external packages. On the other hand, some users may like the normative nature as providing assurance of correctness.

It was noted that there was a software maintance problem previously for 3D-AVC, such that the living codebase was removed from the web.

There would be less naming convenience on the ITU side as H.266.1 and H.266.2.

Revisit.

[JVET-T0006](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10479) JVET AHG report: 360° video coding, software and test conditions (AHG6) [J. Boyce, Y. He, K. Choi, J.-L. Lin, Y. Ye]

The document summarizes activities on 360-degree video content conversion software development between the 19th (22 June – 1 July 2020) and the 20th (7–16 Oct. 2020) JVET meetings.

*Activities*

The 360Lib-11.0 software package released on Aug. 28, 2020 included following changes:

* Support frame packing specified by VVC GCMP SEI messages (from JVET-S0257)
* Provide configuration files for those frame packings

Additionally, at the verification testing AHG meeting held on Sept 4th, it was identified that 360Lib did not support blending along the edges of cubemap faces for the GCMP project format using the padded samples. A request was made for this AHG to add this functionality to improve subjective quality of extracted viewports containing the face edges. Subsequently the 360Lib-11.0 software had been further enhanced to support blending along cubemap face edges using the padded samples, and preliminary expert viewing indicates some subjective quality improvement. It is requested to consider integrating this blending functionality into the next release of the 360Lib software.

*Software repository and versions*

The 360Lib software is developed using a Subversion repository located at:

https://jvet.hhi.fraunhofer.de/svn/svn\_360Lib/

The released version of 360Lib-11.0 can be found at:

https://jvet.hhi.fraunhofer.de/svn/svn\_360Lib/tags/360Lib-11.0/

360Lib-11.0 testing results can be found at:

ftp.ient.rwth-aachen.de/ahg/testresults/360Lib-11.0

360Lib bug tracker

https://hevc.hhi.fraunhofer.de/trac/jem/newticket?component=360Lib

*360Lib-11.0 results*

The first table below is for the projection formats comparison using VTM-10.0 according to 360-degree video CTC (JVET-L1012). It compares padded hybrid equi-angular cubemap (PHEC) coding and padded equi-rectangular projection (PERP) coding using VTM-10.0.

**VTM-10.0 PHEC vs PERP (PERP as anchor)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **PHEC over PERP (VTM-10.0)** | | | | | |
|  | **End-to-end WS-PSNR** | | | **End-to-end S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | -11.78% | -6.81% | -7.30% | -11.71% | -6.73% | -7.24% |
| Class S2 | -5.35% | -1.83% | -1.48% | -5.34% | -1.73% | -1.41% |
| **Overall** | -9.20% | -4.81% | -4.98% | -9.16% | -4.73% | -4.91% |

The next table is for PERP coding comparison between VTM-10.0 and HM-16.16. Table 3 is to compare PHEC coding with VTM-10.0 with and CMP coding with HM-16.16.

**VTM-10.0 PERP vs HM-16.16 PERP (HM-16.16 PERP as anchor)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **VTM-10.0 PERP - Over HM-16.16 PERP** | | | | | |
|  | **End-to-end WS-PSNR** | | | **End-to-end S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | -25.62% | -33.61% | -36.36% | -25.60% | -33.63% | -36.33% |
| Class S2 | -35.07% | -32.96% | -35.62% | -35.06% | -32.99% | -35.67% |
| **Overall** | -29.40% | -33.35% | -36.06% | -29.38% | -33.37% | -36.07% |

**VTM-10.0 PHEC vs HM-16.16 CMP (HM-16.16 CMP as anchor)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **VTM-10.0 PHEC - Over HM-16.16 CMP** | | | | | |
|  | **End-to-end WS-PSNR** | | | **End-to-end S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | -30.07% | -35.71% | -38.02% | -29.96% | -35.69% | -37.98% |
| Class S2 | -37.78% | -35.87% | -38.11% | -37.78% | -35.86% | -38.14% |
| **Overall** | -33.15% | -35.77% | -38.06% | -33.09% | -35.75% | -38.05% |

The AHG recommended to continue software development of the 360Lib software package.

It was commented that there is a small improvement for PHEC in VTM 10.0 relative to VTM 9.

A new version of 360Lib was requested with the new blending functionality.

+ an output document revision to describe it

It was also suggested to move the 360Lib software to a GitLab server. This was agreed.

[JVET-T0007](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10480) 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 19th meeting by teleconference (22 June – 1 July 2020) and the 20th meeting by teleconference (7–16 October 2020).

The primary activity of the AhG was related to the mandates of studying and evaluating available HDR/WCG test content, comparing the performance of the VTM for HDR/WCG content, and coordinating with AHG4 in preparation for verification testing for HDR video content.

The AhG generated CTC anchors for the VTM according to JVET-P2011. A summary of the performance is provided below, and more detailed information may be found in Excel spreadsheets attached with the AHG report.

NOTE: It is noted that the Class H2 sequences were updated at the 19th JVET meeting. This update consisted of changing the chroma sampling location for the sequences. In the tables below, we report the performance of the VTM10 and VTM9 using the configurations provided in the VTM repository to provide a reference for participants tracking the performance. Similarly, the HM16.18 results use the VTM9 sequences. Please note that the difference in Class H2 is largely due to the sequence change.

Insert results from revised report.

The chroma position issue was discussed. An additional version of report was requested to be provided with additional data to clarify this. The test sequences had changed in preparation for the verification tests. (See the notes of the meeting report of the previous meeting.)

The AHG had significant coordination with AHG4, including sharing results on the new HDR content and expert viewing procedures.

Two contributions related to HDR video coding were noted to have been registered for the meeting.

* JVET-T0082 New HLG sequences for high bit-depth video coding [K. Kondo, M. Ikeda, and A. Browne]
* JVET-T0106 AHG4/AHG7: Report on Display Settings for HDR/WCG Verification Tests [A. Segall, M. Wien and V. Baroncini]

The AHG recommended to review the relevant input contributions

[JVET-T0008](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10481) JVET AHG report: Layered coding and resolution adaptivity (AHG8) [S. Wenger, A. Segall, M. M. Hannuksela, Hendry, S. McCarthy, Y.-C. Sun, P. Topiwala, Y.-K. Wang]

This document summarizes the activity of AHG08: Layered coding and resolution adaptivity, between the 19th JVET meeting (22 June – 1 July 2020) and the 20th JVET meeting by teleconference (7–16 October 2020).

No email traffic on the JVET reflector was observed that was labelled as belonging to AHG8.

Four contributions related to the AHG are provided as input to the current meeting. The contributions are as follows:

* JVET-T0049 Composite Picture Information SEI Message [Hendry, H. Jang, S. Kim, J. Lim (LGE)]
* JVET-T0065 Conformance Points for Multilayer 8K [F. Bossen (Sharp)]
* JVET-T0070 AHG9/AHG8: Scalability dimension SEI message [Y.-K. Wang, L. Zhang, K. Zhang, Z. Deng (Bytedance)]
* JVET-T0080 AHG9: Multilayered OLS decoded picture assembling SEI message [E. Thomas (TNO)]

The AHG recommended reviewing all contributions related to layered coding and resolution adaptivity.

It was commented that, with the amount of activity that was recently devoted to getting these features into VVC, there should be more activity investigating it.

[JVET-T0009](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10482) JVET AHG report: SEI message studies (AHG9) [S. McCarthy, J. Boyce, P. de Lagrange, A. Luthra, A. Tourapis, Y.-K. Wang, S. Wenger]

This AHG report summarizes the activity of AHG9: SEI message studies between the 19th JVET Meeting (teleconference, 22 June – 1 July 2020) and the 20h JVET Meeting (teleconference, 7–16 Oct. 2020).

The following contributions were identified as related to AHG9. It was noted that some of the contributions were not identified as AHG9 in their titles.

* [JVET-T0048](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10427) On the syntax design for extending SEI messages and VUI for HEVC, VVC and VSEI [Y.-K. Wang (Bytedance), G. J. Sullivan (Microsoft)]
* [JVET-T0049](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10428) Composite Picture Information SEI Message [Hendry, H. Jang, S. Kim, J. Lim (LGE)]
* [JVET-T0050](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10429) AHG9: HEVC Annotated Regions SEI message software update for HM [P. Guruva reddiar, J. Boyce (Intel)]
* [JVET-T0051](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10430) AHG9: Object tracking information file format for use with Annotated Regions SEI messages [P. Guruva reddiar, J. Boyce (Intel)]
* [JVET-T0052](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10431) AHG9: AVC Annotated Regions SEI message software for JM [P. Guruva reddiar, J. Boyce (Intel)]
* [JVET-T0053](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10432) AHG9: VVC and VSEI Annotated Regions SEI message [J. Boyce, P. Guruva reddiar (Intel)]
* [JVET-T0054](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10433) Some errata items for HEVC and AVC [Y.-K. Wang, L. Zhang (Bytedance), A. M. Tourapis (Apple)] (Aspects #7 and #8 relate to AHG9)
* [JVET-T0055](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10434) Some errata items for VVC [Y.-K. Wang, Z. Deng (Bytedance)] (Aspects #2, #3, and #4 relate to AHG9)
* [JVET-T0056](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10435) AHG9: SEI manifest and SEI prefix indication SEI messages [Y.-K. Wang (Bytedance), T. Stockhammer (Qualcomm), A. M. Tourapis, D. Singer (Apple)]
* [JVET-T0059](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10438) AHG9: object representation SEI message [J. Chen, Y. Ye, J. Hu, K. Li, L. Hu, Y. Long (Alibaba)]
* [JVET-T0066](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10445) AHG9: IRAP only HRD SEI message [R. Skupin, Y. Sanchez, K. Suehring, T. Schierl (Fraunhofer HHI), Y.-K. Wang, L. Zhang (Bytedance)]
* [JVET-T0070](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10449) HG9/AHG8: Scalability dimension SEI message [Y.-K. Wang, L. Zhang, K. Zhang, Z. Deng (Bytedance)]
* [JVET-T0071](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10450) AHG9: SEI messages for support of cross RAP referencing based video coding [Y.-K. Wang (Bytedance)]
* [JVET-T0075](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10454) AHG9: Errata and editorial clarifications for AVC and HEVC film grain characteristics SEI message semantics [S. McCarthy, P. Yin, W. Husak, T. Lu, F. Pu, T. Chen (Dolby)]
* [JVET-T0076](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10455) AHG9: ALF control SEI message [H.-B. Teo, H.-W. Sun, C.-S. Lim (Panasonic)]
* [JVET-T0080](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10459) AHG9: Multilayered OLS decoded picture assembling SEI message [E. Thomas (TNO)]
* [JVET-T0090](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10469) AHG9: Digital signature SEI message [J. Xu, Y.-K. Wang, L. Zhang, K. Zhang (Bytedance)]
* [JVET-T0092](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10471) AHG9/AHG11: Neural network based super resolution SEI message [T. Chujoh, E. Sasaki, T. Ikai]

There were no relevant emails besides the AHG kickoff message sent to the JVET reflector during the AHG period.

The AHG recommended to review the related contributions and continue SEI messages studies.

There were relevant input contributions on both errata and SEI messages.

There was a discussion about getting SEI showcase software that’s not otherwise part of the reference software package. – For showcase purposes. A Python script was mentioned for this. It was commented that we already have some software tools that are bundled with the reference software, and if such software is maintained appropriately, we could do this with such SEI showcase software.

End of session #2.

[JVET-T0010](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10483) JVET AHG report: Encoding algorithm optimization (AHG10) [A. Duenas, A. Tourapis, S. Ikonin, A. Norkin, R. Sjöberg, J. Le Tanou, J.-M. Thiesse]

This AHG report was discussed at 0500 Friday 9 October 2020 (chaired by GJS & JRO).

This document summarizes the activities of AHG10: Encoding algorithm optimizations, between the 19th meeting (teleconference, 22 June – 1 July 2020) and the 20th meeting (teleconference, 7–16 October 2020).

No e-mail related to AHG10 activity was sent to the JVET reflector during the AHG period.

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

* [JVET-T0062](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10441)**: Extension of rate control to support random access configuration with GOP size of 32**

This contribution extends the rate control model in the VTM software, that is based on contribution JVET-M0600, to also support the GOP 32 that was presented in JVET-S0180. Compared to a fixed bit ratio rate control model, the extended rate control model presented appears to provide coding gains, in terms of Y-PSNR BD-Rate improvements of ~16% and ~19% for Class C and Class D content respectively..

* [JVET-T0063](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10442): Suggestion to upgrading random access configuration in CTC with GOP 32 configuration

This contribution provides additional information to what was presented at the last meeting in JVET-S0180. In particular, at the previous JVET teleconference in June, a GOP 32 configuration was included into the VTM. During the interim period between that meeting and this meeting, the GOP 32 configuration was tested for 4K SDR and was selected to be used as part of the verification testing for 4K SDR. In this contribution it is suggested to upgrade all other random access common test conditions, i.e. conditions applicable to HD SDR as well as HDR sequences, to also use a GOP hierarchy of 32. It is asserted that doing so would result in an improved performance for the VTM compared to its current usage. In particular, the BD-rate gains over the current SDR RA common test conditions of -3.0%, -6.3%, and -6.4% for luma, Cb, and Cr, respectively, are reported. Similarly, for the HDR RA common test conditions BD-rate gains of -7.9%, -5.2, -4.4%, and -4.5% for deltaE, psnrL, wpsnrY, and psnrY, respectively, are reported. It is further asserted that visual quality is also improved.

It is not mentioned in the contribution if similar changes should be performed to the anchor reference, i.e. for encodings performed using the HM reference software of HEVC.

* [JVET-T0064](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10443): Addition of ALF filter strength control to VTM

This contribution asserts that currently, there is no possibility to control the ALF filter strength in the VTM and that it is only possible to turn it on or off completely. This proposal suggests including a configurable ALF filter strength parameter in the VTM reference software, which allows the user to have some control of the ALF filter strength. If the provided ALF filter strength parameter is set to be less than 1, then the magnitude of the filter coefficients is decreased accordingly. Similarly, if the ALF filter strength parameter is set to be larger than 1, then the magnitude of the filter coefficients is increased. Finally, if the ALF filter strength parameter is set to be equal to 1, the current behavior of ALF remains unchanged.

The BDR performance of the proposed modification using the CTC and with the ALF filter strength parameter set to be equal to 0.75, i.e. reducing the filter strength by 25%, is as follows:

AI: 0,27%, 0,25%, 0,31% (Y,U,V)

RA: 0,33%, 0,10% , 0,16% (Y,U,V)

LDB: 0,30%, 0,11%, 0,23% (Y,U,V)

Further decreasing the filter strength to 50%, by using an ALF filter strength parameter equal to 0.5, has the following impact on BDR:

RA: 0,90%, 0,74%, 0,89% (Y,U,V)

LDB: 0,67%, 0,90%, 1,50% (Y,U,V)

It is claimed that a reduction of the ALF filter strength can, in some cases and especially for inter predictive coding, improve subjective quality.

* [JVET-T0078](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10457): GDR Software

This contribution provides source code for enabling GDR functionality in the VTM reference software and proposes its inclusion in an upcoming release. The software offers the following functionality:

* Encoding input video sequences into VVC compliant GDR bitstreams using the VTM encoder with the following supported features:
  + Flexible GDR period configured through user specified parameters,
  + Allocation of a similar number of bits per picture, allowing for delays as low as one frame interval,
  + Progressive intra refresh over a GDR period with even distribution of forced intra areas over pictures within the GDR period using the virtual boundary syntax supported in picture header,
  + Application of the necessary (encoding) constraints on coding tools to prevent "leaks" when performing a random access operation, and
  + Exact match at the recovery point (or leak-free).
* Decoding the leak-free GDR bitstream using the VTM decoder.

The VVC specification supports gradual decoding refresh (GDR). However, it is claimed that the VTM reference software does not fully support this functionality. It is claimed that it is desirable to make the source code of the GDR software available and maintained as a part of the VTM package.

* [JVET-T0099](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10496): Open optimized VVC encoder (VVenC) and decoder (VVdeC) implementations

In September 2020, two months after the first version of VVC was finalized, Fraunhofer HHI released the source code for its optimized VVC encoder (VVenC) and decoder (VVdeC) on GitHub. For VVenC, the goal was to demonstrate comparable or even subjectively better coding efficiency than the VTM reference software at much lower runtime. In addition, VVenC provides real-world encoder features like rate control and multi-threading.

Without QP adaptation for subjective optimization and 6 threads the following PSNR-based YUV BD-rates and speedup factors compared to the HM and the VTM (see figure below) are reported for four different presets:

Faster HD: -11.0%, 20x (HM), 160x (VTM) UHD: -15.9%, 36x (HM), 270x (VTM)

Fast HD: -25.5%, 16x (HM), 130x (VTM) UHD: -28.8%, 26x (HM), 200x (VTM)

Medium HD: -34.4%, 8.5x (HM), 69x (VTM) UHD: -37.4%, 14x (HM), 110x (VTM)

Slow HD: -37.9%, 2.7x (HM), 22x (VTM) UHD: -40.6%, 5x (HM), 38x (VTM)

With QP adaptation for subjective optimization and 6 threads, the following MS-SSIM-based YUV BD-rates and speedup factors compared to the HM and the VTM (see figure below) are reported for the same presets:

Faster HD: -18.4%, 20x (HM), 160x (VTM) UHD: -15.4%, 36x (HM), 270x (VTM)

Fast HD: -28.4%, 16x (HM), 130x (VTM) UHD: -29.3%, 26x (HM), 200x (VTM)

Medium HD: -37.5%, 8.5x (HM), 69x (VTM) UHD: -39.2%, 14x (HM), 110x (VTM)

Slow HD: -40.2%, 2.7x (HM), 22x (VTM) UHD: -42.1%, 5x (HM), 38x (VTM)

Chart, line chart

Description automatically generated Chart, line chart

Description automatically generated

*Quality vs Encoding speed tradeoffs achieved using VVenC and the VTM compared HM-16.22*

The Fraunhofer Versatile Video Encoder (**VVenC**) and Decoder (**VVdeC**) development was initiated to provide publicly available, fast, and efficient VVC software implementations. The VVenC and VVdeC software are based on the VTM, with optimizations including software redesign to mitigate performance bottlenecks, extensive SIMD optimizations, and basic multi-threading support to exploit parallelization.

It is claimed that the VVenC further contains improved encoder search algorithms and supports real-world encoder features, including frame-level rate control and perceptually optimized encoding, and that the bitstreams generated are fully compliant with the VVC specification and its reference decoder (VTM-10.0). Some of the features that this software encoder currently provides include the following:

* Encoder implementation with four predefined quality/speed presets.
* Optimized VVC encoder providing speedups between 22x and 270x over VTM-10.0 for HD and UHD test sequences depending on the chosen quality/speed preset.
* Perceptual optimization to improve subjective video quality.
* Frame-level rate control supporting VBR encoding.

Expert mode encoder interface available, allowing fine-grained control of the encoding process.

* [JVET-T0074](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10453): LMCS encoder algorithm for YUV4:4:4 content

This contribution proposes an automatic non-normative encoder optimization algorithm for LMCS to improve coding performance for SDR YUV4:4:4 video content. In VTM10.0, the default LMCS encoder algorithm is optimized for YUV4:2:0 content. However, YUV4:2:0 and YUV4:4:4 signals have very different characteristics. Similarly, YUV4:4:4 camera (natural) content and screen content have different characteristics. The YUV4:4:4 LMCS encoder algorithm proposed in this contribution provides the following coding performance for SDR YUV4:4:4 CTC (reference: LMCS off; test: LMCS on) under VTM10.0 in terms of BDRate and processing time:

* + For camera content (YUV4:4:4):
    - AI: { -0.32%, 0.69%, 1.04% }, Enc: 104%, Dec: 101%
    - RA: { -1.13%, 2.98%, 0.20% }, Enc: 104%, Dec: 100%
  + For screen content (YUV4:4:4):
    - AI: { -1.22%, 2.89%, 3.08% }, Enc: 104%, Dec: 101%
    - RA: { -1.64%, -0.01%, 0.16% }, Enc: 101%, Dec: 100%
* [JVET-T0103](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10500): Information on and analysis of the VVC encoders in the SDR UHD verification test (Late)

Intends to answer the following questions:

Is there any difference in perceptual quality between VTM 9 and 10 decoder compliant bitstreams?

Which objective metrics can model the outcome of the subjective evaluation best?

The AHG recommends that the related input contributions are reviewed and to further continue the study of encoding algorithm optimizations in JVET and to coordinate the review with of some of the AHG 12 related to Rice parameter derivation for high bit-depth coding JVET-T0072, JVET-T0085, JVET-T0089 and SIMD implementation of BDOF for high bit-depth coding JVET-T0086.

[JVET-T0011](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10484) JVET AHG report: Neural network-based video coding (AHG11) [A. Alshina, S. Liu, J. Pfaff, M. Wien, P. Wu, Y. Ye]

This AHG report was reviewed in session 6 at 0730 UTC Thursday 8 October (chaired by GJS & JRO).

This document summarizes the activities of AHG11: Neural network-based video coding between the 19th Meeting (22 June –1 July 2020) and the 20th meeting (7–16 Oct. 2020), both held by teleconference.

One teleconference meeting was held on July 30 with total 160 experts participated. Among the 160 participants, more than 140 stayed in the meeting for longer than 1 hour, about 120 participated the full (or nearly full) 2 hr meeting.

In this meeting, the following input document was reviewed and discussed. Valuable inputs were received and a couple of revisions of this document were generated and uploaded after the telco meeting.

JVET-T0041 Methodology and reporting template for neural network coding tool testing [S. Liu, A. Segall, E. Alshina, J. Boyce]

A collection of training datasets was reviewed in this telco meeting.

The telco meeting report (JVET-T0042) was uploaded to the JVET document web site shortly after the meeting.

Anchor generation

Following the suggestion and agreement reached in the AHG telco meeting, AHG11 should generate anchor results using five QP values (22, 27, 32, 37, 42). B-D rates should be calculated using all five QP values. It was suggested to provide Low Delay 4K anchor results and make it optional for proponents to test and report.

The anchor results were generated and announced through JVET reflector on September 22. Thanks to Tencent and Huawei for anchor generation and crosscheck.

Training data

A couple of datasets were identified to be suitable for AHG11 training process, such as the ones from xiph.org and JVET. In addition, University of Bristol would like to contribute a video training dataset, BVI-DVC, which contains 800 sequences at various spatial resolutions from 270p to 2160p and has been evaluated on ten existing network architectures for four different coding tools (https://arxiv.org/abs/2003.13552). Among these sequences, 104 are sourced by University of Bristol and have been granted copyright permission to JVET. University of Bristol is working on collecting copyright permissions from the third party source owners. The progress was reported positive.

In addition, discussions with respect to quality metrics, such as including MS-SSIM besides the conventional PSNR, and related issues were held among interested experts.

Input contributions

AHG11 related input contributions to this meeting included:

* JVET-T0041, Methodology and reporting template for neural network coding tool testing, S. Liu, A. Segall, E. Alshina, J. Boyce, M. Wien, D. Grois (AHG)
* JVET-T0042, Report of AHG11 Meeting on Neural Network-based Video Coding, S. Liu, E. Alshina, J. Pfaff, M. Wien, P. Wu, Y. Ye (AHG)
* JVET-T0057, AHG11: A case study to reduce computation of a neural network based in-loop filter by pruning, C. Auyeung, W. Wang, W. Jiang, X. Li, S. Liu (Tencent)
* JVET-T0058, AHG11: Information on inter-prediction coding tool with deep neural network, B. Choi, Z. Li, W. Wang, W. Jiang, X. Xu, S. Liu (Tencent)
* JVET-T0069, AHG11: SSIM based CNN model for in-loop filtering, T. Ouyang, F. Liu, H. Zhu, Z. Chen (Wuhan Unvi.), X. Xu, S. Liu (Tencent)
* JVET-T0073, AHG11: Neural Network-based intra prediction with transform selection in VVC, T. Dumas, F. Galpin, P. Bordes, F. Leleannec (Interdigital)
* JVET-T0079, AHG11: Neural Network-based In-Loop Filter, H. Wang, M. Karczewicz, J. Chen, A.M. Kotra (Qualcomm)
* JVET-T0088, AHG11: Convolutional neural networks-based in-loop filter, Y. Li, L. Zhang, K. Zhang, Y. He, J. Xu (Bytedance)
* JVET-T0092, AHG9/AHG11: Neural network based super resolution SEI, T. Chujoh, E. Sasaki, T. Ikai (Sharp)
* JVET-T0094, AHG11: In-loop filtering based on neutral network, T.-C. Ma, W. Chen, X. Xiu, Y.-W. Chen, H.-J. Jhu, C.-W. Kuo, X. Wang (Kwai)
* T0096

Summary results of technical proposals

Coding performance and runtime complexity

[Insert table]

Network information and complexity

Network information in training and inference stages [update a few table entries]

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Doc. #** | **Category** | Training time | Training data | Total Parameters | Memory Parameter (MB) | Memory Temp. (MB) | MAC (Giga) |
| JVET-0057 | Complexity reduction | 24h | DIV2K | 16,778 | 0.067 | 15238.7 | 129.61 |
| JVET-0058 | Virtual ref. picture | 7 days | Xiph.org and BVI-DVC (subset) | 125,993,047 | 504 | - | 200 |
| JVET-0069 | CNN in-loop filter | 57h | DIV2K | 250,971 | 1 | 32399.36 | 830 |
| JVET-0073 | NN intra prediction and LFNST | 8h | ILSVRC2012  DIV2K  CLIC2020 | 11,614,376 | 50.3 | - | 10.9 |
| JVET-0079 | CNN in-loop filter | - | - | 1million | - | - | - |
| JVET-0088 | CNN in-loop filter | 30h | DIV2K, BVI-DVC | 4,876,181 | 18.6 | 1012.5 | 3.05 |
| JVET-0094 | CNN in-loop filter | - | DIV2K | - | - | - | - |

The AHG recommended:

* To review input contributions;
* To continue investigating neural network-based video coding tools, including coding performance, complexity and quality metrics;
* To further discuss, define, refine and regulate common testing as well as training conditions for neural network-based video coding.

[JVET-T0012](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10485) JVET AHG report: High bit depth, high bit rate, and high frame rate coding (AHG12) [A. Browne, T. Ikai, X. Xiu]

This AHG report was reviewed in session 5 at 0500 UTC Thursday 8 October (chaired by GJS & JRO).

This document summarizes the activity of AHG12: High bit depth, high bit rate, and high frame rate coding between the 19th meeting by teleconference (22 June – 1 July 2020) and the 20th meeting by teleconference (7–16 October 2020).

In total there were 11 contributions noted to be related to high bit depth and one on high bit rate, but none related to high frame rate. The contributions are listed in the following section.

* Test conditions and sequences
  + JVET-T0047, AHG12: Preliminary common test conditions for high bit depth video coding, T. Ikai, A.Browne, X. Xiu
  + JVET-T0082, New HLG sequences for high bit-depth video coding, K. Kondo, M. Ikeda, A. Browne (Sony)
* Tools and transforms
  + JVET-T0084, AHG12: Transform coefficients range extension for high bit-depth coding, T. Zhou, T. Chujoh, E. Sasaki, T. Ikai (Sharp)
  + JVET-T0086, AHG12: SIMD implementation of BDOF for high bit-depth coding, T. Chujoh, E. Sasaki, T. Ikai (Sharp)
  + JVET-T0087, AHG12: VVC coding tool evaluation for high bit-depth coding, T. Ikai, T. Zhou, T. Hashimoto (Sharp)
  + JVET-T0091, AHG12: Fix on encoder overflow issue of the LMCS at high bit-depth coding, X. Xiu, H.-J. Jhu, Y.-W. Chen, T.-C. Ma, X. Wang (Kwai)
  + JVET-T0093, AHG12: On signalling for high-bit depth, Y. Kidani, K. Kawamura, K. Unno (KDDI)
* Golomb-Rice coding
  + JVET-T0072, AHG12: Rice parameter selection for high bit depths, A. Browne, S. Keating, K. Sharman (Sony)
  + JVET-T0085, AHG12: Rice parameter derivation for high bit-depth coding, T. Hashimoto, T. Ikai (Sharp)
  + JVET-T0089, AHG12: Slice based Rice parameter selection for transform skip residual coding, H.-J. Jhu, X. Xiu, Y.-W. Chen, T.-C. Ma, C.-W. Kuo, X. Wang (Kwai Inc.)
  + JVET-T0105, AHG12: On the Rice parameter derivation for high bit-depth coding, L. Pham Van, D. Rusanovskyy, M. Karczewicz (Qualcomm)
* Lossless coding (RGB and YCgCo)
  + JVET-T0111 YCgCo-R: Observations and findings [D. Buitenhuis (VideoLAN), A.M. Tourapis (Apple Inc)]

The following results comparing the performance of VTM 10.0 and HM 16.22 were obtained using the preliminary test conditions (JVET-T0047). The preliminary test conditions include SVT content previously included in the HEVC RExt test conditions, and new 12 bit PQ content converted from the same original sources as the PQ content in the VVC 10 bit HDR test conditions.

*Extended Precision Enabled*

The following results were obtained with extended precision enabled (--ExtendedPrecision=1) in both VTM 10.0 and HM16.22. Note that the VVC specification does not have this flag, although the VTM supports it.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **AI-12** |  |  |
|  |  |  | **Over HM16.22 EP1** |  |  |
|  | Y | U | V | EncT | DecT |
| SVT | 0.01% | 0.24% | 0.14% | 12387% | 132% |
| PQ444 | -4.78% | -6.16% | -6.84% | 4460% | 182% |
| PQ422 | -8.43% | -12.76% | -13.44% | 4460% | 182% |
|  |  |  |  |  |  |
| **Overall** | -4.40% | -6.23% | -6.71% | 6040% | 162% |
|  |  |  | **LDB-12** |  |  |
|  |  |  | **Over HM16.22 EP1** |  |  |
|  | Y | U | V | EncT | DecT |
| SVT | -3.59% | -1.19% | -1.19% | 1464% | 173% |
| PQ444 | -7.18% | -8.58% | -8.23% | 1055% | 182% |
| PQ422 | -9.47% | -14.41% | -15.61% | 1055% | 182% |
|  |  |  |  |  |  |
| **Overall** | -6.75% | -8.06% | -8.35% | 1298% | 180% |
|  |  |  | **RA-12** |  |  |
|  |  |  | **Over HM16.22 EP1** |  |  |
|  | Y | U | V | EncT | DecT |
| SVT | -3.68% | -1.42% | -1.43% | 2414% | 206% |
| PQ444 | -7.86% | -9.43% | -9.21% | 1512% | 200% |
| PQ422 | -10.26% | -15.29% | -16.84% | 1512% | 200% |
|  |  |  |  |  |  |
| **Overall** | -7.27% | -8.71% | -9.16% | 1903% | 202% |
|  |  |  | **AI-16** |  |  |
|  |  |  | **Over HM16.22 EP1** |  |  |
|  | Y | U | V | EncT | DecT |
| SVT | 36.73% | 34.40% | 34.33% | 11495% | 145% |
|  |  |  |  |  |  |
| **Overall** | 36.73% | 34.40% | 34.33% | 11495% | 145% |
|  |  |  | **LDB-16** |  |  |
|  |  |  | **Over HM16.22 EP1** |  |  |
|  | Y | U | V | EncT | DecT |
| SVT | 19.93% | 21.07% | 21.13% | 1604% | 219% |
|  |  |  |  |  |  |
| **Overall** | 19.93% | 21.07% | 21.13% | 1604% | 219% |
|  |  |  | **RA-16** |  |  |
|  |  |  | **Over HM16.22 EP1** |  |  |
|  | Y | U | V | EncT | DecT |
| SVT | 17.74% | 19.96% | 19.94% | 2494% | 219% |
|  |  |  |  |  |  |
| **Overall** | 17.74% | 19.96% | 19.94% | 2494% | 219% |

*Extended Precision Disabled*

The following results were obtained with extended precision disabled (--ExtendedPrecision=0) in both VTM 10.0 and HM16.22.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **AI-12** |  |  |
|  |  |  | **Over HM-16.22** |  |  |
|  | Y | U | V | EncT | DecT |
| SVT | 0.55% | 0.55% | 0.43% | 11780% | 144% |
| PQ444 | -5.09% | -6.23% | -6.86% | 4414% | 191% |
| PQ422 | -8.33% | -12.95% | -13.56% | 4414% | 191% |
|  |  |  |  |  |  |
| **Overall** | -4.29% | -6.21% | -6.66% | 5886% | 172% |
|  |  |  | **LDB-12** |  |  |
|  |  |  | **Over HM-16.22** |  |  |
|  | Y | U | V | EncT | DecT |
| SVT | -3.09% | -0.82% | -0.80% | 1228% | 145% |
| PQ444 | -7.84% | -8.88% | -8.50% | 1038% | 191% |
| PQ422 | -9.72% | -14.96% | -15.99% | 1038% | 191% |
|  |  |  |  |  |  |
| **Overall** | -6.88% | -8.22% | -8.43% | 1212% | 175% |
|  |  |  | **RA-12** |  |  |
|  |  |  | **Over HM-16.22** |  |  |
|  | Y | U | V | EncT | DecT |
| SVT | -3.20% | -1.16% | -1.15% | 2030% | 148% |
| PQ444 | -8.35% | -9.75% | -9.40% | 1456% | 192% |
| PQ422 | -10.51% | -15.72% | -17.03% | 1456% | 192% |
|  |  |  |  |  |  |
| **Overall** | -7.35% | -8.88% | -9.19% | 1759% | 177% |
|  |  |  | **AI-16** |  |  |
|  |  |  | **Over HM-16.22** |  |  |
|  | Y | U | V | EncT | DecT |
| SVT | 29.80% | 36.03% | 34.31% | 7995% | 162% |
|  |  |  |  |  |  |
| **Overall** | 29.80% | 36.03% | 34.31% | 7995% | 162% |
|  |  |  | **LDB-16** |  |  |
|  |  |  | **Over HM-16.22** |  |  |
|  | Y | U | V | EncT | DecT |
| SVT | 21.90% | 23.63% | 23.80% | 922% | 166% |
|  |  |  |  |  |  |
| **Overall** | 21.90% | 23.63% | 23.80% | 922% | 166% |
|  |  |  | **RA-16** |  |  |
|  |  |  | **Over HM-16.22** |  |  |
|  | Y | U | V | EncT | DecT |
| SVT | 19.66% | 21.56% | 21.69% | 1482% | 167% |
|  |  |  |  |  |  |
| **Overall** | 19.66% | 21.56% | 21.69% | 1482% | 167% |

The results show that currently VTM 10.0 offers gains over HM 16.22 for most 12 bit processing except SVT AI. However, there are significant losses for 16 bit content. This performance was not unexpected and some of the contributions listed above directly address this loss.

The AHG recommended the following:

* To review all related contributions
* To continue studying the benefits and characteristics of VVC coding tools for high bit depth and high bit rate coding.
* To continue investigating the requirements for future extensions of VVC to support high bit depth and high bit rate coding
* To continue refining the preliminary test conditions for high bit depth and high bit rate coding including the evaluation of new and modified test sequences
* To identify new test sequences for high frame rate coding which might be included in future test conditions

It was commented that for up to 12 bit, the extended precision may not be necessary. However, another participant said that some gain is shown for using this for some content.

It was commented that for the 16 bit SVT content, in some cases the extended precision flag is not helping. Another participant commented that without the extended precision flag enabled, the design basically does not support a full 16 bits of content, so the results on such content should be viewed skeptically.

See JVET-T0111 providing information about lossless coding, which reports that HEVC SCC is outperforming VVC for intra coding in some cases.

[JVET-T0013](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10486) JVET AHG report: Tool reporting procedure and testing (AHG13) [W.-J. Chien, J. Boyce, Y.-W. Chen, R. Chernyak, K. Choi, R. Hashimoto, Y.-W. Huang, H. Jang, R.-L. Liao, S. Liu]

This AHG report was discussed at 0520 Friday 9 October 2020 (chaired by GJS & JRO).

This document summarizes the activity of AHG13: “Tool reporting procedure” between the 19th meeting by teleconference (22 June – 1 July 2020) and the 20th Meeting by teleconference (7 – 16 Oct. 2020). Tool on/off experimental results vs. VTM anchor are provided for the tools specified in JVET-S2005.

The initial version of JVET-S2005 “Methodology and reporting template for tool testing” was provided on 20 July 2020.

All tests described in JVET-S2005 were conducted. VTM tool tests were conducted on VTM-10.0 software with VTM configuration by switching off or on specific tool either in configuration files or macros.

The tested tools, testers, and cross-checkers are listed in tables in the AHG report.

The results of the tests are summarized in Table 3-8 below. The attached spreadsheet provides additional data. Table 9 shows tool test results across several VTM versions. The method of computing combined BD-Rate\_YUV is similar to the suggested method in JVET-Q2016[2]. Instead of computing PSNR\_YUV for each frame and then averaging frame PSNR\_YUVs for a sequence, PSNR\_YUV is directly calculated from average PSNR\_Y, PSNR\_U, and PSNR\_V. The difference of the two methods is due to neglectable rounding error. Scatter plots (Figs. 1-3) are also provided for the tested tools in random access configuration, comparing PSNR-Y based bd-rate on the Y axis vs. each of Enc runtime ratio, Dec runtime ratio, and a weighted average of Enc and Dec runtime ratio, (*Enc + a\*Dec)/(a+1)*, with a configurable weight, *a*. The exemplary weighting is set to 6 and can be adjusted in the spreadsheet attached to this report.

Full experimental results and configuration files can be found at the link below:

<https://hevc.hhi.fraunhofer.de/svn/svn_VVCTestConfig/branches/VTM-10.0/>

There was no bitrate or PSNR differences between testers and cross-checkers.

Encoder and Decoder runtime ratios provided by both the testers and cross-checkers are included in the reporting template, to identify if there were significant runtime differences.

Table 3 Simulation results in all intra configuration (AI) of VTM tool tests. (VTM anchor)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | **AI** |  |  |  |
| **Acronym** | **BDR-Y** | **BDR-U** | **BDR-V** | **Tester EncTime** | **Tester DecTime** | **XChecker EncTime** | **XChecker DecTime** |
| CST | 0.41% | 9.87% | 9.55% | 147% | 99% | 153% | 104% |
| DQ | 1.71% | 1.41% | 1.24% | 94% | 97% | 94% | 102% |
| CCLM | 1.54% | 13.89% | 14.76% | 100% | 101% | 99% | 97% |
| MTS | 1.33% | 0.98% | 1.05% | 88% | 99% | 95% | 98% |
| ALF | 2.20% | 12.81% | 12.43% | 100% | 92% | 86% | 82% |
| MRLP | 0.33% | 0.12% | 0.14% | 99% | 101% | 100% | 103% |
| ISP | 0.48% | 0.28% | 0.23% | 85% | 99% | 86% | 100% |
| LMCS | 0.96% | 0.39% | 0.70% | 98% | 98% | 99% | 99% |
| MIP | 0.63% | 0.17% | 0.18% | 90% | 101% | 88% | 101% |
| LFNST | 0.97% | 2.02% | 2.28% | 111% | 100% | 107% | 98% |
| JCCR | 0.57% | 0.57% | 0.64% | 99% | 101% | 98% | 102% |
| SAO | 0.01% | 0.15% | 0.20% | 100% | 91% | 101% | 97% |
| CCALF | -0.14% | 9.69% | 8.55% | 100% | 97% | 101% | 101% |

Table 4 Simulation results in random access configuration (RA) of VTM tool tests. (VTM anchor)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | **RA** |  |  |  |
| **Acronym** | **BDR-Y** | **BDR-U** | **BDR-V** | **Tester EncTime** | **Tester DecTime** | **XChecker EncTime** | **XChecker DecTime** |
| CST | 0.12% | 4.12% | 4.57% | 98% | 99% | 100% | 99% |
| DQ | 1.60% | 0.95% | 0.66% | 97% | 99% | 99% | 102% |
| CCLM | 1.02% | 11.52% | 12.59% | 99% | 100% | 99% | 100% |
| MTS | 0.75% | 0.66% | 0.64% | 90% | 99% | 93% | 99% |
| ALF | 4.34% | 19.31% | 19.06% | 97% | 89% | 95% | 84% |
| AFFINE | 3.10% | 2.30% | 2.15% | 82% | 97% | 81% | 97% |
| SbTMC | 0.43% | 0.29% | 0.38% | 101% | 100% | 99% | 98% |
| AMVR | 1.42% | 2.15% | 2.36% | 85% | 101% | 88% | 104% |
| GPM | 0.67% | 1.14% | 1.16% | 91% | 100% | 97% | 101% |
| BDOF | 0.76% | 0.33% | 0.26% | 98% | 97% | 94% | 99% |
| CIIP | 0.26% | 0.00% | -0.02% | 97% | 99% | 94% | 97% |
| MMVD | 0.52% | 0.44% | 0.49% | 92% | 100% | 92% | 101% |
| BCW | 0.40% | 0.46% | 0.46% | 91% | 100% | 95% | 102% |
| MRLP | 0.17% | 0.08% | 0.10% | 100% | 100% | 98% | 102% |
| ISP | 0.28% | 0.29% | 0.33% | 96% | 100% | 95% | 100% |
| DMVR | 0.83% | 1.11% | 1.14% | 99% | 97% | 100% | 96% |
| SBT | 0.41% | -0.03% | -0.02% | 94% | 99% | 94% | 99% |
| LMCS | 1.38% | 1.14% | 0.99% | 95% | 97% | 93% | 98% |
| SMVD | 0.25% | 0.23% | 0.23% | 97% | 102% | 96% | 101% |
| MIP | 0.33% | 0.35% | 0.37% | 96% | 100% | 94% | 98% |
| LFNST | 0.70% | 0.78% | 1.08% | 95% | 100% | 92% | 99% |
| JCCR | 0.53% | 0.40% | 0.14% | 99% | 100% | 97% | 98% |
| SAO | 0.08% | 0.14% | 0.31% | 97% | 96% | 100% | 98% |
| PROF | 0.48% | 0.14% | 0.08% | 99% | 100% | 99% | 99% |
| CCALF | -0.13% | 13.88% | 13.73% | 100% | 99% | 99% | 99% |

Table 5 Simulation results in low delay B configuration (LDB) of VTM tool tests. (VTM anchor)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | **LDB** |  |  |  |
| **Acronym** | **BDR-Y** | **BDR-U** | **BDR-V** | **Tester EncTime** | **Tester DecTime** | **XChecker EncTime** | **XChecker DecTime** |
| CST | 0.07% | 1.64% | 2.36% | 97% | 94% | 102% | 97% |
| DQ | 1.64% | 0.36% | -0.04% | 96% | 95% | 100% | 100% |
| CCLM | 0.00% | 3.22% | 3.58% | 100% | 100% | 100% | 98% |
| MTS | 0.61% | -0.02% | -0.14% | 88% | 98% | 98% | 99% |
| ALF | 4.12% | 25.06% | 19.58% | 99% | 92% | 92% | 81% |
| AFFINE | 3.06% | 2.27% | 2.54% | 72% | 92% | 73% | 93% |
| SbTMC | 0.70% | 0.69% | 0.63% | 101% | 97% | 108% | 100% |
| AMVR | 0.56% | 0.70% | 0.98% | 87% | 100% | 90% | 99% |
| GPM | 1.54% | 1.79% | 1.99% | 89% | 103% | 95% | 100% |
| CIIP | 0.43% | 0.41% | 0.82% | 94% | 100% | 93% | 93% |
| MMVD | 0.49% | 0.22% | 0.26% | 90% | 101% | 96% | 98% |
| BCW | 0.24% | 0.20% | 0.52% | 92% | 94% | 117% | 106% |
| MRLP | 0.08% | -0.13% | 0.08% | 99% | 99% | 99% | 97% |
| ISP | 0.11% | 0.04% | 0.47% | 99% | 99% | 100% | 100% |
| SBT | 0.63% | -0.47% | -0.28% | 92% | 99% | 92% | 98% |
| LMCS | 0.94% | -0.30% | -0.15% | 95% | 96% | 95% | 96% |
| MIP | 0.14% | 0.27% | 0.50% | 95% | 101% | 96% | 100% |
| LFNST | 0.27% | 1.20% | 0.89% | 92% | 102% | 93% | 102% |
| JCCR | 0.13% | 1.92% | 2.44% | 100% | 98% | 99% | 100% |
| SAO | 0.11% | 0.43% | 1.23% | 97% | 93% | 100% | 96% |
| PROF | 0.32% | 0.01% | 0.20% | 96% | 97% | 97% | 95% |
| CCALF | -0.15% | 18.58% | 13.79% | 100% | 97% | 101% | 100% |

Table 6 Simulation results for screen coding tools for ClassF and ClassTGM (VTM anchor)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | **AI** |  |  |  |
| **Acronym** | **BDR-Y** | **BDR-U** | **BDR-V** | **Tester EncTime** | **Tester DecTime** | **XChecker EncTime** | **XChecker DecTime** |
| IBC Class F | 15.26% | 14.97% | 15.02% | 54% | 99% | 49% | 99% |
| IBC Class TGM | 47.74% | 45.15% | 45.20% | 63% | 101% | 61% | 106% |
| BDPCM ClassF | 0.91% | 0.92% | 0.98% | 98% | 101% | 98% | 98% |
| BDPCM ClassTGM | 1.44% | 1.57% | 1.57% | 100% | 101% | 97% | 98% |
|  |  |  |  | **RA** |  |  |  |
| IBC Class F | 12.49% | 12.53% | 12.68% | 86% | 100% | 84% | 100% |
| IBC Class TGM | 22.48% | 22.10% | 22.49% | 90% | 101% | 86% | 100% |
| BDPCM ClassF | 0.70% | 0.68% | 0.93% | 100% | 101% | 98% | 97% |
| BDPCM ClassTGM | 0.79% | 1.03% | 0.98% | 100% | 100% | 97% | 97% |
|  |  |  |  | **LD** |  |  |  |
| IBC Class F | 6.17% | 5.80% | 6.99% | 86% | 103% | 81% | 98% |
| IBC Class TGM | 11.38% | 11.97% | 12.23% | 87% | 107% | 83% | 101% |
| BDPCM ClassF | 0.35% | -0.37% | 0.35% | 100% | 99% | 102% | 100% |
| BDPCM ClassTGM | 0.29% | 0.42% | 0.37% | 100% | 99% | 103% | 106% |

Table 7 Simulation results of coding tools for color space 4:4:4 (VTM anchor)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | **AI** |  |  |  |
| **Acronym** | **BDR-Y** | **BDR-U** | **BDR-V** | **Tester EncTime** | **Tester DecTime** | **XChecker EncTime** | **XChecker DecTime** |
| PALETTE, YUV | 13.58% | 16.71% | 18.20% | 98% | 107% | 97% | 107% |
| PALETTE, RGB | 15.43% | 15.39% | 14.67% | 98% | 110% | 97% | 109% |
| ACT, RGB, Sensor | 7.91% | 2.53% | 3.54% | 98% | 101% | 97% | 101% |
| ACT, RGB, Computer | 15.33% | 0.74% | 6.68% | 76% | 100% |  |  |
|  |  |  |  | **RA** |  |  |  |
| PALETTE, YUV | 9.48% | 12.93% | 14.51% | 100% | 100% | 97% | 97% |
| PALETTE, RGB | 10.90% | 11.08% | 10.62% | 101% | 101% | 100% | 100% |
| ACT, RGB, Sensor | 12.41% | 4.94% | 6.00% | 95% | 99% | 93% | 96% |
| ACT, RGB, Computer | 26.42% | 4.00% | 11.16% | 87% | 99% |  |  |
|  |  |  |  | **LD** |  |  |  |
| PALETTE, YUV | 5.16% | 8.47% | 9.80% | 97% | 99% | 93% | 95% |
| PALETTE, RGB | 5.74% | 5.95% | 5.79% | 97% | 99% | 97% | 98% |
| ACT, RGB, Sensor | 17.76% | 7.41% | 8.29% | 93% | 95% | 95% | 95% |
| ACT, RGB, Computer | 38.43% | 6.80% | 17.92% | 91% | 102% |  |  |

Table 8 Luma sample usage and memory bandwidth results of VTM tool “off” test. (VTM anchor)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | AI |  | RA |  |  | LDB |  |
| Acronym | Sample usage | Sample usage | Ave mem BW | Max mem BW | Sample usage | Ave mem BW | Max mem BW |
| CCLM | 48.48% | 3.73% |  |  | 0.76% |  |  |
| ALF | 99.00% | 70.77% |  |  | 67.66% |  |  |
| AFFINE |  | 17.84% |  |  | 26.99% |  |  |
| SbTMC |  | 10.40% |  |  | 12.92% |  |  |
| AMVR |  | 5.43% |  |  | 2.51% |  |  |
| GPM |  | 2.44% |  |  | 6.00% |  |  |
| BDOF |  | 44.99% |  |  |  |  |  |
| CIIP |  | 0.90% |  |  | 1.55% |  |  |
| MMVD |  | 6.84% |  |  | 8.44% |  |  |
| BCW |  | 9.94% |  |  | 8.28% |  |  |
| MRLP | 6.51% | 0.59% |  |  | 0.23% |  |  |
| DMVR |  | 40.35% |  |  |  |  |  |
| SBT |  | 2.61% |  |  | 4.27% |  |  |
| SMVD |  | 2.89% |  |  |  |  |  |
| MIP | 23.76% | 5.09% |  |  | 2.40% |  |  |
| LFNST | 9.86% | 0.79% |  |  | 0.35% |  |  |
| JCCR | 10.95% | 0.50% |  |  | 0.11% |  |  |
| SAO | 31.39% | 6.92% |  |  | 7.76% |  |  |

Table 9 test results of VTM tool “off” test on various VTM versions

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | **VTM** | **RA** |  |  |  |
| **Abbreviation** | **VTM3** | **VTM4** | **VTM5** | **VTM6** | **VTM7** | **VTM8** | **VTM9** | **VTM10** |
| CST | 0.72% | 1.08% | 1.22% | 0.96% | 0.88% | 0.80% | 0.96% | 0.96% |
| DQ | 1.41% | 1.39% | 1.27% | 1.27% | 1.28% | 1.41% | 1.42% | 1.42% |
| CCLM | 3.94% | 4.01% | 3.84% | 3.57% | 3.60% | 3.26% | 3.43% | 3.43% |
| MTS | 1.25% | 0.82% | 0.37% | 0.68% | 0.70% | 0.71% | 0.73% | 0.73% |
| ALF | 3.61% | 3.71% | 4.78% | 4.65% | 4.63% | 7.06% | 7.18% | 7.18% |
| AFFINE | 2.43% | 2.47% | 2.39% | 2.82% | 2.80% | 2.80% | 2.93% | 2.93% |
| SbTMC | 0.52% | 0.43% | 0.40% | 0.48% | 0.43% | 0.43% | 0.41% | 0.41% |
| AMVR | 0.97% | 1.11% | 1.13% | 1.60% | 1.59% | 1.56% | 1.60% | 1.60% |
| GPM | 0.43% | 0.43% | 0.40% | 0.39% | 0.44% | 0.74% | 0.77% | 0.77% |
| BDOF | 1.02% | 0.63% | 0.67% | 0.67% | 0.66% | 0.66% | 0.66% | 0.66% |
| CIIP | 0.43% | 0.51% | 0.32% | 0.24% | 0.23% | 0.22% | 0.20% | 0.20% |
| MMVD | 0.81% | 0.52% | 0.59% | 0.52% | 0.51% | 0.51% | 0.50% | 0.50% |
| BCW | 0.48% | 0.45% | 0.46% | 0.43% | 0.41% | 0.40% | 0.41% | 0.41% |
| MRLP | 0.24% | 0.18% | 0.17% | 0.18% | 0.14% | 0.12% | 0.15% | 0.15% |
| ISP |  | 0.24% | 0.12% | 0.20% | 0.30% | 0.28% | 0.28% | 0.28% |
| DMVR |  | 0.80% | 0.87% | 0.87% | 0.89% | 0.88% | 0.90% | 0.90% |
| SBT |  | 0.33% | 0.34% | 0.31% | 0.31% | 0.31% | 0.32% | 0.32% |
| LMCS |  | 0.64% | 0.61% | 0.97% | 1.36% | 1.32% | 1.32% | 1.32% |
| SMVD |  | 0.26% | 0.24% | 0.27% | 0.26% | 0.26% | 0.24% | 0.24% |
| MIP |  |  | 0.28% | 0.32% | 0.37% | 0.34% | 0.34% | 0.34% |
| LFNST |  |  | 0.75% | 0.60% | 0.74% | 0.75% | 0.75% | 0.75% |
| JCCR |  |  | 0.28% | 0.35% | 0.32% | 0.41% | 0.41% | 0.40% |
| SAO | 0.80% | 0.63% | 0.16% | 0.13% | 0.12% | 0.10% | 0.11% | 0.11% |
| PROF |  |  |  | 0.41% | 0.39% | 0.38% | 0.40% | 0.40% |
| CCALF |  |  |  |  |  | 2.30% | 2.53% | 2.53% |

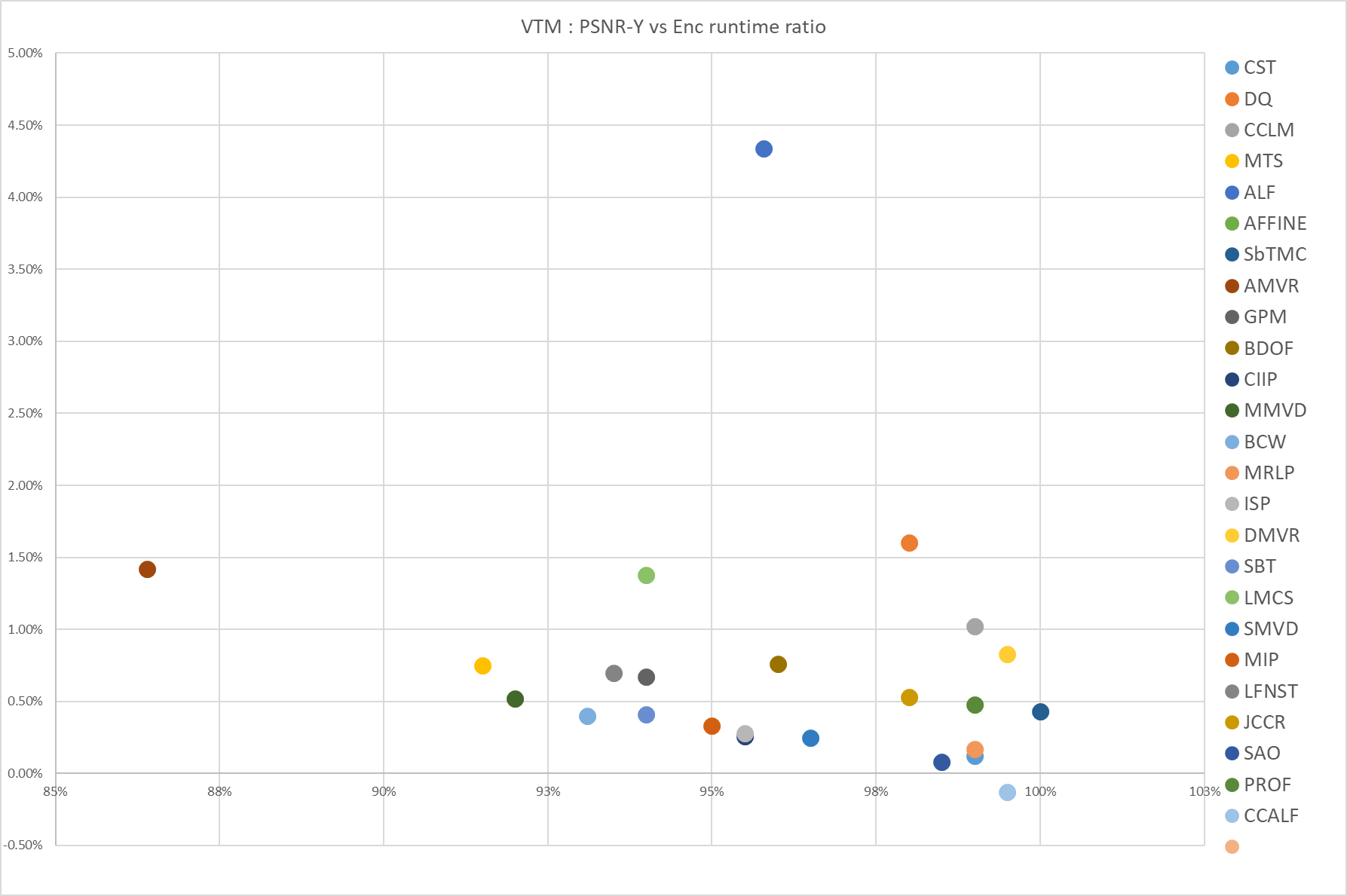


Figure 1. PSNR-Y vs encoding runtime ratio of VTM with VTM tool tests (VTM anchor)

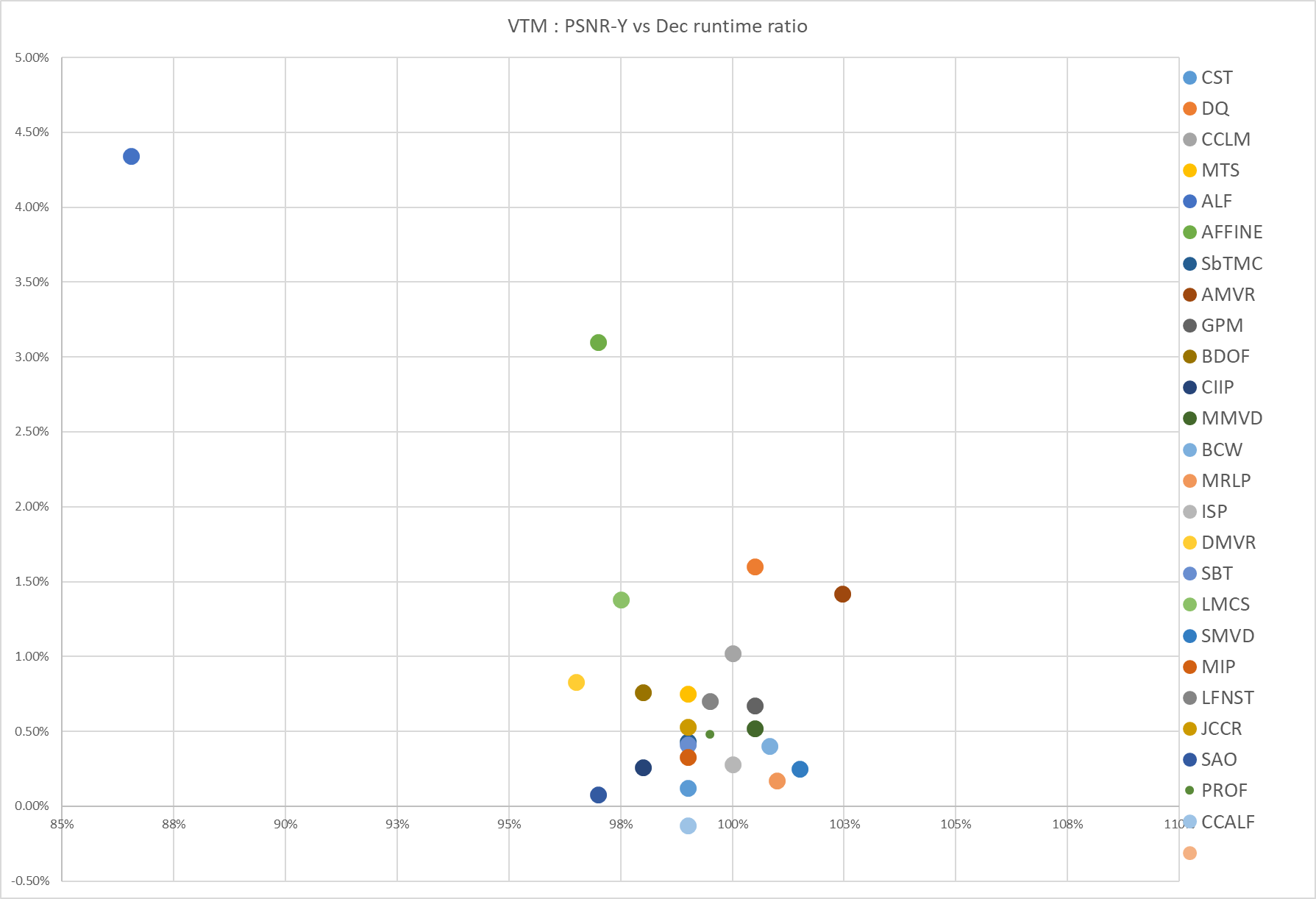


Figure 2. PSNR-Y vs decoding runtime ratio of VTM with VTM tool tests (VTM anchor)

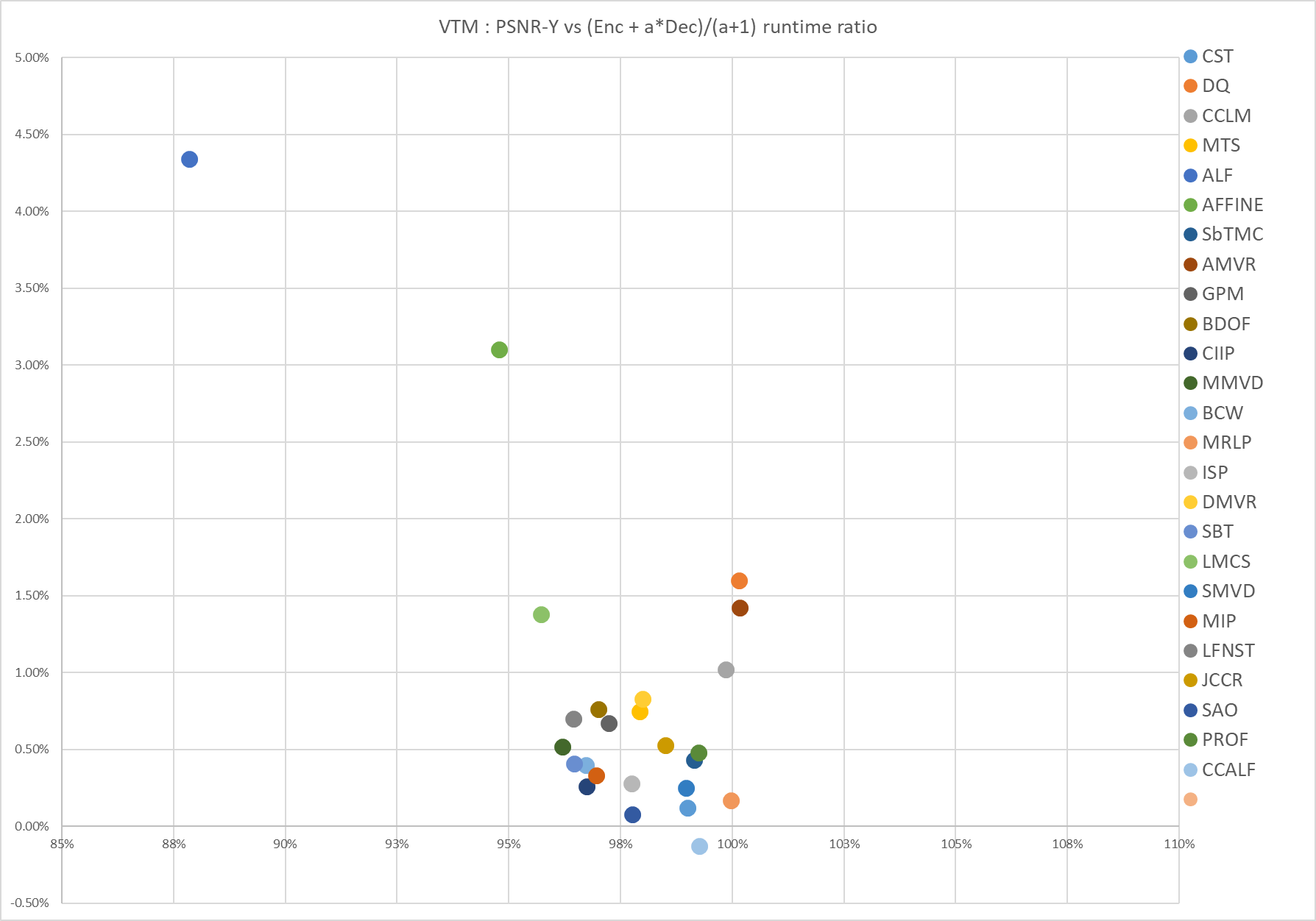


Figure 3. PSNR-Y vs weighted runtime ratio (a = 6) of VTM with VTM tool tests (VTM anchor)

[One point missing for ACT, fill in when available]

The AHG recommended the following:

* Review the performance of the coding tools
* Discontinue the AHG

No surprises were noted in the test results. For all aspects, the average test results were within 0.02% for luma. For JCCR there was a slightly larger different for chroma.

[JVET-T0021](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10487) JCT-VC AHG report: Project management (AHG1) [G. J. Sullivan, J.-R. Ohm]

This AHG report was discussed at 0530 Friday 9 October 2020 (chaired by GJS & JRO).

This document reports on the work of the JCT-VC ad hoc group on Project Management, including an overall status report on the project and the progress made during the interim period since the preceding meeting.

It is noted that, starting from the twentieth JVET meeting, work items which had previously been conducted in JCT-VC will be continued 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.

Seven input documents (JVET-T0048, JVET-T0050, JVET-T0051, JVET-T0052, JVET-T0054, JVET-T0075, JVET-T0113) relate to these items. Reports of the 5 AHGs which had been established at the 40th JCT-VC meeting were also submitted to JVET as JVET-T0021 through JVET-T0025. It is suggested to investigate possible merging of former JCT-VC AHGs with JVET AHGs to avoid duplication of efforts.

The JCT-VC email reflector had approx. 1294 subscribers as of 7 October 2020. It is suggested to keep it open, but send an email encouraging users to subscribe to the JVET reflector, and to continue future discussions on issues related to AVC, HEVC, and CICP on the JVET reflector as well.

The output documents from the preceding meeting had been made available at the "Phenix" site (<http://phenix.int-evry.fr/jct/>) or the ITU-based JCT-VC site (<http://wftp3.itu.int/av-arch/jctvc-site/2020_06_AN_Virtual/>), particularly including the following:

* The meeting report (JCTVC-AN1000), posted 2020-10-07
* High Efficiency Video Coding (HEVC) Test Model 16 (HM 16) Encoder Description Update 14 (JCTVC-AN1002), posted 2020-10-06
* Errata report items for HEVC, AVC, Video CICP, and CP usage TR (JCTVC-AN1004), posted 2020-10-06
* Annotated regions and shutter interval information SEI messages for AVC (Draft 1) (JCTVC-AN1006), posted 2020-09-11
* Usage of video signal type code points (Draft 1 for version 3) (JCTVC-AN1008), posted 2020-10-07

It is suggested to reserve JVET numbers, or a number space, to “living” JCT-VC documents, such as HM, ongoing standards under development, CTC, etc.

Software maintenance for AVC and HEVC generally was progressing according to plans, and should further be continued in JVET.

Since the approval of software copyright header language at the March 2011 parent-body meetings, that topic seems to be resolved.

Released versions of the software are available on the gitlab server at the following URL:  
https://vcgit.hhi.fraunhofer.de/jct-vc/HM/-/tags/*version\_number*,  
where *version\_number* corresponds to one of the versions described below – e.g., HM-16.22.

Intermediate code submissions can be found on a variety of branches available at:  
https://vcgit.hhi.fraunhofer.de/jct-vc/HM/-/branches/*branch\_name*,  
where *branch\_name* corresponds to a branch (eg., HM-16.22-dev).

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

The ftp site at ITU-T is used to exchange draft conformance testing bitstreams. The ftp site for downloading bitstreams is <http://wftp3.itu.int/av-arch/jctvc-site/bitstream_exchange/>.

A spreadsheet to summarize the status of bitstream exchange, conformance bitstream generation is available in the same directory. It includes the list of bitstreams, codec features and settings, and status of verification.

[JVET-T0022](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10488) JCT-VC AHG report: Test model editing and errata reporting (AHG2) [B. Bross, C. Rosewarne, J.-R. Ohm, K. Sharman, G. J. Sullivan, A. Tourapis, Y.-K. Wang]

This AHG report was discussed at 0535 Friday 9 October 2020 (chaired by GJS & JRO).

This document reports the work of the JCT-VC ad hoc group on (HEVC and AVC) test model editing and errata reporting (AHG2) between the 40th JCT-VC meeting by teleconference (Jun. 2020) and the 20th JVET meeting by teleconference.

Status of AHG2 work

JCT-VC output documents

[**JCTVC-AN1002**](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=11029) **High Efficiency Video Coding (HEVC) Test Model 16 (HM 16) Encoder Description Update 14**

The HM encoder description document was revised to include the GOP size 8 low delay structure.

[**JCTVC-AN1004**](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=11030)**Errata report items for HEVC, AVC, Video CICP, and CP usage TR**

All errata items noted in the meeting minutes of the 40th JCT-VC meeting have been integrated into the errata report.

JVET input documents related to JCTVC AHG2

**[JVET-T0048](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10427) On the syntax design for extending SEI messages and VUI for HEVC, VVC and VSEI [Bytedance, Microsoft]**

[**JVET-T0054**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10433) **Some errata items for HEVC and AVC (Bytedance, Apple)**

[**JVET-T0075**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10454) **AHG9: Errata and editorial clarifications for AVC and HEVC film grain characteristics SEI message semantics (Dolby)**

Text ticket review

*This table is reproduced from the previous AHG2 report (JCTVC-AN0002).*

|  |  |  |
| --- | --- | --- |
| **Ticket** | **Summary** | **Status** |
| 1507 | Duplicate syntax element ctxInc assignments in Table 9-48. | Confirmed present in H.265 (11/19)  (new ticket, filed 2020-04-18). |
| 1505 | Relates to semantics of AUs appearing after an ‘end of sequence’ or ‘end of bitstream’ NAL unit. | See previous meeting notes regarding JCTVC-AM0002. |
| 1504 | Small typos in profile\_tier\_level syntax in tabular form (7.3.3) | Aspect 1: Extra ‘)’ is still present in H.265 (11/19) Aspect 2: sub\_layer\_profile\_compatibility\_flag missing[ i ] issue confirmed fixed in H.265 (11/19). |
| 1500 | typo in equation (8-69),(8-70) (palette\_transpose\_flag not transposing) | Confirmed present in H.265 (11/19). |
| 1498 | Typos in Table 9-43 (for palette\_run\_suffix PalletMaxRun should be PalletMaxRunMinus1) | Confirmed present in H.265 (11/19). |
| 1491 | duplicate invocation of 9.3.4.3 arithmetic decoding process (invoked both in 9.3.4.1 and also in 9.3.4.2). | Confirmed present in H.265 (11/19) – confirm whether action is needed) |
| 1427 | Eqns (8-185) and (8-187) could be removed as editorial cleanup. | The ticket is marked as enhancement and note the equation numbers are updated since the ticket was filed. |

Decision: It was confirmed that these should be included in JVET-T1004.

The following ticket was closed previously but is marked for further study in JCTVC-AN1004:

|  |  |  |
| --- | --- | --- |
| 1372 | IntraPredModeY is not set to be DC for palette and IntraBC mode. | *(Included in JCTVC-AN1004)* Not present in JCTVC-W1005. |

It was agreed that #1372 this is an obsolete report and does not need to be further tracked.

The recommendations of the HEVC test model editing and errata reporting JCT-VC AHG are for JVET to:

* Encourage the use of the issue tracker to report issues with the text of both the HEVC specification and the encoder description.
* Review the above-identified errata reports.
* Confirm resolutions of open tickets (if any) in the issue tracker and close them.

[JVET-T0023](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10489) JCT-VC AHG report: Software development and software technical evaluation (AHG3) [K. Sühring, B. Li, K. Sharman, V. Seregin, G. Tech, A. Tourapis]

This AHG report was discussed at 0605 Friday 9 October 2020 (chaired by GJS & JRO).

This report summarizes the activities of the JCT-VC AhG on HEVC, AVC and HDRTools software development and software technical evaluation that have taken place between the 39th and 40th JCT-VC meetings.

There had not been any releases of software during the reporting period.

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

* [HM-16.22](https://vcgit.hhi.fraunhofer.de/jct-vc/HM/-/releases/HM-16.22) (Jul. 2020)
  + (svn [HM 16.20](https://hevc.hhi.fraunhofer.de/trac/hevc/browser/tags/HM-16.20) (Sep. 2018) )
* [HM-16.21+SCM-8.8](https://vcgit.hhi.fraunhofer.de/jct-vc/HM/-/tags/HM-16.21+SCM-8.8) (Mar. 2020)
  + (svn [HM 16.20 + SCM 8.8](https://hevc.hhi.fraunhofer.de/trac/hevc/browser/tags/HM-16.20%2BSCM-8.8) (Mar. 2018) )
* [SHM 12.4](https://hevc.hhi.fraunhofer.de/trac/shvc/browser/SHVCSoftware/tags/SHM-12.4) (Jan. 2018) [svn]
* [HTM 16.3](https://hevc.hhi.fraunhofer.de/trac/3d-hevc/browser/3DVCSoftware/tags/HTM-16.3) (Jul. 2018) [svn]
* [JM 19.0](https://vcgit.hhi.fraunhofer.de/jct-vc/JM/-/tags/JM-19.0)
* [JSVM 9.19.15](https://vcgit.hhi.fraunhofer.de/jct-vc/jsvm/-/tags/JSVM_9_19_15) (Note that this item was recently recovered)
* [JMVC 8.5](https://vcgit.hhi.fraunhofer.de/jct-vc/jmvc/-/tags/JMVC_8_5) (Note that this item was recently recovered)
* [3DV ATM 15.0](https://vcgit.hhi.fraunhofer.de/jct-vc/3dv-atm/-/tags/3DV-ATM_v15.0)
* [HDRTools 0.19.1](https://gitlab.com/standards/HDRTools/-/tags/v0.19.1) (Sep. 2019)

A repository containing MFC and MFCD had not been identified.

HM related activities

HM16.22 was tagged on Jul. 3, 2020. . Changes include:

* JCTVC-AK0030 (Change to random-access encoder configuration).
* Changes to LowDelay GOP-8 QP structure to match those used by VTM.
* JCTVC-AK1005 (Shutter interval information SEI)
* Additional checks to warn if DPB limits would be exceeded by a configuration.
* Porting of JVET’s parcat software for concatenating simulations run in parallel.
* Removal of macros.
* Updates to the software reference manual for the new cmake build process.
* Addition of encoder controls for some SEIs (from author of Shutter interval SEI), namely ambient view environment SEI, content light level SEI, and film grain characteristics SEI.
* Shutter interval info SEI message.

The following actions had yet to be included:

* JCTVC-AM0023 (Illustration of the film grain characteristics SEI message in HEVC) – a merge request had been made, and the software coordinators had made a first pass of the software; a second pass is required to verify the function.
* JCTVC-AJ0028 (Encoder-only Supplemental Motion Vector Estimation for Point cloud Coding content) – some minor changes remain.

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

Except for the addition above, the support of SEI messages had not changed and is summarized in JCTVC-AM0003. It was noted that the SEI manifest SEI message software support was not yet in the HM, although it may be in the JSVM. It was commented that some software may have been provided for that at the Marrakech meeting of 2019-01.

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:

* 39 tickets for “HM”, most of which are more than 5 years (see in particular a recent ticket about CABAC termination),
* 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

Codebases without recent action

* SCC SCM
* SHM
* HTM of MV-HEVC and 3D-HEVC
* HDRTools
* JM, JSVM, JMVM

Recommendations

* Continue to develop reference software based on HM 16.22, HM 16.21 + SCM 8.8, SHM 12.4, HTM 16.3 and HDRTools 0.19.1 and improve their quality.
* Test the reference software more extensively outside of common test conditions.
* Add more conformance checks to the decoder to more easily identify non-conforming bit-streams, especially for profile and level constraints.
* Encourage people who are implementing HEVC based products to report all (potential) bugs that they are finding in that process.
* Encourage people to submit bit-streams that trigger bugs in the HM. Such bit-streams may also be useful for the conformance specification.
* Encourage people to submit configuration files that trigger bugs in HDRTools.
* Continue to investigate the merging of branches.
* Keep common test conditions aligned for the different standards.

It was suggested to merge the two software maintenance AHGs.

[JVET-T0024](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10490) JCT-VC AHG report: Supplemental enhancement information (AHG4) [J. Boyce, C. Fogg, S. McCarthy, G. J. Sullivan]

This AHG report was discussed at 0620 Friday 9 October 2020 (chaired by GJS & JRO).

This document summarizes the activity of AHG4: Supplemental enhancement information between the 40th JCT-VC meeting held by teleconference, and with the transfer of the scope of the AHG into JVET, the 20th JVET meeting.

The key activity in the AHG was generation of output document JCTVC-AN1006 “Annotated regions and shutter interval information SEI messages for AVC”.

There were five AVC and HEVC SEI related input contributions. Two contributions have errata for existing SEI messages in the AVC and/or HEVC specifications. Three contributions related to software support for the Annotated Regions SEI message in HEVC and the JCTVC-AN1006 output document for AVC.

* Errata to existing HEVC and AVC SEI message
  + [**JVET-T0075**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10454) **AHG9: Errata and editorial clarifications for AVC and HEVC film grain characteristics SEI message semantics [S. McCarthy, P. Yin, W. Husak, T. Lu, F. Pu, T. Chen (Dolby)]**
  + [**JVET-T0048**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10427) **On the syntax design for extending SEI messages and VUI for HEVC, VVC and VSEI [Y.-K. Wang (Bytedance), G. J. Sullivan (Microsoft)]**
* Software related to Annotated Regions SEI message

[**JVET-T0050**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10429) **AHG9: HEVC Annotated Regions SEI message software update for HM [P. Guruva reddiar, J. Boyce (Intel)]**

[**JVET-T0051**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10430) **AHG9: Object tracking information file format for use with Annotated Regions SEI messages [P. Guruva reddiar, J. Boyce (Intel)]**

[**JVET-T0052**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10431) **AHG9: AVC Annotated Regions SEI message software for JM [P. Guruva reddiar, J. Boyce (Intel)]**

The AHG recommended the following:

* Review input contributions
* Consider if the scope of this AHG should be rolled into the VVC SEI AHG

[JVET-T0025](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10491) JCT-VC AHG report: Test sequence material (AHG5) [T. Suzuki, V. Baroncini, E. François, P. Topiwala, S. Wenger]

This AHG report was discussed at 0630 Friday 9 October 2020 (chaired by GJS & JRO).

There was basically no update from the last meeting.

A revised version of the report adds discussion of high bit depth test sequences. The copyright status for use of the monochrome high bit depth medical test material outside of work on HEVC was reported to be under study, and at the moment is expressed as restricted to HEVC development work in JCT-VC. It was commented that DICOM (or its members) may have suitable test material.

The lists of available test sequences was provided in the AHG report.

It was noted that the University of Hannover ftp site address has changed.

# Project development (XX)

## General (0)

## Text development and errata reporting (6, 1 TBP, 1 revisit)

[JVET-T0048](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10427) On the syntax design for extending SEI messages and VUI for HEVC, VVC and VSEI [Y.-K. Wang (Bytedance), G. J. Sullivan (Microsoft)]

This contribution was discussed 1900 Wednesday 7 October 2020 (chaired by JRO).

This contribution discusses the SEI payload syntax in AVC, HEVC, and VVC/VSEI, and the VUI payload syntax in VVC/VSEI. It reports an asserted bug on the HEVC buffering period SEI message syntax extension and proposes a fix for the asserted bug, and discusses how extensions, when needed, should be specified for the syntax of SEI messages in HEVC and VSEI, and to the syntax of VUI in VSEI.

The first method was suggested to be “cleaner”, future proof, and easier to understand, while the second would isolate the changes to the one place where there is an asserted problem in HEVC.

For VVC, it was asserted that the first proposed approach would enable detection in the VVC spec of whether the VSEI spec had been extended.

It was discussed whether a third approach could be used, or a NOTE could be sufficient to clarify the corner case. It was sugested that the problem could exist only if the use\_alt\_cpb\_params\_flag is equal to 1.

The existing extension mechanism for SEI messages (which is there from HEVC and was also used in VVC/VSEI) was reported to cause a problem, as an extension of less than one byte with a pattern 10..0 would be interpreted as a byte alignment mechanism. The issue could only occur for very short extension payloads, as otherwise the presence of the extension would be identified by the payload size. Currently, only the buffering period SEI message could have a case where it is affected.

Two possible solutions are suggested: a) Define a longer payload extension for the SEI messages that are potentially affected (“second method”). b) Define for the buffering period SEI message (and for future SEI messages that might have a similar problem) an extra flag that clearly signals the presence of an extension (“first method”).

During the discussion, it is suggested that another solution could be to put a note into the spec. pointing out this case for buffering period. This would not affect existing implementations, but make encoder manufacturers aware that a decoder would ignore an intended extension in this corner case (which a more smart encoder could avoid). This would however not avoid similar cases in newly defined SEI messages (if not carefully designed, or proponents of such SEI not being aware of this problem), for future extensions the “first method” appears to be cleaner and better to understand.

Offline study and revisit was requested.

[JVET-T0054](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10433) Some errata items for HEVC and AVC [Y.-K. Wang, L. Zhang (Bytedance), A. M. Tourapis (Apple)]

This contribution was discussed 1945 Wednesday 7 October 2020 (chaired by JRO & GJS).

*Errata items for both HEVC and AVC*

1. Change the range notation from the following:

x = y..z x takes on integer values starting from y to z, inclusive, with x, y, and z being integer numbers and z being greater than y.

to be as follows (same as in VVCv1):

x = y..z x takes on integer values starting from y to z, inclusive, with x, y, and z being integer numbers and z being greater than or equal to y.

*Errata items for HEVC only*

1. Search and replace "satisify" with "satisfy".
2. In clause 9.3.3.6, fix the wrong indention for the paragraph starting with "Otherwise".
3. Search and check "CVS", particularly in semantics, and fix to be CLVS as needed.
4. Search and replace instances of "clasue" with "clause".
5. The variables InitCpbRemovalDelay[ SchedSelIdx ] and InitCpbRemovalDelayOffset[ SchedSelIdx ] are initially derived in clause C.2.1 and updated in clause C.2.3, and the second variable is not even used in clause C.2.3. However, currently, the wording in clause C.2.3 is that these variables are derived, which can confuse readers as to why the second variable is derived but not used in clause C.2.3 thus why not just remove the mentioning of the second variable from clause C.2.3.

Similar issue was identified during the development of VVCv1, and the final text in there was to use the phrase "updated" instead of "derived" for these two variables in clause C.2.3.

It is therefore proposed to change the following phrase in clause C.2.3:

The variables InitCpbRemovalDelay[ SchedSelIdx ], InitCpbRemovalDelayOffset[ SchedSelIdx ], CpbDelayOffset and DpbDelayOffset are derived as follows:

to be as follows:

The variables InitCpbRemovalDelay[ SchedSelIdx ] and InitCpbRemovalDelayOffset[ SchedSelIdx ] are updated, and the variables CpbDelayOffset and DpbDelayOffset are derived, as follows:

1. Replace all the eight instances of "decoder conformance to profiles" with "the decoding process". These instances are as follows:

**vps\_extension\_data\_flag** may have any value. Its presence and value do not affect decoder conformance to profiles specified in Annex A. Decoders conforming to a profile specified in Annex A but not supporting the INBLD capability specified in Annex F shall ignore all vps\_extension\_data\_flag syntax elements.

**sps\_extension\_data\_flag** may have any value. Its presence and value do not affect decoder conformance to profiles specified in this version of this Specification. Decoders conforming to this version of this Specification shall ignore all sps\_extension\_data\_flag syntax elements.

**pps\_extension\_data\_flag** may have any value. Its presence and value do not affect decoder conformance to profiles specified in this version of this Specification. Decoders conforming to this version of this Specification shall ignore all pps\_extension\_data\_flag syntax elements.

**slice\_segment\_header\_extension\_data\_byte** may have any value. Decoders shall ignore the value of slice\_segment\_header\_extension\_data\_byte. Its value does not affect decoder conformance to profiles specified in Annex A.

**vps\_extension\_data\_flag** may have any value. Its presence and value do not affect decoder conformance to profiles specified in Annexes A, G or H. Decoders conforming to a profile specified in Annexes A, G or H shall ignore all vps\_extension\_data\_flag syntax elements.

**vps\_non\_vui\_extension\_data\_byte** may have any value. Decoders shall ignore the value of vps\_non\_vui\_extension\_data\_byte. Its value does not affect decoder conformance to profiles specified in this version of this Specification.

**slice\_segment\_header\_extension\_data\_bit** may have any value. Decoders shall ignore the value of slice\_segment\_header\_extension\_data\_bit. Its value does not affect decoder conformance to profiles specified in this version of this Specification.

**vps\_extension\_data\_flag** may have any value. Its presence and value do not affect decoder conformance to profiles specified in Annexes A, G, H, or I. Decoders conforming to a profile specified in Annexes A, G, H, or I shall ignore all vps\_extension\_data\_flag syntax elements.

1. In the informative Table D.1, change the persistence scope for the SEI manifest SEI message and the SEI prefix indication SEI message from "The CLVS containing the SEI message" to "The CVS containing the SEI message", to be aligned with the semantics of the SEI messages. Note that this is not needed for AVC as AVC does not have such an informative table.
2. Remove a tab from each of the two middle rows below:

|  |  |
| --- | --- |
| content\_light\_level\_info( payloadSize ) { | **Descriptor** |
| **max\_content\_light\_level** | u(16) |
| **max\_pic\_average\_light\_level** | u(16) |
| } |  |

Decision (BF): Adopt into next available revisions.

[JVET-T0055](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10434) Some errata items for VVC [Y.-K. Wang, Z. Deng (Bytedance)]

This contribution was discussed 1955 Wednesday 7 October 2020 (chaired by JRO & GJS).

This contribution reports some asserted bugs in VVC 1st edition and proposes fixes for the asserted bugs. The asserted bugs and proposed fixes are summarized as follows:

1. The current constraint on POC values in clause C.4 unnecessarily requires the prevTid0Pic to be within the POC distance of MaxPicOrderCntLsb / 2 even when the current picture has the ph\_poc\_msb\_cycle\_val signalled or has an inter-layer reference picture, and under either condition the inference of POC MSB is not needed. Consequently, the constraint disallows the intended use of the syntax element vps\_max\_tid\_il\_ref\_pics\_plus1[ i ][ j ] equal to 0 without having a great value of sps\_log2\_max\_pic\_order\_cnt\_lsb\_minus4.

The proposed fix for this asserted bug is to require the prevTid0Pic to be within the POC distance of MaxPicOrderCntLsb / 2 only when the current picture does not have the ph\_poc\_msb\_cycle\_val signalled and does not have an inter-layer reference picture.

It was commented that this is primarily just removing a constraint in a case where the constraint is not necessary, and it may not be strictly necessary to relax the constraint. At least the wording of two NOTEs needs changing, as they imply that the constraint is not imposed.

The proponent said that the design intent was to not impose the tight constraint, as that is what is expressed in the two NOTEs.

It was commented that the POC can be used for detecting AU boundaries, and asked if removing the constraint could confuse AU detection.

Relevant use cases would include multilayer scenarios, such as inter-layer prediction from IRAP pictures only.

If we do not relax the constraint, the worst impact would be that the encoder would need to use more bits for POC LSBs than would have been otherwise necessary.

1. In the general sub-bitstream extraction process, the step 9, which removes from outBitstream all SEI NAL units that contain a scalable nesting SEI message that has sn\_ols\_flag equal to 0 and there is no value in the list NestingLayerId equal to a value in the list LayerIdInOls[ targetOlsIdx ] is redundant.

The proposed fix for this asserted bug is to remove the step 9 of the general sub-bitstream extraction process.

1. In the subpicture sub-bitstream extraction process, when at least one VCL NAL unit has been removed by the step 2, the step 5 removes from outBitstream all SEI NAL units that contain scalable nesting SEI messages with sn\_subpic\_flag equal to 0, including nested SLI SEI messages. However, nested SLI SEI messages, when present, may be needed by the step 6 for parameter sets rewritting.

The proposed fix for this asserted bug is to move step 5 after step 6 in the subpicture sub-bitstream extraction process.

1. It is allowed that both nested and non-nested HRD-related SEI messages are present and apply to the same OLS. Similarly, it is also allowed that both nested and non-nested non-HRD-related SEI messages are present and apply to the same layer. However, there lacks a constraint to require that the content of such SEI messages of a particular payload type to be the same.

The proposed fix for this asserted bug is to require such nested and non-nested SEI messages to have the same SEI payload content.

Decision (BF): Adopt into next available revisions. For aspect #1, rephrase the NOTEs without removing the normative constraint.

[JVET-T0075](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10454) AHG9: Errata and editorial clarifications for AVC and HEVC film grain characteristics SEI message semantics [S. McCarthy, P. Yin, W. Husak, T. Lu, F. Pu, T. Chen (Dolby)]

This contribution was discussed 2030 Wednesday 7 October 2020 (chaired by JRO & GJS).

The contribution proposes editorial corrections and clarifications for the semantics of the film grain characteristics (FGC) SEI message in the current versions of AVC and HEVC. The proposed editorial corrections and clarifications are the same as those already made for the semantics of the FGC SEI message in the current version of VSEI.

The contribution proposed a sentence saying “When VUI parameters are not present for the CVS, and equivalent information is not conveyed by external means, separate\_colour\_description\_present\_flag shall be equal to 1.”

There was discussion of this sentence, regarding whether this case is possible in the AVC and HEVC contexts and whether it should discuss VUI parameters presence or colour description syntax presence.

There was an error in subtracting 8 from something that was already specified as “\_minus8”.

It was commented that the variable names and usage of multiple index variables could be improved.

Terminology should be double-checked in final editing.

Decision (BF): Adopt into next available revisions, with editorial adjustments as discussed above.

[JVET-T0110](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10507) AHG2: Editorial and errata input on VVC draft text [B. Bross, J. Chen, S. Liu, Y.-K. Wang]

This contribution was discussed 2100 Wednesday 7 October 2020 (chaired by JRO & GJS).

-v3 of the contribution was discussed.

This document provides integrations of a number of bug fixes reported in the bug tracking system as well as some other editorial changes made by the authors, based on the latest JVET output draft VVC text in JVET-S2001-vH.

List of logged changes:

* Replaced Fig. 20 with the correct one (now the same as in the final FDIS text).
* Fixed tickets [#1364](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1364), [#1365](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1365), [#1366](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1366), [#1367](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1367), [#1368](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1368), [#1369](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1369), [#1371](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1371), [#1372](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1372), [#1379](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1379), #1380, [#1381](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1381), [#1386](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1386), [#1387](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1387), [#1388](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1388)
* Fixed block vector used for IBC-coded chroma blocks in equations (1109, 1110) for 4:2:2 and 4:4:4 9 (Thanks to Li Zhang).
* Fixed typos ‘mesage’ and a couple of wrong minus sign characters for the interpolation filter coefficients.
* Added missing “NAL unit” (3 instances) after “AUD ”.
* Added a dependency on y component to completely fix the issue reported in [#1372](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1372).
* Added a missing condition for the inference of tu\_y\_coded\_flag to completely fix the issue reported in [#1366](https://jvet.hhi.fraunhofer.de/trac/vvc/ticket/1366).

See also the notes in AHG2 report. This is a collection of fixes done in response to tickets beyond -vH of S2001.

Decision (BF): Adopt into next available revision.

Additional offline study for double-checking was encouraged.

It was agreed to produce an output text to include all fixes relative to JVET-S2001-vH and for the editors to work with the publication staff in the parent bodies to get as many of these fixes into the publication processes as possible.

End of session 3.

[JVET-T0113](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10510) Errata for draft report on usage of video signal type code points [Y. Syed (Comcast)] [late]

This document lists the document errata or clarifications in the planned 3rd edition of ISO/IEC DTR 23091-4: 202x.

Decision: Agreed changes relative to balloted text:

* Ballot comments: Removal of version numbers: Agreed, and similarly remove version numbers from BT.656-5, BT.1120-9, BT.2077-2 (Tables 8 & 9).
* Change “Interstitial” to “Vertically interstitial”.
* Typo: “ST.2036-4”
* Copy-paste typo: Bottom line of Table 3 “ChromaLocType = 0” should be “ChromaLocType = 2”
* Left alignment on bibliographic entry for SMPTE RA.

## Test conditions (1)

See also T0103 and T0099 on VVenC and VVdeC optimized implementations and discussion of MCTF.

[JVET-T0063](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10442) Suggestion to upgrading random access configuration in CTC with GOP 32 configuration [K. Andersson, J. Enhorn, R. Sjöberg, J. Ström, L. Litwic, P. Wennersten (Ericsson)]

[Add summary.]

This was discussed in session 15 at 2120 Tuesday 13 October (chaired by GJS & JRO).

3% gain for SDR, 4% for some HDR metrics.

This is already being done in the VT for SDR 4K.

(This change should not be done for HM CTC, as it violates the DPB constraint for operation at large picture sizes.)

Decision (CTC): Adopt into VVC CTC.

## Verification test (7)

See also T0060 and T0120 for test material, T0063 on CTC, and T0103 and T0099 on VVenC and VVdeC optimized implementations and discussion of MCTF.

Contributions in this area were initially discussed in Session 8 at 0720 UTC on Friday 9 October 2020 (chaired by GJS and JRO).

[JVET-T0043](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10422) Status Report on SDR HD Low Delay Verification Test Preparation [M. Wien, V. Baroncini]

This was an input to an interim AHG meeting, and was reviewed there.

[JVET-T0044](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10423) Status Report on 360 Video Verification Test Preparation [M. Wien, V. Baroncini]

This was an input to an interim AHG meeting, and was reviewed there.

[JVET-T0045](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10424) AHG4: Agenda and report of the AHG meeting on the SDR HD Verification Tests on 2020-09-03 [M. Wien, V. Baroncini]

This was a report of an interim AHG meeting. The meeting was held online on Thursday, 2020-09-03, 16:00h UTC.

It was noted that the gains observed for SDR HD in the low-delay (LD) configuration appear to be lower than what has been reported for the SDR UHD sequences in the random-access (RA) configuration before.

It was commented that this may potentially be related to the LD configuration that has been used in this setup. It was suggested to investigate this issue by looking at the performance of the same sequences using the RA configuration and by downscaling and coding UHD sequences using LD which have been assessed under RA configuration before.

It was suggested to do testing of SDR HD content in both, LD and RA configurations. It was asserted that RA coding of HD content is a relevant application case which should be addressed in the VVC verification tests.

In terms of test content, it was suggested broaden the set of test sequences under consideration and to seek for suitable content representing conversational applications, as well as mobile streaming applications.

In the AHG meeting it was commented that the QP settings used in the LD-B configuration from the JVET common test conditions might not be optimized for visual assessment. It was noted that for HD LD, the behaviour of the MOS curves is following the behaviour of the PSNR curves much closer than for the UHD RA case where the visual quality was found to be higher, by tendency, compared to the PSNR curve. It was suggested to investigate the LD configuration files in order to address potential issues with visual quality. It was mentioned that the GOP size used in the LD configuration could impact the visual quality. A GOP size change from 4 to 8 pictures in the VTM was adopted into the common testing conditions based on contribution JVET-P0345. The LD configuration of HM-16.20 uses a GOP size of 4. It was suggested that HM and VTM should be comparable in terms of configuration (if both specifications support the corresponding settings).

It was commented in the AHG meeting that the content currently under investigation does not match the LD application case well (the Witcher1 test sequence might be closest). It was noted that the LD configuration should be demonstrated on conversational and gaming content.

Next steps identified as needed were discussed as follows:

* Investigate the appropriateness of the available LD configuration for subjective evaluation, e.g. by expert viewing.
* See contribution T0097 on RA configuration.
* Include SDR HD RA as a verification test category, besides SDR HD LD. The test sequences for both cases could be selected independently. The current test sequences for HD VT are no typical for low delay applications.
  + More conferencing content is needed. See T0060, which proposes new test sequences. Another test sequence had recently been provided by Ericcson, and it was requested for a corresponding document to be provided.
  + A set of tests for HD RA is desired.
* It was commented that HM 16.22 should be used to be able to use the latest configurations of LD and RA GOP structures.
* Seek test sequences suitable for visual assessment both types of application scenarios, conversational/gaming (LD) and streaming (RA). Some sequences of the current SDR HD test set might be suitable if encoded with the RA configuration. JVET experts are invited to suggest and provide test sequences to the verification test coordinators.

It was planned to conduct remote expert viewing during the meeting to identify the suitable test points in the HD SDR RA category.

Verification test on HD SDR RA could likely be conducted during the upcoming meeting cycle (if the pandemic situation allows).

[JVET-T0046](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10425) AHG4: Agenda and report of the AHG meeting on the 360° Video Verification Tests on 2020-09-04 [M. Wien, V. Baroncini, Y. Ye]

This was a report of an interim AHG meeting. The meeting was held online on Friday, 2020-09-04, 16:00h UTC.

It was noted that the set of sequences used in the PCMP/GCMP formats was reduced due to cubemap artefacts which were considered to be too strong for use in the subjective evaluation. It was pointed out that in the used implementation, no blending of the padded cube faces was applied during the rendering process. It was agreed that blending should be enabled here. Offline coordination of the required implementation task was recommended.

It is expected that the visual quality of the rendered sequences would thereby be improved. It is noted that in case of PERP, blending is already applied for this reason.

It was noted that the quality at high QPs was too low especially for the GT\_Sheriff sequence. The interpretation of the MOS data at the low end generally needs to be treated with care. It was commented that both, quality and rate-based evaluation is planned to be applied. This is the reason why five rate points have been selected. It was further commented that some narrowing of the rate and quality ranges is likely to be applied during the analysis.

It was suggested to modify the rate point selection by shifting the VTM rate points up by one step. This might allow for rate-matching in the low-bitrate range and for quality matching in the high-bitrate area. Thereby, both, rate savings and quality improvement could be demonstrated using the same curves.

It was suggested to use the latest version of the VTM when re-encoding the test sequences.

Next steps:

* Resolving the blending issue for the CMP-variants (implies software integration, should be doable by the October meeting)
* Perform subjective assessment of the test sequences HarbourBiking2 and KiteFliteWalking2 with the improved blending process. (Can be done with the available bitstreams)
* Encoding of the test sequences with the latest versions of the HM and VTM test models
* Complete analysis of the currently available data (with more results from visual assessments to be added)
* Updated selection of rate points based on the results of this analysis
* Run experts viewing sessions for confirmation.

It was planned to conduct expert viewing during the meeting. As a first step, some additional VTM QP points need to be encoded. Mathias will contact the usual volunteers for that.

Verification test on 360° video could likely be conducted during the upcoming meeting cycle (if the pandemic situation allows).

It was suggested that the contributors of the blending method in the 360lib software should submit an input contribution describing their approach. This was later registered as JVET-T0118.

[JVET-T0097](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10494) Report on VVC compression performance verification testing in the SDR UHD Random Access Category [M. Wien (RWTH), V. Baroncini (VABTECH)] [late]

This document reports the first set of verification test results comparing VVC to its predecessor, HEVC. The compression performance capabilities of the two specifications are compared based on the HEVC reference software HM-16.22, and two VVC implementations, the VVC reference software VTM-9.0 and the open-source VVC implementation VVenC-0.1.0. For HEVC, the random-access configuration provided with the configuration file cfg/encoder\_randomaccess\_main10.cfg of HM-16.22 was used. For the VTM, the random-access configuration provided with the configuration file cfg/encoder\_randomaccess\_vtm\_gop32.cfg of VTM-9.0 was employed. The results of a visual assessment of VVC compared to HEVC by naïve test subjects is reported. The assessment included five SDR UHD test sequences encoded in random-access configuration with a random-access interval of 1.07s. The measured MOS figures indicate a significant improvement of VVC over HEVC for both VVC implementations, VTM-9.0 and VVenC-0.1.0, resulting in an overall average bitrate saving of -43% and -49%, respectively. It was noted that the VVenC-0.1.0 software has been used at a medium configuration resulting in more than 100 times faster runtime than the VTM encoder, while having an overall average bitrate saving of about 50% of VVC over the HM.

It was commented that having graphs that have only VTM vs. HM, as well as graphs with all three encoders, could be desirable.

It was noted that the purpose of including a third encoder in the test is justified in order to show that a real-world encoder can obtain a quality similar to that shown by the reference software (and even somewhat better).

Another suggestion was to compute Bjontegaard deltas within a limited range of MOS values – e.g., the range of 5 or 6 to 8 or 8.5, to avoid comparing where quality is too poor to be reasonable in an application and to avoid the quality saturation range.

It was agreed to produce a report of these tests as an output.

[JVET-T0106](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10503) AHG4/AHG7: Report on Display Settings for HDR/WCG Verification Tests [A. Segall, M. Wien, V. Baroncini] [late]

[Add summary]

Discussed briefly in session 20 at 2020.

Detailed presentation was not considered necessary. This was provided for information and should be considered in future VT planning.

A participant commented that this was valuable information.

## Coding studies and tools on specific use cases (0)

## Test material (3 - completed)

[JVET-T0060](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10439) AHG4: new video conference sequences for HD verification testing [S. Xu, R.-L. Liao, J. Chen, Y. Ye (Alibaba)]

[Add summary]

Discussed briefly in session 20 at 2010.

To be studied in AHG toward potential VT usage. Test plan for the HD VT to be finalized in January.

[Add summary]

Discussed briefly in session 20 at 2010.

To be studied in AHG toward potential VT usage. Test plan for the HD VT to be finalized in January.

[JVET-T0082](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10461) New HLG sequences for high bit-depth video coding [K. Kondo, M. Ikeda, A. Browne (Sony)]

[Add summary]

Discussed briefly in session 20 at 2010.

Discussed in HBD BoG (add sections cross-references).

## Conformance test development (1 - 1 revisit)

[JVET-T0100](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10497) AHG5: On gaps in conformance bitstreams [F. Bossen (Sharp)] [late]

This was discussed in session 15 at 2145 Tuesday 13 October (chaired by GJS & JRO).

An independent implementation of a VVC decoder is used to analyse bitstreams submitted for conformance testing. Several gaps in the conformance set are identified, including syntax elements that never appear, syntax elements that exhibit little variety, and combinations of syntax element values that are never exercised.

Some bitstreams were in the pipeline for consideration, but had not yet been used in this testing.

The reported coverage issues were categorized as follows:

* Syntax elements not present in any bitstream
* Syntax elements always taking same value in every bitstream (when present)
* Syntax elements with limited range use
* Combination of pps\_rect\_slice\_flag = 0, pps\_loop\_filter\_across\_tiles\_enabled\_flag = 1 and pps\_loop\_filter\_across\_slices\_enabled\_flag = 0 (see ticket #1372)
* Combination of sps\_palette\_enabled\_flag = 1 and sps\_entropy\_coding\_sync\_enabled\_flag = 1 (see section 9.3.2.7)

This was appreciated as good input toward increasing conformance bitstream coverage.

It was again discussed if the conformance part should better be informative (TR) – see previous discussion under AHG5 report. One concern that had been raised before was less visibility on the ITU site, as it would have a different number than the main spec. It was however commented that on the ITU site, the conformance tests could perhaps be called “test vectors”. Some experts expressed the opinion that implementers might consider an informative conformance spec less relevant. On the other hand, compliance with streams from a normative spec is not giving a full proof about a decoder’s compliance with all aspects of a spec. Generally, a TR would develop more dynamic development. In general, no serious concerns were raised against defining a non-normative conformance. Revisit for a decision (also for software).

## Software development (5 - closed)

[JVET-T0064](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10443) Addition of ALF filter strength control to VTM [K. Andersson, J. Ström (Ericsson)]

This was discussed in session 15 at 2245 Tuesday 13 October (chaired by GJS & JRO).

This contribution asserts that currently, there is no possibility to control the ALF filter strength in VTM, it is only possible to turn it on or off completely. This proposal suggests including a configurable ALF filter strength parameter to VTM to enable control of the ALF filter strength. The ALF filter strength parameter scales the magnitudes of the filter coefficients according to the parameter such that a value less than 1 decreases the magnitudes of the filter coefficients and a value larger than 1 increases the magnitudes of the filter coefficients. If the strength parameter is equal to 1 the current behaviour is unchanged.

The BDR performance on CTC with the ALF filter strength parameter equal to 0.75, e.g. reducing the filter strength by 25%, is as follows:

* AI: 0.27%, 0.25%, 0.31% (Y, U, V)
* RA: 0.33%, 0.10%, 0.16% (Y, U, V)
* LDB: 0.30%, 0.11%, 0.23% (Y, U, V)

Further decreasing the filter strength to 50% by using the ALF filter strength parameter equal to 0.5 has the following impact on BDR:

* RA: 0.90%, 0.74%, 0.89% (Y, U, V)
* LDB: 0.67%, 0.90%, 1.50% (Y, U, V)

It is claimed that a reduction of the ALF filter strength in some cases can improve the subjective quality of VTM for inter predictive coding.

It was commented that having some more automatic method of tuning the strength would be desirable. This proposal offers some ability for controlling this. Controlling based on slice type and separate control of luma and chroma and control over CCALF were also mentioned possibilities.

Decision (SW): Adopt into software (not used by default, not in CTC).

[JVET-T0109](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10506) Crosscheck of JVET-T0064 (Addition of ALF filter strength control to VTM) [N. Hu (Qualcomm)] [late] [miss]

[JVET-T0078](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10457) GDR Software [S. Hong, L. Wang, K. Panusopone (Nokia)]

This was discussed in session 15 at 2300 Tuesday 13 October (chaired by GJS & JRO).

This contribution includes the source code to enable GDR functionality for VVC and proposes adding the software in the VTM source code package. The software offers the following setups:

* Encoding input video sequences into GDR bitstreams using the VTM encoder with the following features:
  + Flexible GDR period configured through parameter set by user,
  + Similar number of bits per picture, implying the delay is as low as one frame interval,
  + Progressive intra refresh over a GDR period with even distribution of the forced intra areas over pictures within the GDR period using virtual boundary syntax in picture header,
  + Necessary (encoding) constraints on coding tools to prevent the leaks and
  + Exact match at the recovery point (or leak-free).
* Decoding the leak-free GDR bitstream using the VTM decoder.

A large amount of effort was involved in developing this. JVET-T0127 reports that it affects or adds about 10,000 lines of encoder code.

It was asked if the software is well written and whether it would be difficult to maintain.

The software was integrated in VTM 8. A proponent said they had prepared a new version based on VTM 10.

It was suggested to put this in its own fork for study in the reference software database.

It was suggested to consider drafting a report on how to use GDR in encoders.

Several participants commented that GDR capability is important and would be highly desirable to support in the software, at least eventually.

Decision (SW): It was agreed to add this as a fork for study in the reference software database (the latest version that is available).

It was further suggested to use the software to generate some conformance bitstreams for GDR, and the proponent agreed to do this.

The contributor was thanked for providing this software which will enable testing and experimentation with GDR capability. It was hoped that after a brief period of study it will be possible to merge this capability into the main branch.

[JVET-T0127](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10524) Crosscheck of JVET-T0078 (GDR Software) [J. Enhorn (Ericsson)] [late]

[JVET-T0081](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10460) AHG3: On StreamMergeApp design [E. Thomas (TNO)]

This was discussed in session 16 at 0900 Wednesday 14 October (chaired by GJS & JRO).

The StreamMergeAp allows to merge single-layer bitstreams into a multi-layer bitstream. This contribution discusses issues encountered when using the StreamMergeApp together with bitstreams generated with VTM 10.0.

Possible directions are proposed to solve these issues.

* At the time the StreamMergeApp was first developed, the VPS NAL was always written even for a single layer bitstream. This way, the StreamMergeApp could use the detection of a VPS NAL as a trigger to write the new VPS in the output bitstream. However, it has been observed that the writing of the VPS in a single layer bitstream is conditioned to the vps identifier being greater than 0 in VTM 10.0. Possible options are:
  + VTM encoder to write the VPS for bitstreams with one layer
    - It was encouraged that this should be possible. It would also help in generation of conformance bitstreams and functionality testing.
  + StreamMergeApp to determine when to write the VPS differently
* The StreamMergeApp must create a VPS for the output bitstream which is the similar task of the VTM encoder when creating a multi-layer bitstream. As a result, the VTM encoder executes functions to 1) initialise VPS parameters according to the encoding configuration (e.g. dependency between layers or not) and 2) derive remaining VPS parameters based on the conditions set in the previous step. Currently, the StreamMergeApp simply creates a VPS object and relies on the default values set in the object constructor and is thus outdated. As a result, it would be advantageous to at least reuse the logic implemented in the encoding library and possibly some part of the encoding app logic although the StreamMergeApp does not parse a config file as the encoding app does. Possible options are:
  + Transfer initialization logic of VPS to a generic place, which seemed suggested as preferable.
  + Duplicate initialization logic of VPS in StreamMergeApp

K. Suehring was commented that the VPS is the only PS where the application does some of the initialization of the content outside of the library, which seemed like an architectural problem. He said he had been considering doing some cleanup work in this area already.

It was commented that there are other problems, including the current syntax of the NAL unit header, decoder capability information syntax, and that start code emulation prevention bytes are not added back when rewriting. See also ticket #918.

Also, generally the tool can only work with particular bitstream combinations, and should be modified to check whether these conditions are present.

Work to fix these problems was encouraged, in coordination with the software coordinators.

[JVET-T0118](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10515) AHG6: Blending with padded samples for GCMP [L. Lee, J.-L. Lin (MediaTek), Y. Wang, Y. He, L. Zhang (Bytedance)] [late] [miss]

This was discussed in session 16 at 0635 Wednesday 14 October (chaired by GJS & JRO).

The blending at decoder side for generalized cubemap projection format (GCMP) is proposed. To reduce the seam artefacts at face boundaries caused by the encoding, the padding is usually applied for encoding at the packed frame boundaries and the discontinuous face boundaries. In this proposal, those padded samples are used for the blending before rendering the viewport extracted from the decoded 360-degree video. The simulation results show that it can achieve -0.21%, -0.31%, and -0.31% BD rate saving for Y, U, and V components in terms of End-to-End WS-PSNR, respectively. The seam artefacts are claimed to be reduced visibly.

It was commented that although the PSNR effect is minor, the visual effect is substantial, and seems to resolve the problem that had been detected in preparation of the verification test to a large extent – the visibility of face boundaries is largely reduced.

The software capability was appreciated.

It was asked whether this is implemented with geometric correction. It is just overlapping with linear blending based on sample lines.

It was asked how, at the encoder side, the padding area is generated.

It was commented that most likely the face boundaries could be made even less visible with a more sophisticated scheme, and further study should be done regarding how the padded areas are generated at the encoder side and how the overlap is performed at the decoder side, and that such study should ideally be conducted before the verification testing is performed.

Decision (SW): Adopt JVET-T0118: Include the implemented blending scheme in the 360Lib software, and consider using it in verification testing. If an improved encoder-side padding can be performed, the encoding should also be modified (this is a matter separate from JVET-T0118, which affects only the decoder).

## Implementation studies (4 - 1 revisit)

[JVET-T0061](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10440) Multi-threaded VTM decoder description and performance analysis [S. Gudumasu, T. Poirier, F. Urban, F. Hiron, R. Jullian, P. de Lagrange (InterDigital)]

This was discussed in session 16 at 0705 Wednesday 14 October (chaired by GJS & JRO).

This contribution describes a VTM-based VVC decoder implementation using parallel processing. The design takes advantage of both task and data parallelization to distribute the decoding modules on a general-purpose multi-core processor. The reported average decoding time ratios compared to the VTM10.0 decoder are 12% and 14% for Class A and Class B sequences respectively, in Random Access test conditions using 16 decoding threads.

The contributor said they could provide their software in a branch of the reference software codebase.

The software includes code cleanup aspects as well as parallelism / speed optimizations.

The information about this decoder optimization was appreciated.

It was suggested that it be published in a fork to be made available for study before being merged into the master version.

Decision (SW): It was agreed to add this as a fork for study in the reference software database (the latest version that is available).

Software was uploaded to vcgit.hhi.fraunhofer.de/delagrangep/VVCSoftware\_VTM/-/tree/JVET-T0061.

[JVET-T0095](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10492) Performance of a VVC software decoder [B. Zhu, S. Liu, X. Xu, X. Zhang, C. Gu, L. Wang, W. Feng (Tencent)] [late]

This was discussed in session 16 at 0650 Wednesday 14 October (chaired by GJS & JRO).

An independent VVC software decoder implemented by Tencent (nicknamed O266) is demonstrated in this contribution. The performance of such a decoder is provided by decoding the CTC bitstreams generated by VTM-10 encoder. The following average performance is reported for CTC RA configuration:

* 48.5 fps can be achieved on an 8-core 8-thread (without overclocking) processor when decoding 4K bitstreams, a 21X speedup over the VTM decoder;
* 182 fps can be achieved on an 8-core processor when decoding full HD (1080p) bitstreams;
* 259 fps can be achieved on an 8-core processor when decoding full HD (1080p) SCC bitstreams.

A video player was implemented based on the presented VVC decoding capability and can display VVC bitstreams in real-time. This player is built based the open source software of FFmpeg/VLC and comes with basic play/pause control. It was reported that this player will use the GPL license and will be released to download soon at <https://mmedia.tencent.com/vvc-player>.

The performance results that were shown were for an x86 architecture.

It was noted that high speeds were also reported in a contribution to the previous meeting.

The information about this decoder optimization was appreciated.

[JVET-T0099](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10496) Open optimized VVC encoder (VVenC) and decoder (VVdeC) implementations [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 (HHI)] [late]

This was discussed in session 16 at 0735 Wednesday 14 October (chaired by GJS & JRO).

In September 2020 ( two months after the first version of VVC was finalized), Fraunhofer HHI released the source code for its optimized VVC encoder (VVenC) and decoder (VVdeC) on GitHub. For VVenC, the goal was to demonstrate comparable or even subjectively better coding efficiency than VTM at much lower runtime. In addition, VVenC provides real-world encoder features like rate control and multi-threading.

Without QP adaptation for subjective optimization and 6 threads the following PSNR-based YUV BD-rates and speedup factors compared to HM and VTM are reported for four different presets:

* Faster HD: -11.0%, 20x (HM), 160x (VTM) UHD: -15.9%, 36x (HM), 270x (VTM)
* Fast HD: -25.5%, 16x (HM), 130x (VTM) UHD: -28.8%, 26x (HM), 200x (VTM)
* Medium HD: -34.4% , 8.5x (HM), 69x (VTM) UHD: -37.4%, 14x (HM), 110x (VTM)
* Slow HD: -37.9% , 2.7x (HM), 22x (VTM) UHD: -40.6%, 5x (HM), 38x (VTM)

With QP adaptation for subjective optimization and 6 threads, the following MS-SSIM-based YUV BD-rates and speedup factors compared to HM and VTM are reported for four different presets:

* Faster HD: -18.4%, 20x (HM), 160x (VTM) UHD: -15.4%, 36x (HM), 270x (VTM)
* Fast HD: -28.4% , 16x (HM), 130x (VTM) UHD: -29.3%, 26x (HM), 200x (VTM)
* Medium HD: -37.5%, 8.5x (HM), 69x (VTM) UHD: -39.2%, 14x (HM), 110x (VTM)
* Slow HD: -40.2%, 2.7x (HM), 22x (VTM) UHD: -42.1%, 5x (HM), 38x (VTM)

VVdeC has been developed with the goal to be able to playback 10bit UHD video at 60 frames per second.

The proponent said they were working toward a more permissive licence for the software release (but probably not contribute improvements back into VTM due to resourcing).

In an encoding setting with the same speed as HM, the encoder had approximately the same PSNR compression performance as the VTM.

In another encoding setting with the same speed as HM, the encoder had better MS-SSIM compression performance as the VTM.

MCTF capability was included, which was reported to be with improved parallelism support. (It was asked whether we should enable this for VTM and HM CTC, as it provides about 3-4% gain and was said to improve “pumping” effect.) See also JVET-T0103 for notes regarding MCTF. The proponent said they might be able to contribute the improved parallelism to the VTM. It was commented that the segment-based parallelism operation should be checked before enabling this in the CTC.

The MCTF processing was further discussed in session 17 at 2015 (chaired by GJS and JRO).

After side activity to check on this, it was reported that there was no parallelism problem and was agreed to enable MCTF in the future CTC and VT work (at least for test sequences without scene cuts and without screen content and without 360° projection mapping).

A fix for MCTF interaction with LMCS is needed before MCTF can be used. This fix should be in the next release.

Discretion is given to the test coordinators regarding whether to apply MCTF in the tests. We would not re-run the SDR UHD tests because of this – only change usage in future work. Ideally, the tool should auto-detect and disable the processing under circumstances like scene cuts. It was commented that the bilateral filtering applied in the MCTF process may, at least somewhat, avoid scene cut and screen content problems.

It was commented that MCTF has no effect below QP value 17.

Revisit after side activity to think about MCTF and generate some test results (consider 360° and HDR, consider VT versus CTC, and which CTC output documents are affected, will it be used with HBD/HBR?).

The test results were without rate control.

The decoder did not yet support subpictures.

The availability of this implementation was greatly appreciated.

[JVET-T0103](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10500) Information on and analysis of the VVC encoders in the SDR UHD verification test [C. Helmrich, B. Bross, J. Pfaff, H. Schwarz, D. Marpe, T. Wiegand (HHI)] [late]

This was discussed in session 16 at 0835 Wednesday 14 October (chaired by GJS & JRO).

This contribution provides further information on the configuration of the VVC reference encoder (VTM) and the alternative encoder (VVenC) evaluated in the VVC verification test for SDR UHD content, specified in JVET-S2009 (test plan, draft 3). Based on an analysis of the bitstreams and Bjøntegaard delta-rate values obtained using several video quality assessment (VQA) methods, the following conclusions can be drawn:

* JVET-S2009 specifies VTM version 9.0 for the SDR UHD verification test, but the decoded video files resulting from newer VTM 10.0 encoded bitstreams were found to be identical to those resulting from the VTM 9.0 encodings. Thus, the SDR UHD verification test was, effectively, conducted using the VTM 10.0 software release, which was the latest VTM version as of late September 2020.
* The results of the SDR UHD verification test, reported in JVET-T0097, indicate that VVenC 0.1.0 (VTM 10 decoder compliant) tends to outperform the VTM encoder in terms of visual quality. It was found that only some of the VQA methods used for Bjøntegaard delta-rate evaluation – namely, the XPSNR, SSIM, and MS-SSIM methods – succeed in modelling this tendency with satisfactory accuracy. The PSNR and VMAF metrics, in particular, fail to model the subjective judgments well.

The VVenC encodings include optimizations for subjective video quality and encoder-only efficiency improvements, namely,

* Signal and in-GOP frame index dependent perceptual QP adaptation (QPA) based on the published JVET documents JVET-H0047, JVET-K0206, and JVET-M0091, QPA related ALF encoder changes JVET-L0181, and more recent XPSNR related enhancements to the spatiotemporal QPA algorithm, as further described in the contribution.
* Encoder-only deblocking optimization (EDO) accounting for the effect of deblocking on the recons­tructed-video signal distortion within the encoding parameter search-loop of each coded frame,
* Encoder-only GOP based motion compensated temporal filtering (MCTF) as a denoising pre-filter to improve the motion prediction performance and thus provide objective and subjective gains.

See further information in the contribution document and also the notes for T0069 regarding “VMAF” having different models and different versions.

A plug-in for the XPSNR (version 1.0, initial release, as a plug-in for FFmpeg) is available at <https://github.com/fraunhoferhhi/xpsnr/tags>.

It was commented that there is a mode decision feature in the VTM for partially accounting for deblocking in the mode decision and that this is not enabled in the current CTC, and perhaps it should be (see JVET-M0428, reporting 0.7% for RA). It was mentioned that this is not available in the HM, and was suggested that we should not use it in VTM-to-HM comparisons unless it is available in both HM and VTM, as it could hypothetically apply to both contexts.

It was commented that part of the subjective benefit could be from the MCTF, which was not enabled in the VTM test cases of the verification test. It is further commented that this design of the test was by purpose, as the non-normative configuration of HM and VTM was designed as similar as possible (under the constraints and tools implemented in the respective standard). If it was enabled in VTM, it should also be enabled in HM.

The proponent agreed that some of the visual benefit might be due to MCTF, but also said that in his opinion QP adaptation also provided a substantial benefit in some cases.

## Complexity analysis (2 - closed)

[JVET-T0067](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10446) Bit Stream Feature Analyzer (BSFA) for Coding Tool Statistics based on VTM-10.0 [M. Kränzler, C. Herglotz, A. Kaup]

This was discussed in session 18 at 0545 Thursday 15 October (chaired by GJS & JRO).

This contribution presents an analyzer add-on for the VTM-10.0 decoder. For any VVC-conforming bitstream, the analyzer counts the occurrence of all major coding tools and coding modes used for improving the compression efficiency. Examples are tools like intra prediction, transformations, residual coefficients, and in-loop filtering operations on slice level, block level, or pel level. The authors report that the tool is lightweight in terms of complexity and that the results of the analyzer can be used for further interpretation and assessment of coding tools. Furthermore, results can be used for novel encoding techniques targeting energy efficient coding or mode decisions based on machine learning.

The tool reportedly integrates into the decoder and increases decoder runtime by about 4%. It contains about 2000 lines of code.

The proponent was reportedly working on adding decoder energy usage estimation (see prior contribution JVET-Q0052 and Green-MPEG prior standard). Also see JVET-T0068.

Use in no-reference quality estimation was suggested (cf. ITU-T P.1204.3).

The source code, aligned to VTM 10.0, is available for download under the BSD License at <https://gitlab.lms.tf.fau.de/LMS/vtm-analyzer>.

Information was requested for studying capabilities relative to commercial bitstream analysers, e.g. from Elecard and VICUE.

It was suggested to use the tool to analyse the conformance bitstreams and use it to identify gaps in conformance testing coverage.

Further study was encouraged, and potential integration into the VTM main branch could be considered in the future.

[JVET-T0068](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10447) Decoding Time and Energy Assessment of VTM-10.0 and VVdeC [M. Kränzler, C. Herglotz, A. Kaup]

This was discussed in session 18 at 0615 Thursday 15 October (chaired by GJS & JRO).

Also see JVET-T0067.

This contribution evaluates the decoding energy and time demand of VVC in relation to HEVC. The authors compare several implementations of both video codecs in terms of decoding energy and time. The evaluation reportedly reveals that the decoding energy demand of the VTM-10.0 decoder has a significant increase compared to HM-16.20, which depends on the coding configuration. For random access, it is reported that the decoding energy is increased by 87% on average. The authors also report an investigation of one practical software decoder implementation for VVC (VVdeC) and HEVC (OpenHEVC).

The VVdeC and OpenHEVC implementations were compared to the reference software, with substantially reduced power consumption reported for the optimized implementations. In the group discussions, energy relationships to clock rate and multithreading were discussed.

It was asked whether the VVdeC (using 12 threads) and OpenHEVC analysis used similar bit rates and numbers of threads.

It was commented that such analysis is very platform dependent, so that the reported estimates are not generally true - e.g. for hardware implementations or hybrid hardware/software implementations. Another participant said that although there is platform dependency in such estimates, such analysis can also be relevant - especially when special purpose hardware is not available on the platform.

## Encoder optimization (2 - completed)

See also the notes on MCTF.

[JVET-T0062](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10441) AHG10: Extension of rate control to support random access configuration with GOP size of 32 [F. Liu, Z. Liu, Y. Li, Z. Chen (Wuhan Univ.)]

This contribution was presented in session 19 at 0920 on Thursday 15 October (chaired by GJS and JRO).

This contribution provides an extension of the rate control in VTM to support the configuration with the GOP size of 32 as described in JVET-S0180. With this proposed extension, rate control reportedly works well on the configuration with GOP size of 32.

The scheme follows JVET-M0600 as the basic design principle.

The experiments were based on VTM 10.1rc1 (git hash: 16bc143c) with a bug fix for rate control and conducted on HPC with Linux platform (ticket #1095).

Some test results were still missing.

It was commented that having the actual software available for study would be helpful.

It was thus reported that we basically just don’t have a functioning rate control for GOP size 32. The software coordinator was given the discretion to further consider this.

[JVET-T0116](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10513) Crosscheck of JVET-T0062 (AHG10: Extension of rate control to support random access configuration with GOP size of 32) [Y. Wang (Bytedance)] [late] [miss]

[JVET-T0074](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10453) AHG10: LMCS encoder algorithm for YUV4:4:4 content [F. Pu, T. Lu, P. Yin, S. McCarthy, W. Husak, T. Chen (Dolby)]

This contribution was presented in session 19 at 0845 on Thursday 15 October (chaired by GJS and JRO).

This contribution proposes an automatic non-normative encoder optimization algorithm for LMCS to improve coding performance for SDR YUV 4:4:4 video content. In VTM 10.0, the default LMCS encoder algorithm is optimized for YUV 4:2:0 content. However, YUV 4:2:0 and YUV 4:4:4 signals have very different characteristics. Similarly, YUV 4:4:4 camera content and screen content have different characteristics. The YUV 4:4:4 LMCS encoder algorithm proposed in this contribution provides the following coding performance for SDR YUV 4:4:4 (reference: LMCS off; test: LMCS on) under VTM 10.0 in terms of BD rate and processing time:

* For camera content (YUV 4:4:4):
  + AI: { -0.32%, 0.69%, 1.04% }, Enc: 104%, Dec: 101%
  + RA: { -1.13%, 2.98%, 0.20% }, Enc: 104%, Dec: 100%
* For screen content (YUV 4:4:4):
  + AI: { -1.22%, 2.89%, 3.08% }, Enc: 104%, Dec: 101%
  + RA: { -1.64%, -0.01%, 0.16% }, Enc: 101%, Dec: 100%

In the anchor, LMCS is off. This is not our current CTC anchor, which has LMCS on.

The encoding algorithm makes a different setting for two test sequences, which previously did not work so well (BirdsInCage and Robot), based on histogram data and chroma variance.

The algorithm operates differently for AI and RA, and classifies screen versus camera content.

It was noted that while luma is improving, chroma is degrading in some of these test results. For 4:4:4, two-thirds of the data is chroma, so chroma is more important than in the 4:2:0 case.

There seemed to be some improvement for the screen content RA case.

It was commented that in the AI case the bit usage was sometimes substantially increasing with the modification.

It was commented that studying the actual code would be helpful.

Further study was encouraged, especially regarding the reference for comparison, the luma/chroma balance, and the generality of the method.

[JVET-T0098](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10495) Crosscheck of JVET-T0074:AHG10: LMCS encoder algorithm for YUV4:4:4 content [J. Chen (Alibaba)] [late] [miss]

## Profile/tier/level specification (1 - 1 revisit)

[JVET-T0065](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10444) Conformance points for multilayer 8K [F. Bossen (Sharp)]

Was presented in joint session with parent bodies on Tuesday – see notes in sec. 8.2.

Initial comments 12 Oct 1915 - No level supporting 4K base layer and 8K enhancement layer. Could change the Multilayer Main profile, add levels, or add profile(s).

Agreed to establish such an additional level - details to be worked out [revisit]

# Low-level tool technology proposals (22)

## AHG12: High bit depth coding for VVC (11)

### Test conditions and test sequences

See also JVET-T0082 in section 4.6 on 12 bit 4K HLG test sequences.

[JVET-T0047](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10426) AHG12: Preliminary common test conditions for high bit depth video coding [T. Ikai, A. Browne, X. Xiu (AHG coordinators)]

This contribution was discussed at 0545 on Thursday 8 October (chaired by GJS & JRO).

This document proposes common test conditions (CTC), conversion practices and software reference configurations to be used in the context of high bitdepth video coding experiments. The object of this contribution is to provide a comparison point for the JVET-T meeting and the basis for a CTC for future meetings. The co-chairs seek the opinions of experts in a number of areas which will allow the direction of further work to be determined. The co-chairs also call on the JVET community for it’s help in providing and identifying new test sequences.

v2 revised SVT md5 value.

v3 fixed the template sheet (PQ422 EncT/DecT) and added VTM 10.0 anchor (ExtendedPrecision flag = 0)

AHG12 considers that a useful CTC should contain the following types of content:

* 12 bit HDR (PQ and HLG) sequences
* 16 bit RGB sequences
* 12/16 bit Domain-specific (e.g. medical) sequences
* High frame rate sequences (>= 100Hz)

In addition, the RExt CTC contained a series of tests for synthesized content involving the addition of MSB to the content. The co-chairs exressed interest in the opinion of other experts on the relevance of such tests to the new CTC.

AHG 12 has reviewed the HEVC RExt CTC (JCTVC-P1006) and the VVC HDR CTC (JVET-S2011), and the AHG suggested:

* Using 12 bit versions of the PQ sequences from the HDR CTC (convert original half float exr to 12bit PQ by HDRTools)
* Keeping the existing SVT 16 bit RGB sequences from the RExt CTC
* Removing all other sequences (until it is confirmed that theyare both available for use and of value to the new CTC).

This and other aspects recommended by the AHG were agreed by JVET unless noted otherwise.

It was commented that there are some 4:0:0 test sequences from Siemens that are in the JCT-VC repository with some different password as medical content, but they are not listed in JVET-T0025, and were reported to be quite old, and their copyright status should be reviewed.

M. Wien was suggested as a potential additional contact for medical content.

DICOM was suggested as a potential source of HBD content.

It was commented that some of the EBU material was also at 200fps. There is even a clip at 300fps

And one at 1000fps

https://tech.ebu.ch/docs/testmaterial/ebu\_test\_sequences\_tech\_info\_170315.pdf

See MPEG contribution M36666.

It was commented that we have some 100 Hz PQ content that have been treated as 50 Hz in other experiments. Another participant said the conversion from EXR for that content may have had some problem with the use of inadequate dynamic range.

It was agreed to produce a CTC output document.

There were reported to still be considerable gaps in the types of sequence in the initial CTC:

* HLG sequences
* Medical or other domain-specific sequences
* High frame rate sequences

Experts were encouraged to fill these gaps and there was hope that new sequences will be made available in submissions to the JVET-T meeting.

In the discussion, there was also interest expressed in camera-phone-type content.

The latest VTM 10.0 should be used and the following macro shall be enabled for both anchor and test

#define JVET\_R0351\_HIGH\_BIT\_DEPTH\_ENABLED 1

Results are to be reported using the following template: JVET-T0047\_AHG.xlsm

The template uses 6 point bdrate calculation in JVET-H0033.

The following tests should be conducted.

* Intra only for 12 bit and 16 bit
* Random access for 12 bit
* Low-delay B for 12 bit and 16 bit

Note: In the all intra configuration, TemporalSubsampleRatio shall be set equal to 8.

The following configuration files provided in the cfg/ folder of the VTM, the common software package, shall be used.

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

These configuration files specify the QP offsets for intra and inter predicted frames.

Note: For IntraPeriod to reflect the intra refresh period in the random access test cases. The intra refresh period is dependent on the frame rate of the source and the GOP size in use: a value 16 shall be used for sequences with a frame rate equal to 20fps, 32 for 24fps, 32 for 30fps, 48 for 50fps, 64 for 60fps, and 96 for 100fps.

It is recognized that some of these tests, particularly for low delay, may prove to be . The opinions of experts are sought

In addition, the following config file or tool settings shall be specified per class (VTM only).

|  |  |
| --- | --- |
| Sequence class | config file |
| SVT | cfg/per-class/formatRGB.cfg |
| PQ422 | cfg/444/yuv444.cfg  and  cfg/per-class/classH1.cfg |
| PQ444 | cfg/444/yuv444.cfg  and  cfg/per-class/classH1.cfg |

Initial tests on VTM 10.0 have shown that at present LMCS shows significant losses for 12 bit processing of PQ context, and will require a different configuration for 12 bit processing than for 10 bit. For at least the JVET T meeting LMCS shall be disabled as follows:

--LMCSEnable=0

It is hoped that at subsequent meetings an alternative configuration of LMCS will be made available and allow the tool to be re-enabled. More generally AHG12 believes collaboration with other experts on a series of tool-on / tool-off tests will be beneficial.

Revisit / BoG to further discuss.

### Tools and transforms

[JVET-T0084](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10463) AHG12: Transform coefficients range extension for high bit-depth coding [T. Zhou, T. Chujoh, E. Sasaki, T. Ikai (Sharp)]

This contribution was discussed at 0630 on Thursday 8 October (chaired by GJS & JRO).

This contribution proposes a format range extension design for high bit-depth coding. In the VVC versin1, the transform coefficient range is always equal to 15 regardless of bit-depth. It is asserted that this design is not sufficient in high bit-depth coding.

This proposal proposes to extend BitDepth + 5 for high bit-depth coding to argue this value would be sufficient and it is consistent in 10 bit-depth case. It is reported that the experimental results show BD-rate gain relative to having the extended precision flag equal to 0 is 1.03 %, 1.05 %, 1.02 % in AI-12, LDB-12, RA-12 (qp = -13, -8, -3, 2, 7, 12) and 14.83 %, 9.72 %, 8.82 % in AI-16, LDB-16, RA-16 (qp= -33, -28, -23, 18, 13, 8) respectively.

It is also reported that this high precision design shows comparable or better coding performance in 12, 14 and 16 bit cases compared to BitDepth + 6 design.

HEVC and VTM use BitDepth + 6.

It was commented that some results show coding loss with increased precision, which may indicate some other problem in the scheme.

Note: The current VTM software provides an extension for transform coefficient range when JVET\_R0351\_HIGH\_BIT\_DEPTH\_ENABLED macro is enabled in build and the extended precision flag is equal to 1 where BitDepth + 6 is used for transform coefficient range instead of 16.

It was commented that a prior theoretical study in the HEVC context had indicated that BitDepth +5 or BitDepth + 6 are theoretically needed in that context.

It was suggested to consider tool on/tool off tests to identify whether there are some coding tools that are not designed properly or not helpful with high bit depths.

It was commented that BitDepth + 5 does not affect the 10 bit case, but BitDepth + 6 does.

The proposal also adjusts bdShift1 and bdShift2 in specific ways, and may preserve more precision throught the first stage by doing this.

It was commented that we should not consider something that introduces a different code path or different logic for lower bit depths.

Further study was recommended on this.

[JVET-T0104](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10501) AHG12: Cross-check of JVET-T0084: Transform coefficients range extension for high bit-depth coding [L. Pham Van (Qualcomm)] [late]

[JVET-T0087](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10466) AHG12: VVC coding tool evaluation for high bit-depth coding [T. Ikai, T. Zhou, T. Hashimoto (Sharp)]

This contribution was discussed at 0650 on Thursday 8 October (chaired by GJS & JRO).

This contribution checks VVC coding tools in high bitrate range. It is reported that MTS, CCLM, DEPQUANT shows more than 0.5 % gain, i.e. 0.80 %, 1.51 %, and 1.41 % gain in average bdrateY in AI-12 (QP= -13, -8, -3, 2, 7, 12) respectively. It is also reported LFNST shows 0.47% loss in terms of coding efficiency.

It is suggested to define common test conditions for high bit-depth coding based on experiments. And it is also recommended experts to study coding tools to confirm if it is performed appropriately or it can be improved by adjustments.

v2 added encoder configuration test, which include CTU size and max MTT size configuration. It is proposed that a configuration with maxmtt size 16 shows better balance between encoding time and gain where the maximum loss with 0.19% loss and EncT is 61% in LDB-16, which changes the maximum encoding time from 3.6 days to 1.7 days.

Information about runtime versus coding gain was reported.

It was commented that LFNST and MTS are somewhat competing, and LFNST is typically not so helpful at high bit rates. This may be partly an encoder optimization matter.

The proponent proposed using some different configuration settings for low QP operation – reducing max block sizes, to speed up the experiments.

Some experiments used shortened test sequences (e.g. LD cases).

BoG was asked to consider this, and further study was encouraged.

[JVET-T0086](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10465) AHG12: SIMD implementation of BDOF for high bit-depth coding [T. Chujoh, E. Sasaki, T. Ikai (Sharp)]

This contribution was discussed in JVET session 11 at 2130 on Monday 12 October (chaired by GJS & JRO).

Since on the current configuration for high bit-depth coding, SIMD code is not applied, the encoding times are very huge. In the current BDOF processing, the internal precision exceeds 16-bit when the bit depth is greater than 12. In this contribution, by limiting the values of two variables, the internal precision of BDOF is within 16-bit and efficient SIMD codes can be implemented. As a result of the experiment, the coding efficiency was not changed, and the encoding and decoding times have been improved. The specification is only influenced by this change in the case of the bit-depth greater than 12.

The primary purpose of the contribution is not just to provide a SIMD implementation, but to alter the standard in a way that enables it.

It was commented that the reported runtimes are not comparing the impact of just the modified part of the design.

It was also commented that there are other parts of the computation that need to go beyond 16 bits, even for bit depths of 8−10. The proponent said that the proposed change is in a particularly important core part of the BDOF processing, while another participant said there are other more critical parts.

The encoding time was said to not be reliable in the low-QP test cases.

BDOF is not helping for coding 16 bit SVT test sequences at low QPs, although it helps at higher QPs.

Further study was encouraged, including more broad study of other places in the design that have bit depth issues.

[JVET-T0114](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10511) Crosscheck of JVET-T0086 (AHG12: SIMD implementation of BDOF for high bit-depth coding) [Y. Kidani, K. Kawamura (KDDI)] [late] [miss]

[JVET-T0091](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10470) AHG12: Fix on encoder overflow issue of the LMCS at high bit-depth coding [X. Xiu, H.-J. Jhu, Y.-W. Chen, T.-C. Ma, X. Wang (Kwai)]

This contribution was discussed in JVET session 11 at 2155 on Monday 12 October (chaired by GJS & JRO).

In this contribution, an encoder overflow problem was identified in the current VTM-10 reference software for the calculation of the LMCS parameters of SDR content when the internal bit-depth is 16-bit. Encoder fixes are provided in the contribution to resolve such overflow issue. Simulation results reportedly show compared to the VTM-10.0 16-bit anchors, the average {Y, U, V} BD-rate performance of the proposed fixes are {-0.88%, 0.40%, 0.12%}, {-1.31%, 1.61%, 1.61%} and {-0.55%, 2.3%, 2.51%} for AI, RA and LDB configurations, respectively.

Some test cases had such a substantial problem that they were left out of the experiment.

A corresponding merge request was to be filed.

This is only an encoder issue, not affecting operation up to a bit depth of 12 bits.

Decision (SW): The contribution was appreciated and agreed to be incorporated in the reference software. (A document contribution is not really necessary for such bug fix assistance.)

[**JVET-T0112**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10509) **Crosscheck of JVET-T0091 (AHG12: Fix on encoder overflow issue of the LMCS at high bit-depth coding) [F. Pu, T. Lu (Dolby)] [late] [miss]**

[JVET-T0093](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10472) AHG12: On signalling for high-bit depth [Y. Kidani, K. Kawamura, K. Unno (KDDI)]

This contribution was discussed in JVET session 11 at 2205 on Monday 12 October (chaired by GJS & JRO).

This contribution proposes a signalling for high-bit depth video coding in VVC. In the previous meeting, JVET-S0243 presentation slides reported that coding gains of some tools are reduced or lost as QP decreases in high-bit depth video sequences. JCCR is one of the reported tools, and the verification effort can be reduced if JCCR is disabled in the low QP range for high-bit depth coding. In this contribution, the additional applicable condition of JCCR with the pre-defined QP threshold is proposed against VVC ver.1. With this condition, JCCR can be disabled when block based QP is equal to or smaller than the pre-defined QP threshold. Simulation results under JVET-T0047 reportedly show the coding loss is negligible without one PQ test sequence (Market33Clip4000r2).

It was noted that such a change would need to be conditioned on high bit depth (or something else that differs from version 1 usage).

It was also noted that losses were observed in some test results (Market3Clip4000r2).

The claimed benefit is reduced conformance testing for high bit depths. For high-bit-depth decoders that decode lower bit depth bitstreams, it would still need to implement the feature.

Another participant said that since this introduces new conditions, and since JCCR needs to be supported for lower bit depths, it might actually complicate conformance testing.

No action was taken on this.

[JVET-T0107](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10504) Crosscheck of JVET-T0093 (AHG12: On signalling for high-bit depth) [T. Chujoh (Sharp)] [late] [miss]

### Golomb-Rice coding

[JVET-T0072](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10451) AHG12: Rice parameter selection for high bit depths [A. Browne, S. Keating, K. Sharman (Sony)]

This contribution was discussed in JVET session 11 at 2225 on Monday 12 October (chaired by GJS & JRO).

This proposal describes a modification to the method for the selection of Rice parameters for residual coding, for consideration in VVC version 2. The aim of the modification is to offer BD-rate gains when residual coefficient values become larger, e.g. for high bit depths and low QPs, whilst offering negligible losses at VVC version 1 operating points. It is suggested that the modification extends the operating point of the existing residual coding method to be applicable at all currently considered bit depths. Gains are rported for all-intra of -26.09%/-21.90%/-20.57% for the proposed 16 bit CTC, -1.67%/-1.55%/-1.57% for the proposed 12 bit CTC, -0.14%/-0.19%/-0.22% for low QP and 0.02%/-0.02%/0.01% for CTC. A simplification was developed which reduces the coupling between the use and update of computed values within the modification, whilst aiming to maintain the efficiency gains. Results for the simplification are -26.60%/-22.07%/-20.74% for 16 bit, -1.50%/-1.37%/-1.38% for 12 bit and -0.03%/-0.08%/-0.11% for low QP.

Applied to both TSRC and RRC

With that change, gain of approx. 25% in case of 16 bit (SVT, AI). Without that, 16 bit coding with VVC is significantly worse than HM. After applying that change gain approx. 1.75% over HM.

Even for 8-10 bit coding, the proponent reported that some benefit could be shown for extending the Rice parameter values in the low QP range.

To decode 8-10 bit video per VVC v1, this method operates differently from what is in VVC v1, but not much, since the same basic LUT approach is used.

The proposed algorithm impacts entropy coding and in particular the selection of Rice parameters for the Golomb-Rice coding of residual coefficients. The algorithm can be considered as having four components as follows:

* A coefficient size estimation function which is backed by a series of counters and an index function to select which counter is used for each coefficient.
* An update function for the selected counter applied after each coefficient.
* A modification to the current algorithm for computing the clipped sum of adjacent nearby coefficients.
* A modification to the current algorithm of selecting Rice parameters.

This would affect all transform coefficients, not just the first in a block. This is also applied to TSRC.

It is reported that there my be some latency problem (due to usage of a counter for “history” and a formula where that is used, and induces a shift on top of the current lookup table for Rice parameter derivation). This might inhibit parallelization. Contributors believe this could be easily avoided. Some ways to mitigate the sequential nature of the processing were described, in case there is concern about that.

[JVET-T0108](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10505) Crosscheck of JVET-T0072 (AHG12: Rice parameter selection for high bit depths) [T. Hashimoto, T. Ikai (Sharp)] [late] [miss]

[JVET-T0085](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10464) AHG12: Rice parameter derivation for high bit-depth coding [T. Hashimoto, T. Ikai (Sharp)]

This contribution was discussed in JVET session 11 at 2240 on Monday 12 October (chaired by GJS & JRO).

This contribution proposes to improve the derivation of rice parameter for regular residual coding (RRC) for high bit-depth coding. VVC derives Rice parameter in RRC with look-up table and its range is restricted from 0 to 3. It is asserted that this restriction causes worse performance of VTM in high bit-depth coding compared to HEVC Format Range extension profile where Rice parameter derivation can be extended when persistent\_rice\_adaptation\_enabled\_flag is equal to 1. To resolve this problem, this contribution introduces a Rice parameter derivation method using a predicted remainder value.

It is reported 1.52 % and 25.02 % in bdrateY in AI-12 qp= -13, -8, -3, 2, 7, and AI-16 -33, -28, -23, 18, 13, and 0.24 % and 16.95 % in bdrateY in LDB-12 qp= -13, -8, -3, 2, 7, and LDB-16 -33, -28, -23, 18, 13 respectively.

Conceptually similar with JVET-T0072, but uses a modified version of the current lookup table for Rice parameter derivation. This avoids the latency problem of JVET-T0072, but requires a second table to be used depending on bit depth.

The contributor said this scheme is consistent with the current design in VVC v1.

A difference is that this scheme does not use the current LUT approach and instead uses a formula.

(A previous editorial proposal for VVC v1 had a formula approach that produces the same results as the current LUT approach, so using a formula is possible with VVC v1.)

It was asked whether this has a backward compatibility issue. The proponent suggested having a new flag to select between the two methods.

JVET-T0072 has a history tracking, basically as in HEVC range extensions. This one does not have that.

It was commented that it would be good to study more test cases before selecting one.

[JVET-T0101](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10498) Crosscheck of JVET-T0085:AHG12: Rice parameter derivation for high bit-depth coding [Adrian Browne (Sony)] [late]

[JVET-T0089](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10468) AHG12: Slice based Rice parameter selection for transform skip residual coding [H.-J. Jhu, X. Xiu, Y.-W. Chen, T.-C. Ma, C.-W. Kuo, X. Wang (Kwai Inc.)]

This contribution was discussed in JVET session 11 at 2300 on Monday 12 October (chaired by GJS & JRO).

In residual coding of a transform skip block, for each sample, the remaining absolute level “abs\_remainder” is binarized using the fixed Rice parameter equal to 1. However, significant performance degradation is reported in VVC high bit depth configuration. It is observed that using different Rice parameters for different conditions (e.g. different sequences, different QPs, different bit depths) improves the coding efficiency especially for high bit depth conditions. In this proposal, it is proposed to explicitly signal the Rice parameter for each slice to indicate the Rice parameter for the binary codewords of abs\_remainder for TSRC.

As proposed, it would be possible to get the same results by just using a built-in setting based on bit depth rather than sending the Rice parameter in syntax.

This proposal is only affecting TSRC, while other proposals affect both ordinary residual coding and TSRC. The proponent said that some adjustment to regular residual coding is also desirable.

It was commented that we should also study whether encoder active optimization of the Rice parameter could provide some gain. Here the encoder did not take advantage of the flexibility proposed for the syntax.

[JVET-T0102](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10499) Crosscheck of JVET-T0089:AHG12: Slice based Rice parameter selection for transform skip residual coding [A. Browne (Sony)] [late]

[JVET-T0105](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10502) AHG12: On the Rice parameter derivation for high bit-depth coding [L. Pham Van, D. Rusanovskyy, M. Karczewicz (Qualcomm)] [late]

This contribution was discussed in JVET session 11 at 2310 on Monday 12 October (chaired by GJS & JRO).

This contribution proposes a modification to the Rice parameters selection method for regular residual coding (RRC), for consideration in VVC version 2. In current VVC, the Rice parameter has been extensively tested for encoding video sources of 8 bit-depth or 10 bit-depth and is limited in the range from 0 to 3 by clipping. It is asserted that the residual coefficient values become larger for high bit depths and low QPs. Therefore, the current restriction of the rice parameter range leads to performance degradation of VTM in high bit-depth coding. To address this issue, this contribution introduces a minor modification in the rice parameter derivation of VVC: a down scaling of the locSumAbs prior to rice parameter derivation.

It is reported -1.36% and 24.80% in bdrateY in AI-12 qp= -13, -8, -3, 2, 7, and AI-16 -33, -28, -23, 18, 13, and -0.12% and -0.26% in bdrateY in LDB-12 qp= -13, -8, -3, 2, 7, and LDB-16 -33, -28, -23, 18, 13 respectively.

This is proposing a change to regular residual coding only.

The proposal was said to retain the current design as a special case when the bit depth is 10 or lower.

Plan: To have a BoG (T. Ikai) to develop test conditions and a CE.

[JVET-T0119](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10516) Crosscheck of JVET-T0105 (AHG12: On the Rice parameter derivation for high bit-depth coding) [H.-J. Jhu (Kwai Inc.)] [late]

### Lossless RGB / YCgCo coding (1 TBP)

[JVET-T0111](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10508) YCgCo-R: Observations and findings [D. Buitenhuis (VideoLAN), A.M. Tourapis (Apple Inc)] [late] [miss]

## AHG11: Neural network-based technology (17)

### General (4)

[JVET-T0041](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10420) Methodology and reporting template for neural network coding tool testing [S. Liu, A. Segall, E. Alshina, J. Boyce, M. Wien, D. Grois]

[insert summary]

This was reviewed in the AHG meeting and was used for most of the incoming contributions.

This contribution was discussed at 0745 on Thursday 8 October (chaired by GJS & JRO).

It was agreed to produce a corresponding output after considering experiences with it.

Data was used from Univ. Bristol, which was posted publicly by them.

There was some discussion of soliciting comments from hardware experts about what complexity and gain target could be acceptable.

BoG work was suggested to come up with a group suggestion.

Revisit for review & agreement (e.g., SSIM aspects).

[JVET-T0042](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10421) Report of AHG11 Meeting on Neural Network-based Video Coding on 2020-07-30 [S. Liu, E. Alshina, J. Pfaff, M. Wien, P. Wu, Y. Ye]

[insert summary]

This was not presented separately from the AHG 11 report.

[JVET-T0117](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10514) AHG11: Methodology additional requirements [F. Galpin, T. Dumas, P. Nikitin, F. Racapé, F. Le Léannec, E. Francois (InterDigital)] [late]

Discussed in session 13 Tue 13 Oct. 0720 (chaired by JRO & GJS)

Document JVET-T0041 describes the recommended JVET methodology and reporting template for neural network (NN) coding tool testing. This contribution suggests amendments to this document related to the reported complexity information template, in order to better evaluate the complexity of proposed NN-based video coding tools.Proposes unique way of reporting worst case MAC/picture and MAC/pixel

Also proposes describing number of cores when reporting runtimes

Memory should be described taking into account the number of models

Suggests an assessment of memory bandwidth – but no concrete way of doing it

Generally agreed – continue improving the template JVET-T0041 in BoG.^

The MPEG DNNVC AHG had a similar template (CTC plus GPU memory usage), with slightly different test conditions (4 QP points)

What would be needed would be some common form that would help understanding the overall network topology, also including possibility of different architectures.

BoG (organized by A. Segall) to work on these documents

* Unified CTC (both end-to-end and hybrid) – rate range rather than QP in case of end to end
* Reporting template where proponents would report about the general architecture/topology of their proposals (including placement in the codec), and provide relevant complexity parameters. This should start from hybrid, and could later be extended to end-to-end approaches
* Training methodology: Using a common training methodology would be useful to make technology comparable (e.g. in EE)

BoG should also define an exploration experiment on loop filters, as a test case to learn how to assess NN technology (which would be the major purpose of this EE).

DNNVC technical documents and the AHG report should be registered as JVET documents. Currently, the AHG had recommend “further study” on these. Revisit to discuss further action.

[JVET-T0129](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10526) [DNNVC] Comments on Common Test Conditions and Reporting Template [E. Alshina, A.Segall, R.-L. Liao, T.Solovyev]

See BoG JVET-T0130

### “Hybrid” concepts (8)

[JVET-T0073](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10452) AHG11: Neural Network-based intra prediction with transform selection in VVC [T. Dumas, F. Galpin, P. Bordes, F. Leleannec (Interdigital)]

This contribution was discussed at 0805 on Thursday 8 October (chaired by GJS & JRO).

This document describes a new neural network-based intra prediction mode inside VVC. This neural network-based intra prediction mode predicts a given block from neighboring reference samples located above and on the left side of this block. Besides, it infers both the LFNST kernel index and whether the primary transform coefficients resulting from the application of the DCT-2 horizontally and the DCT-2 vertically to the residue of the neural network prediction are transposed, which completely defines the LFNST in case LFNST is selected.

When integrated into VTM-8.0, for the all-intra configuration, -3.36%, -2.95%, and -2.97% of Y BD-rate gain, U BD-rate gain, and V BD-rate gain respectively are reported. In this configuration, the encoding and decoding runtimes of VTM-8.0 with the neural network-based intra prediction mode with respect to VTM-8.0 are 395% and 3575% respectively.

For the random-access configuration, -1.52%, -1.00%, and -1.26% of Y BD-rate gain, U BD-rate gain, and V BD-rate gain respectively are reported. In this configuration, the encoding and decoding runtimes of VTM-8.0 with the neural network-based intra prediction mode with respect to VTM-8.0 are 159% and 723% respectively.

The prediction uses 8 bits of precision, and the result has a mean added and is multiplied by 4 with clipping to produce a 10-bit prediction.

LFNST kernel selection was modified to use a selection from the NN. LFNST kernel unchanged from VVC.

The overall architecture consists of a processing chain where reference areas from left and above are first processed separately and then merged for prediction and LFNST index derivation. For small blocks, a fully connected network (3 layers) is used, for large blocks a convolutional network (6 layers).

BD-rate and coding time in the inference stage when the neural network-based intra prediction mode infers the LFNST to be applied

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All intra over VTM-8.0** | | | | |
|  | BD-rate (Y) | BD-rate (U) | BD-rate (V) | encoding time (CPU) | decoding time (CPU) |
| Class A1 | -4.89% | -4.16% | -4.61% | 366% | 3007% |
| Class A2 | -2.35% | -2.08% | -2.28% | 411% | 2955% |
| Class B | -2.99% | -2.59% | -2.50% | 429% | 3897% |
| Class C | -2.71% | -2.25% | -2.38% | 364% | 3916% |
| Class E | -4.30% | -4.14% | -3.57% | 400% | 3943% |
| **Overall** | -3.36% | -2.95% | -2.97% | 395% | 3575% |
| Class D | -3.02% | -2.63% | -2.71% | 348% | 4010% |
| Class F | -1.91% | -1.35% | -1.49% | 234% | 1844% |
|  | **Random access over VTM-8.0** | | | | |
|  | BD-rate (Y) | BD-rate (U) | BD-rate (V) | encoding time (CPU) | decoding time (CPU) |
| Class A1 | -2.06% | -1.00% | -1.87% | 150% | 761% |
| Class A2 | -1.04% | -0.42% | -0.55% | 150% | 577% |
| Class B | -1.42% | -1.16% | -1.17% | 160% | 593% |
| Class C | -1.61% | -1.24% | -1.44% | 173% | 1057% |
| Class E |  |  |  |  |  |
| **Overall** | -1.52% | -1.00% | -1.26% | 159% | 723% |
| Class D | -1.19% | -0.41% | -1.22% | 160% | 731% |
| Class F | -0.78% | -1.08% | -0.45% | 155% | 445% |

BD-rate and coding time in the inference stage when the neural network-based intra prediction mode predicts the LFNST to be applied (this gives better performance)

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra over VTM-8.0** | | | | | | | | | |
|  | BD-rate (Y) | | BD-rate (U) | BD-rate (V) | | encoding time (CPU) | | | decoding time (CPU) | |
| Class A1 | -5.14% | | -4.01% | -4.74% | | 404% | | | | 3047% |
| Class A2 | -2.35% | | -2.39% | -2.46% | | 435% | | | | 2942% |
| Class B | -3.11% | | -2.74% | -2.72% | | 492% | | | | 3792% |
| Class C | -2.82% | | -2.43% | -2.55% | | 441% | | | | 5153% |
| Class E | -4.48% | | -4.01% | -3.29% | | 494% | | | | 4704% |
| **Overall** | -3.49% | | -3.04% | -3.07% | | 455% | | | | 3889% |
| Class D | -3.16% | | -2.73% | -2.80% | | 426% | | | | 4941% |
| Class F | -1.96% | | -1.75% | -1.40% | | 268% | | | | 2046% |
|  | **Random access over VTM-8.0** | | | | | | | | | |
|  | BD-rate (Y) | BD-rate (U) | | BD-rate (V) | encoding time (CPU) | | | decoding time (CPU) | | |
| Class A1 | -2.14% | -0.94% | | -1.79% | 167% | | 880% | | | |
| Class A2 | -1.09% | -0.22% | | -0.67% | 161% | | 497% | | | |
| Class B | -1.48% | -1.30% | | -0.93% | 173% | | 547% | | | |
| Class C | -1.62% | -1.12% | | -1.30% | 194% | | 992% | | | |
| Class E |  |  | |  |  | |  | | | |
| **Overall** | -1.57% | -0.96% | | -1.15% | 175% | | 692% | | | |
| Class D | -1.30% | -0.63% | | -1.68% | 181% | | 788% | | | |
| Class F | -0.81% | -1.43% | | -0.48% | 167% | | 440% | | | |

It was commented that there were similar gains for a previous HHI proposal (which was later stripped down for lower complexity and became MIP).

One question was about the number of reference lines used in this scheme. The maximum was 16 to the left and 16 above. Some cases use downsampling.

A second question was how much of the coding gain comes from the LFNST part. About 0.5 % AI.

It was noted that this uses fairly large model sizes in some cases. It was remarked that the bottleneck would be small blocks. Even though they use a lower data size for the hidden layers, the number of operations per sample is larger than for large blocks.

The handling of chroma was discussed. Chroma uses this mode when luma does. Adding signalling to control them separately would increase search requirements.

Further study was encouraged.

[**JVET-T0058**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10437) **AHG11: Information on inter-prediction coding tool with deep neural network [B. Choi, Z. Li, W. Wang, W. Jiang, X. Xu, S. Liu (Tencent)]**

This contribution was discussed at 0840 on Thursday 8 October (chaired by GJS & JRO).

This informative contribution reports some preliminary results of deep neural network (DNN) utilization for inter-prediction. The idea is inserting of a new (virtual) reference picture, which is generated by inference processes of trained networks, into reference picture list (RPL). The entire networks consist of several sub-network models for edge-detection, optical flow estimation/compensation and detail enhancement. Since networks were trained with small size training sequences, classes C&D show coding gains of 2.18%, 3.93% respectively in luma for RA, in comparison to other classes.

This basically uses the concept of frame-rate up-conversion (FRUC). A virtual reference picture is generated by NN inference and is inserted into the RPL. The extra picture is treated as a LTRP in the DPB, with MVs set to 0.



Proposed neural network process for generation of virtual reference picture

BD-rate and Coding Time for NN-based Video Coding Tool Testing in Inference Stage   
(Tested 33frames with GOP=32 and patch size 512x512)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access Over VTM-10.0** | | | | | | |
| BD-rate (Y) | BD-rate (U) | BD-rate (V) | Encoding Time (CPU) | Decoding Time (CPU) | Encoding Time (GPU) | Decoding Time (GPU) |
| Class A1 |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |
| Class B | -0.68% | 1.43% | 1.52% |  |  | 133.82% | 45103.39% |
| Class C | -2.18% | -0.27% | 1.37% |  |  | 116.60% | 34643.88% |
| Class E | O | O | O | O | O | O | O |
| **Overall** |  |  |  |  |  |  |  |
| Class D | -3.93% | 4.27% | 6.51% |  |  | 128.06% | 59024.62% |
| Class F |  |  |  |  |  |  |  |

Largest gain for class D. Was trained on small resolution video. However, even for class D numbers for chroma seem to indicate loss.

The entire processing chain consists of edge detection, optical flow estimation/compensation, and detail enhancement by NN.

The model was reportedly trained using lower resolution video, and the proponent said it might do better for higher resolution if trained for that case. There could be a library of such separately tuned models.

The work was said to be somewhat preliminary, and provided for information.

It was commented that in VVC we have made tools operate, e.g., on a 4x4 level rather than a single-sample level, and some of the gain is probably just coming from allowing sample-by-sample processing rather than block-based processing.

It was suggested to consider the interaction with other coding tools such as BDOF and DMVR, and with higher-complexity variations of those other methods (which were exercised e.g. in JEM) as anchors.

Adding other anchor methods of reference picture generation was suggested.

The proponent said that this scheme does not seem to suffer from blurring that is found in other methods.

Further study was encouraged.

[JVET-T0057](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10436) AHG11: A case study to reduce computation of a neural network based in-loop filter by pruning [C. Auyeung, W. Wang, W. Jiang, X. Li, S. Liu (Tencent)]

This contribution was discussed at 0900 on Thursday 8 October (chaired by GJS & JRO).

Main target of this contribution is not an improved technology, but demonstrating for an example how the complexity could be reduced by pruning.

In general, the method of zeroing out weights, i.e. pruning, is applicable to reduce the number of multiply-adds in a convolutional neural network. To study the effect of pruning to coding efficiency, this contribution uses the dense residual convolutional neural network based in-loop filter (DRNLF) from JVET-O0101 as a case study. In this case study, the DRNLF was integrated with VTM-10.0 and tested according to testing condition in JVET-T0041, with and without pruning of 25% of the filters in the convolution layers.

Without 25% filter pruning, the DRNLF resulted in BD-rate of -0.96% / -0.77% / -0.81% / -1.31% for RA/LDB/LDP/AI configuration.

With 25% filter pruning, BD-rate became -0.79% / -0.64% / -0.65 % / -1.07%, respectively.

In general, 25% pruning can lead to a reduction of 25% of the number of multiply-add in a convolutional neural network based coding tool and save 25% of neural network parameters.

This is done as an additional stage of filtering between deblocking and SAO.

Results with pruning

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access Over VTM-10.0** | | | | | | |
| BD-rate (Y) | BD-rate (U) | BD-rate (V) | Encoding Time (CPU) | Decoding Time (CPU) | Encoding Time (GPU) | Decoding Time (GPU) |
| **No Pruning** | -0.96% | -3.88% | -3.82% | 105% | 3334% |  |  |
| **25% Filter Pruning** | -0.79% | -3.70% | -2.82% | 105% | 3244% |  |  |
|  | **Low delay B Over VTM-10.0** | | | | | | |
| BD-rate (Y) | BD-rate (U) | BD-rate (V) | Encoding Time (CPU) | Decoding Time (CPU) | Encoding Time (GPU) | Decoding Time (GPU) |
| **Overall** | -0.77% | -4.88% | -4.64% | 107% | 3904% |  |  |
| **25% Filter Pruning** | -0.64% | -4.58% | -4.17% | 106% | 3848% |  |  |
|  | **Low delay P Over VTM-10.0** | | | | | | |
| BD-rate (Y) | BD-rate (U) | BD-rate (V) | Encoding Time (CPU) | Decoding Time (CPU) | Encoding Time (GPU) | Decoding Time (GPU) |
| **Overall** | -0.81% | -5.06% | -5.12% | 110% | 4370% |  |  |
| **25% Filter Pruning** | -0.65% | -4.95% | -4.49% | 109% | 4157% |  |  |
| (Optional) | **All Intra Over VTM-10.0** | | | | | | |
| BD-rate (Y) | BD-rate (U) | BD-rate (V) | Encoding Time (CPU) | Decoding Time (CPU) | Encoding Time (GPU) | Decoding Time (GPU) |
| **Overall** | -1.31% | -2.69% | -3.38% | 108% | 3431% |  |  |
| **25% Filter Pruning** | -1.07% | -2.79% | -2.74% | 108% | 3418% |  |  |

Without pruning

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access Over VTM-10.0** | | | | | | |
| BD-rate (Y) | BD-rate (U) | BD-rate (V) | Encoding Time (CPU) | Decoding Time (CPU) | Encoding Time (GPU) | Decoding Time (GPU) |
| Class A1 | -1.05% | -3.35% | -5.38% | 107% | 4473% |  |  |
| Class A2 | -1.30% | -4.22% | -2.58% | 106% | 3201% |  |  |
| Class B | -0.74% | -4.13% | -3.56% | 105% | 2772% |  |  |
| Class C | -0.93% | -3.70% | -3.89% | 104% | 3475% |  |  |
| Class E | O | O | O | O | O | O | O |
| **Overall** | -0.96% | -3.88% | -3.82% | 105% | 3334% |  |  |
| Class D | -1.05% | -2.34% | -3.28% | 104% | 2011% |  |  |
| Class F | -0.31% | -1.70% | -1.74% | 109% | 1476% |  |  |
|  | **Low delay B Over VTM-10.0** | | | | | | |
| BD-rate (Y) | BD-rate (U) | BD-rate (V) | Encoding Time (CPU) | Decoding Time (CPU) | Encoding Time (GPU) | Decoding Time (GPU) |
| Class A1 | O | O | O | O | O | O | O |
| Class A2 | O | O | O | O | O | O | O |
| Class B | -0.51% | -5.33% | -4.53% | 106% | 4346% |  |  |
| Class C | -0.58% | -6.43% | -6.70% | 103% | 4978% |  |  |
| Class E | -1.44% | -2.06% | -2.05% | 114% | 2361% |  |  |
| **Overall** | -0.77% | -4.88% | -4.64% | 107% | 3904% |  |  |
| Class D | -0.91% | -3.85% | -4.63% | 103% | 2783% |  |  |
| Class F | -0.33% | -3.29% | -2.35% | 111% | 2247% |  |  |
|  | **Low delay P Over VTM-10.0** | | | | | | |
| BD-rate (Y) | BD-rate (U) | BD-rate (V) | Encoding Time (CPU) | Decoding Time (CPU) | Encoding Time (GPU) | Decoding Time (GPU) |
| Class A1 | O | O | O | O | O | O | O |
| Class A2 | O | O | O | O | O | O | O |
| Class B | -0.53% | -5.78% | -5.15% | 108% | 4951% |  |  |
| Class C | -0.59% | -6.58% | -6.53% | 105% | 5279% |  |  |
| Class E | -1.57% | -1.83% | -3.17% | 120% | 2759% |  |  |
| **Overall** | -0.81% | -5.06% | -5.12% | 110% | 4370% |  |  |
| Class D | -0.98% | -4.43% | -4.12% | 105% | 3552% |  |  |
| Class F | -0.37% | -3.64% | -3.49% | 112% | 2216% |  |  |
| (Optional) | **All Intra Over VTM-10.0** | | | | | | |
| BD-rate (Y) | BD-rate (U) | BD-rate (V) | Encoding Time (CPU) | Decoding Time (CPU) | Encoding Time (GPU) | Decoding Time (GPU) |
| Class A1 | -0.96% | -2.05% | -4.46% | 114% | 3405% |  |  |
| Class A2 | -1.22% | -2.70% | -1.95% | 108% | 2886% |  |  |
| Class B | -0.89% | -2.91% | -3.50% | 106% | 3037% |  |  |
| Class C | -1.58% | -3.63% | -4.24% | 104% | 3353% |  |  |
| Class E | -2.11% | -1.69% | -2.36% | 108% | 5192% |  |  |
| **Overall** | -1.31% | -2.69% | -3.38% | 108% | 3431% |  |  |
| Class D | -2.05% | -2.09% | -4.53% | 103% | 2982% |  |  |
| Class F | -0.67% | -2.16% | -2.07% | 104% | 2180% |  |  |

The method was trained on intra only.

Without pruning: Roughly 22,500 MAC per sample, pruning reduces by 25%. Pruning removes entire channels (filters). This also incurs some drop in performance

It was commented that the number of MACs per sample was still quite high (~16k per sample). It was commented that the original schem was using about 100k MACs per sample for the smallest blocks.

It was commented that the memory size is one of the major concerns in hardware implementation. Unlike the more “random” pruning in fully connected networks, this approach would give benefit in memory size.

It was commented that memory size is quite important, and the pruning reduces the memory size as well as the number of operations.

The basic idea was to encourage work in the direction of considering pruning and other such methods of complexity reduction.

It was asked whether precision control had also been considered. This was all done using floating point. Some other prior contributions had discussed that; this one focuses on structural pruning.

It was commented that it would be desirable to investigate complexity/performance drops for more cases (more aggressive pruning).

Here the pruning is more helpful than in some other schemes, since it is done in a systematic way to remove entire filters rather than individual weights.

Among the proposed schemes, this one was said to be simpler than most others.

Further study was encouraged.

[JVET-T0069](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10448) AHG11: SSIM based CNN model for in-loop filtering [T. Ouyang, F. Liu, H. Zhu, Z. Chen (Wuhan Univ.), X. Xu, S. Liu (Tencent)]

This contribution was discussed in session 9 at 2100 on Friday 9 October (chaired by GJS & JRO).

This contribution provides a CNNLF (convolutional neural network based in-loop filter) for VVC which focus on improving the subjective quality of encoded videos. The performance is evaluated using VMAF in the VTM-10.0. Simulation results report BD-rate savings for luma, and both chroma components compared with VTM-10.0 under AI, RA and LDB configuration.

The proposed CNNLF is introduced into VTM as an additional filter between deblocking filter (DF) and sample adaptive offset (SAO).

Simulation results report 7.50%, 6.01%, 2.30% BD-rate savings on VMAF under AI, RA and LDB configuration.

The first set of tables below use Loss = 0.8× “SSIM”+0.2×MAD for training, then measured using “VMAF”

“SSIM” = C \* (1 − SSIM), with some unknown scaled factor C, not using the multiscale SSIM.

The Y, U, V columns are PSNR.

“VMAF” here means what is done in FFMPEG. There are several VMAFs, different versions and different picture resolutions. A particular VMAF “model” was chosen. The contribution does not say which VMAF was used. Version 0.6.1 was used. The “model” and the version are not the same thing.

It was commented that there is a default model in VMAF.

Experimental results (All Intra)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | | |  |
|  | VMAF | Y | U | V | EncT-CPU | DecT-CPU | DecT-GPU |
| Class A1 | −5.84% | −1.39% | −3.72% | −4.11% | 170% | 41444% | 3363% |
| Class A2 | −7.22% | −1.56% | −7.54% | −8.19% | 137% | 21508% | 2276% |
| Class B | −6.51% | −0.67% | −7.97% | −6.85% | 131% | 21345% | 2043% |
| Class C | −8.11% | −2.03% | −8.53% | −7.34% | 117% | 21533% | 1852% |
| Class E | −10.29% | −2.86% | −4.65% | −5.29% | 139% | 37679% | 2804% |
| **Overall** | −7.50% | −1.61% | −6.76% | −6.46% | 136% | 26294% | 2331% |
| Class D | −9.26% | −2.70% | −7.64% | −8.16% | 115% | 19974% | 1984% |
| Class F | −3.08% | −0.34% | −4.70% | −3.50% | 115% | 20602% | 1826% |

Experimental results (Random Access)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main 10** | | | | | |  |
|  | VMAF | Y | U | V | EncT-CPU | DecT-CPU | DecT-GPU |
| Class A1 | −4.66% | −0.67% | 3.94% | 4.45% | 145% | 51787% | 4174% |
| Class A2 | −7.14% | −1.06% | −8.02% | −9.09% |  | 27635% | 2639% |
| Class B | −6.02% | −0.47% | −7.98% | −5.34% | 141% | 12545% | 1132% |
| Class C | −6.17% | −0.60% | −5.86% | −4.44% | 125% | 16030% | 1450% |
| Class E |  |  |  |  |  |  |  |
| **Overall** | −6.01% | −0.66% | −5.04% | −3.89% |  | 20826% | 1859% |
| Class D | −6.58% | −0.51% | −4.58% | −4.66% | 124% | 29333% | 2407% |
| Class F | −1.57% | 0.39% | −2.15% | −1.34% | 161% | 40522% | 3314% |

Experimental results (Low Delay B)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Low delay B Main 10** | | | | | |  |
|  | VMAF | Y | U | V | EncT-CPU | DecT-CPU | DecT-GPU |
| Class A1 |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |
| Class B | −2.11% | 0.14% | −8.86% | −6.42% | 137% | 14937% | 1307% |
| Class C | −3.07% | 0.55% | −8.36% | −6.76% | 126% | 24803% | 2043% |
| Class E | −1.59% | −0.16% | −3.85% | −3.07% | 181% | 24373% | 2228% |
| **Overall** | −2.30% | 0.20% | −7.44% | −5.69% | 143% | 19991% | 1733% |
| Class D | −1.82% | 0.49% | −6.83% | −6.70% | 123% | 29599% | 2480% |
| Class F | −0.15% | 1.52% | −3.02% | −1.66% | 173% | 66822% | 4844% |

Below tables are for a different loss weighting: Loss=0.2×SSIM+0.8×MAD.

Experimental results (All Intra)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra Main 10** | | | | | | |
|  | VMAF | Y | U | V | EncT-CPU | DecT-CPU | DecT-GPU |
| Class A1 | −4.68% | −1.79% | −5.88% | −6.16% | 156% | 22366% | 2423% |
| Class A2 | −4.41% | −1.94% | −8.81% | −7.77% | 130% | 33088% | 2789% |
| Class B | −3.96% | −1.48% | −8.85% | −8.31% | 124% | 23956% | 2163% |
| Class C | −4.48% | −2.40% | −8.47% | −8.37% | 115% | 20816% | 1829% |
| Class E | −6.68% | −3.28% | −5.47% | −7.27% | 130% | 34118% | 2790% |
| **Overall** | −4.72% | −2.11% | −7.70% | −7.70% | 129% | 25695% | 2312% |
| Class D | −6.69% | −3.14% | −7.24% | −9.38% | 112% | 19564% | 1939% |
| Class F | −1.82% | −0.78% | −4.59% | −3.92% | 114% | 9311% | 1190% |

Experimental results (Random Access)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Random access Main 10** | | | | | | |
|  | VMAF | Y | U | V | EncT-CPU | DecT-CPU | DecT-GPU |
| Class A1 | −4.89% | −2.40% | −7.31% | −6.02% |  | 38157% | 3240% |
| Class A2 | −4.04% | −2.29% | −11.11% | −9.08% | 131% | 25447% | 2202% |
| Class B | −3.54% | −1.53% | −10.26% | −8.49% | 128% | 27498% | 2425% |
| Class C | −2.97% | −1.49% | −7.12% | −6.61% | 120% | 22800% | 1961% |
| Class E |  |  |  |  |  |  |  |
| **Overall** | −3.76% | −1.84% | −9.00% | −7.61% |  | 27499% | 2382% |
| Class D | −3.96% | −1.73% | −6.69% | −7.44% | 120% | 13263% | 1292% |
| Class F | −0.76% | −0.37% | −2.79% | −2.47% | 151% | 5186% | 886% |

Experimental results (Low Delay B)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Low delay B Main 10** | | | | | | |
|  | VMAF | Y | U | V | EncT-CPU | DecT-CPU | DecT-GPU |
| Class A1 |  |  |  |  |  |  |  |
| Class A2 |  |  |  |  |  |  |  |
| Class B | −2.61% | −1.03% | −12.13% | −9.81% | 127% | 32380% | 2847% |
| Class C | −2.43% | −0.91% | −10.84% | −9.41% | 120% | 33890% | 2804% |
| Class E | −2.56% | −2.55% | −5.33% | −7.24% | 171% | 18160% | 1616% |
| **Overall** | −2.54% | −1.37% | −10.00% | −9.04% | 134% | 28450% | 2459% |
| Class D | −3.09% | −1.17% | −10.34% | −10.32% | 118% | 20621% | 1812% |
| Class F | −0.77% | −0.31% | −4.95% | −3.77% | 154% | 11838% | 1326% |

Thus, the first weighting is better for VMAF impact (in overall average), while the second one does better for MAD (although this was in JEM and a patch had been offered for VTM in MPEG DNNVC activity in June 2020). It was commented that it would be better to modify the VTM to use the implementation from HDRTools, which should not be difficult since something similar is done for HDR testing.

It was commented that there are different versions of SSIM as well (e.g. windowing type, windowing size, how to handle edge regions).

HDRTools has SSIM (with a Gaussian window) and MS-SSIM (without the Gaussian window); VTM does not.

Chroma was upsampled to feed to the NN and then downsampled at the output of that stage (staying in YUV).

A QP is an input to the model. Training used 5 QP values.

Network parameters are 32-bit float.

There was discussion of potentially just abandoning PSNR as a metric.

It was commented that it would be good to check whether visual quality is aligned with the analysis.

Another participant said that in some testing VMAF was not so stable - e.g., with deblocking filtering being harmful to VMAF despite being beneficial to visual quality. Some work has recently been conducted to develop new VMAF metrics to try to improve the behaviour.

Temporal consistency would also be good to study.

SSIM ignores chroma.

A participant emphasized that there is no 100% reliable metric in this world. VMAF, MS-SSIM, PSNR all have issues. It was suggested not to push or restrict any metric to be studied, but such study should not only include Excel table, but also proper viewing with MOS collection and computing correlation between MOS and all of the candidate metrics. The most difficult parts here would be collecting materials for viewing and conducting the viewing. The rest (metrics computation and correlation) is relatively easy. Without having those numbers in our hands, discussions such as which metric is better could be endless.

There was a question about how important is the “SE block”, and the proponent said they believed it to be quite important.

It was suggested to establish a BoG and to have a joint discussion with the parent bodies on the future NNVC exploration plans.

Further study on JVET-T0069 was encouraged.

[JVET-T0079](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10458) AHG11: Neural Network-based In-Loop Filter [H. Wang, M. Karczewicz, J. Chen, A. M. Kotra (Qualcomm)]

This contribution was discussed in session 9 at 2230 on Friday 9 October (chaired by GJS & JRO).

This contribution proposes a Neural Network-based In-loop filter (NN filter). The proposed NN filter adds an additional in-loop filter stage and is positionally placed after the Adaptive loop filter (ALF).

It is reported that average 4.11%, 13.83% and 13.28% BD rate saving for RA, and 3.63%, 9.47% and 9.98% BD rate saving for AI, for Y, Cb and Cr components respectively, are achieved.

Test results are shown with models refined for inter frames (where training included data from the coding loop).

*Simulation results of using 2 DIV2K models + 2 BVI-DVC refined models*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Random access Main 10** | | | | |
|  | **Over VTM-10.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -3.53% | -9.41% | -11.44% | 138% | 13880% |
| Class A2 | -4.21% | -15.13% | -13.49% | 135% | 12801% |
| Class B | -3.88% | -14.26% | -13.08% | 139% | 14983% |
| Class C | -4.09% | -13.56% | -13.21% | 128% | 13711% |
| Class E |  |  |  |  |  |
| **Overall** | -3.93% | -13.28% | -12.87% | 135% | 13964% |
| Class D | -5.05% | -14.64% | -15.36% | 132% | 14524% |
| Class F | -1.65% | -8.65% | -7.88% | 165% | 19218% |
|  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | |
|  | **Over VTM-10.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | ? | ? | ? | ? | ? |
| Class A2 | ? | ? | ? | ? | ? |
| Class B | ? | ? | ? | ? | ? |
| Class C | -3.91% | -17.00% | -16.78% | 127% | 14589% |
| Class E | -3.55% | -9.79% | -8.82% | 191% | 25849% |
| **Overall** | ? | ? | ? | ? | ? |
| Class D | -4.86% | -18.78% | -18.47% | 129% | 14197% |
| Class F | -1.82% | -10.58% | -9.78% | 166% | 23844% |
|  |  |  |  |  |  |
|  | **All Intra Main 10** | | | | |
|  | **Over VTM-10.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -3.16% | -6.52% | -8.94% | 160% | 10573% |
| Class A2 | -2.90% | -9.26% | -8.92% | 140% | 7506% |
| Class B | -3.22% | -10.10% | -10.33% | 134% | 7813% |
| Class C | -4.02% | -12.00% | -12.46% | 119% | 5873% |
| Class E | -5.01% | -8.20% | -8.24% | 141% | 8935% |
| **Overall** | -3.63% | -9.47% | -9.99% | 136% | 7834% |
| Class D | -4.32% | -11.63% | -13.77% | 117% | 6222% |
| Class F | -2.46% | -9.57% | -8.56% | 116% | 7159% |

*Simultaion results of using 6 DIV2K models + 6 BVI-DVC refined models*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Random access Main 10** | | | | |
|  | **Over VTM-10.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -3.85% | -9.38% | -10.75% | 203% | 13808% |
| Class A2 | -4.26% | -15.93% | -14.63% | 195% | 12877% |
| Class B | -3.99% | -14.73% | -13.44% | 207% | 14986% |
| Class C | -4.33% | -14.48% | -13.96% | 178% | 14049% |
| Class E |  |  |  |  |  |
| **Overall** | -4.11% | -13.83% | -13.28% | 196% | 14058% |
| Class D | -5.40% | -15.20% | -16.30% | 188% | 15153% |
| Class F | -1.68% | -9.88% | -9.29% | 280% | 19437% |
|  |  |  |  |  |  |
|  | **Low delay B Main 10** | | | | |
|  | **Over VTM-10.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | ? | ? | ? | ? | ? |
| Class A2 | ? | ? | ? | ? | ? |
| Class B | ? | ? | ? | ? | ? |
| Class C | -4.20% | -19.19% | -17.92% | 175% | 14636% |
| Class E | -3.66% | -11.01% | -9.43% | 362% | 26010% |
| **Overall** | ? | ? | ? | ? | ? |
| Class D | -5.31% | -20.55% | -19.82% | 184% | 14396% |
| Class F | ? | ? | ? | ? | ? |
|  |  |  |  |  |  |
|  | **All Intra Main 10** | | | | |
|  | **Over VTM-10.0** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | -3.18% | -6.48% | -8.86% | 278% | 10550% |
| Class A2 | -2.89% | -9.28% | -8.94% | 211% | 7516% |
| Class B | -3.22% | -10.10% | -10.34% | 196% | 7812% |
| Class C | -4.02% | -12.00% | -12.47% | 155% | 5995% |
| Class E | -5.01% | -8.20% | -8.25% | 220% | 8991% |
| **Overall** | -3.63% | -9.47% | -9.98% | 204% | 7877% |
| Class D | -4.32% | -11.64% | -13.77% | 151% | 6450% |
| Class F | -2.45% | -9.67% | -8.67% | 146% | 7194% |

The feature is designed to produce a residual difference to be added to the video content.

A CTU-level on/off is supported.

6 QP values were used for training.

A model selection is encoded.

It was commented that including QP in the model and using a single model would be preferred, to avoid needing to swap models.

In some variation, a picture-level selection of which model to use was performed.

4:2:0 to 4:4:4 conversion is avoided by separating the luma samples into four groups.

The proponent said the location of the filter was partly chosen to enable comparison to post filtering. The CTU-level on/off was not used in the post filtering case.

The proponent said that the scheme included a residue scaling factor and that this aspect turned out to be very useful and could be used in other approaches as well.

Most proponents put a NN filter between deblocking and SAO. This one put it after ALF. Another replaced SAO, deblocking and ALF with the proposed filter (for certain pictures).

Further study on JVET-T0079 was encouraged. It is pointed out that not only comparison of the benefits of different loop filter concepts, but potentially also combination of some of their elements might be worthwhile to study.

It was again suggested to organize a joint meeting with the parent bodies as early as possible during the next week, to clarify how to further conduct the explorations on NN video coding. It was also commented that the progress in the “hybrid” approaches (with regard to complexity vs. performance tradeoff, novel elements, and overall additional compression gain) has only been moderate since the time when the topic was studied in the last JVET AHG two years ago.

It was also mentioned that some other tools that were proposed in the CfP and were not adoptred to VVC due to complexity considerations had brought similar compression gains with comparably muh lower complexity impact than the NN tools.

The relation between the “end to end” approach investigated in the MPEG DNN AHG and the “hybrid” approach also needs further clarification. It is mentioned that quite a few of the discussions above are duplicated in JVET and the MPEG AHG.

Stopped here in session 9.

[JVET-T0088](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10467) AHG11: Convolutional neural networks-based in-loop filter [Y. Li, L. Zhang, K. Zhang, Y. He, J. Xu (Bytedance)]

This contribution was discussed in JVET session 12 at 0610 on Tuesday 13 October (chaired by GJS & JRO).

This contribution presents a convolutional neural network-based in-loop filtering method. The proposed CNN-architecture features a convolutional layer with a stride of 2 and several residual blocks to increase the valid receptive field and enable a smooth optimization. To avoid overfiltering, the proposed method is only applied on I-slices and those B-slices with temporal identifier equal to 0. Compared with VTM-9.0, the proposed method reportedly shows on average 8.24%, 15.72%, and 14.33% BD-rate reductions for Y, Cb, and Cr, respectively, under AI configuration, and on average 3.86%, 13.42%, 11.98% BD-rate reductions for Y, Cb, and Cr, respectively, under RA configuration.

Notes from discussion:

* Replaces all other loop filters.
* Proponents claim subjective benefit at low bit rates
* For I slices, partitioning information is fed into the network
* Different models for I and B slices, different models for each of the five “base QP”
* Decoding time increase (CPU) roughly 150x forAI, roughly 10x for RA.
* Different models are trained for the I slices and B slices and for each of 5 QP values.
* The proponent said they could see some improvement in visual quality.
* It was commented that the contribution should report total memory size rather than size per model.

[JVET-T0094](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10473) AHG11: In-loop filtering based on neural network [T.-C. Ma, W. Chen, X. Xiu, Y.-W. Chen, H.-J. Jhu, C.-W. Kuo, X. Wang (Kwai)] [late]

This contribution was discussed in JVET session 12 at 0625 on Tuesday 13 October (chaired by GJS & JRO).

A deep-learning based in-loop filter is proposed in this contribution. The proposed in-loop filter is located between the de-blocking and SAO stages. The implementation is based on VTM9.0, the average {Y, U, V} BD-rate performance of the proposed deep-learning based filter is reported as {-4.99%, -16.39%, -17.34%} and {-3.92%, -18.09%, -16.93%} for AI and RA configurations, respectively.

It was commented that the decoding times are extremely high. The proponent said this may be because they did not precompile the pytorch library. It was commented that the inference method may be a reason for high runtime.

[JVET-T0092](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10471) AHG9/AHG11: Neural network based super resolution SEI [T. Chujoh, E. Sasaki, T. Ikai (Sharp)]

This contribution was discussed in JVET session 12 at 0640 on Tuesday 13 October (chaired by GJS & JRO).

In this contribution, a framework for introducing a super-resolution post-filter using NN (Neural Network) is proposed. In VVC, RPR (Reference Picture Re-sampling) has been introduced. As a result of some experiments to change the resolution of the whole sequence, in several 4K sequences, there were some coding gains. In this experiment, the up-sampling filter is used the interpolation filter of VVC. If some super-resolution post-filters using NN are introduced, the visual quality will be more improved. For example, MPEG NNR which is under development in SC29/WG4 can describe neural networks efficiently. In this contribution, an SEI message which can encode and decode a neural network for super-resolution is proposed.

The contributor recommended that AHG studies should consider a method of using RPR.

There was some odd behaviour for chroma.

Visual quality investigation was recommended.

It was commented that there could be some denoising effect, and that it may be desirable to study use of the denoising prefilter in VTM.

It was further commented that there has been a bug in the temporal prefilter software’s interaction with another feature, which will soon be fixed.

It was commented that there was a University of Bristol contribution to the CfP, and perhaps another contribution at the CfP stage, that contained some similarities.

The “convex hull” style of comparison across resolutions was suggested to be considered.

See also T0096.

[JVET-T0096](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10493) AHG11: Deep neural network for super-resolution [J. Y. Lee (Sejong University), W. Lim, G. Bang (ETRI)] [late]

This contribution was discussed in JVET session 12 at 0655 on Tuesday 13 October (chaired by GJS & JRO).

This contribution proposes a deep neural network to upsample a low resolution image. To avoid excessive computational complexity, the proposed method cascades two CNN layers and trains residuals between the original and upsampled images. Experimental results show that it improves PSNR of the upsampled image with relatively small complexity. Therefore, it can be used for frame interpolation, such as reference picture resampling.

The testing was just for the upsampling stage of the process, reporting PSNR gains in the upsampling process. At low QP, the PSNR difference was up to 1.5 dB, although at high QP no gain was evident.

Rather preliminary work, demonstrating the benefit of NN based upsampling versus bicubic approach. Try to implement a low complex network for that. Works best at low QP, no gain at high. Training was done with uncoded material.

See also T0092.

[JVET-T0117](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10514) AHG11: Methodology additional requirements [F. Galpin, T. Dumas, P. Nikitin, F. Racapé, F. Le Léannec, E. Francois (InterDigital)] [late]

This contribution was discussed in JVET session 13 at 0730 on Tuesday 13 October (chaired by JRO & GJS).

It proposes some differences relative to T0041. [Consider moving notes next to T0041.]

### “End to end” concepts (5, TBP or BoG?)

Contributions in this category were initially submitted for consideration in MPEG’s AHG on DNNVC, and were duplicated as JVET documents after the parent bodies agreed that JVET should consider this topic.

[JVET-T0121](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10518) [DNNVC] AhG on deep neural networks based video coding [Y. Ye, E. Alshina, J. Chen, S. Liu, J. Pfaff, S. S. Wang (??)]

[JVET-T0122](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10519) [DNNVC] Experimental Results of End-to-end DL Compression in Different Colour Spaces [C. Ma, R.-L. Liao, Z. Wang, Y. Ye (??)] [late]

[JVET-T0123](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10520) AHG11: A study of handling YUV420 input format for DNN-based video coding [A. K. Singh, H. E. Egilmez, M. Coban, M. Karczewicz (Qualcomm)] [late]

[JVET-T0125](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10522) [DNNVC] Decomposition, Compression, Synthesis (DCS)-based Framework: A New Exploration on Video Coding [M. Lu, Z. Ma (Nanjing Univ.)] [late]

[JVET-T0128](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10525) [DNNVC] Preliminary results of Neural Network Loop Filter [Z. Wang, R.-L. Liao, C.Y. Ma, Y. Ye (Alibaba)] [late]

## Other (0)

# High-level syntax (HLS) proposals (17)

## AHG8: Layered coding and resolution adaptivity (2)

See also section 6.2.

[JVET-T0070](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10449) AHG9/AHG8: Scalability dimension SEI message [Y.-K. Wang, L. Zhang, K. Zhang, Z. Deng (Bytedance)]

See section 6.2.

[JVET-T0080](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10459) AHG9: Multilayered OLS decoded picture assembling SEI message [E. Thomas (TNO)]

See section 6.2.

## AHG9: SEI message studies and proposals (14, 2 revisit, 3 TBP)

See section 4.2 for errata considerations and section 8.1 for joint meeting discussion.

[JVET-T0049](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10428) Composite Picture Information SEI Message [Hendry, H. Jang, S. Kim, J. Lim (LGE)]

This contribution was discussed 2130 Wednesday 7 October 2020 (chaired by JRO & GJS).

This contribution proposes a design for a composite picture information (CPI) SEI message. A composite picture is a picture created / generated from rectangular patches from pictures in an access unit. Similar SEI messages were proposed by JVET-S0123 and JVET-S0107. It is asserted that the design proposed in this contribution differs from the previous contributions as follows:

* The proposed design works for composite picture from not only multi layers but also from a single layer bitstream.
* The proposed design provides more details about handling of issues related to gap and overlap of patches in composite picture.
* In the proposed design, the association between CPI SEI message and its associated output layer set (OLS) is made by carrying CPI SEI message in scalable nesting SEI message. It is asserted that this design is more aligned with how other SEI messages that are associated with OLS.
* The proposed design provides ways for reducing bit cost for signalling.

The design principles of the proposed CPI SEI message are provided in section 2 of this document.

A scene could be composited from pictures in different layers or from regions of pictures.

Examples were shown in two figures.

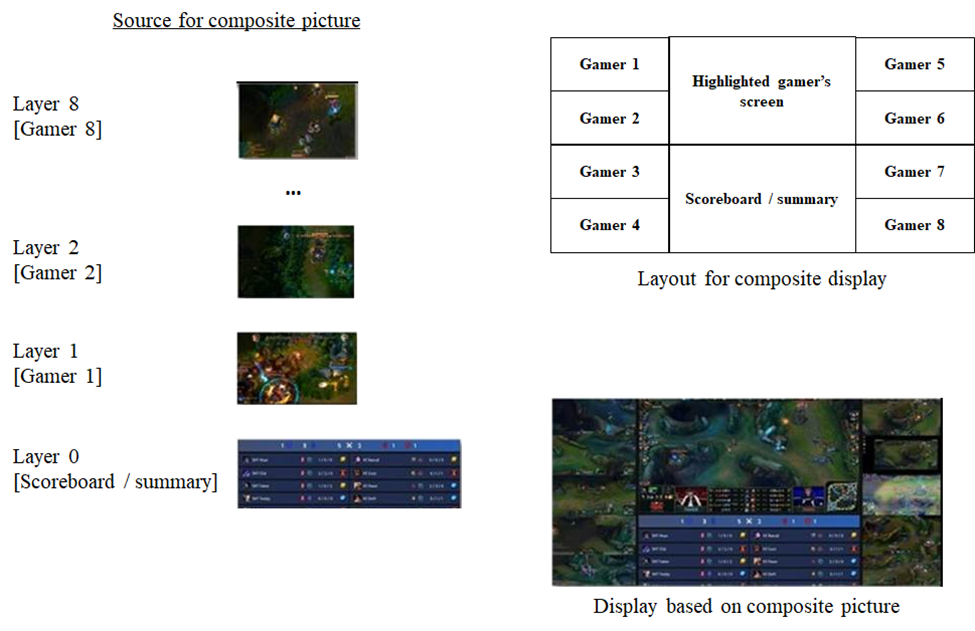


Figure – Use of composite picture in broadcasting / streaming online game program

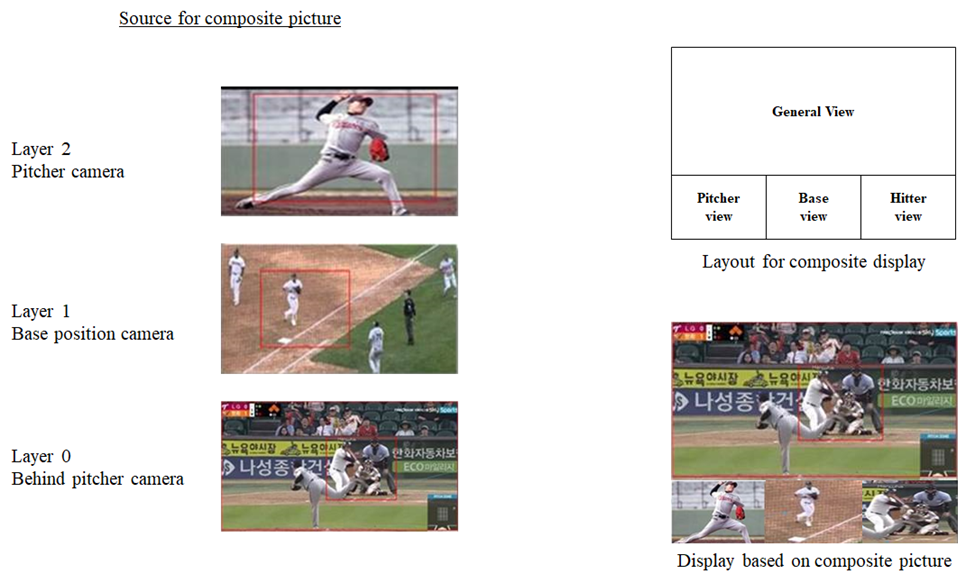


Figure – Use of composite picture in broadcasting / streaming live baseball game

For supporting composite picture based for multilayer bitstream, the following SEI message design principles were proposed:

1. Composite picture information (CPI) SEI specifies information to create / generate a composite picture from one or more pictures in an access unit.
2. A composite picture contains one or more rectangular shaped patches.

No particular relationship between the patches and subpictures or tiles was proposed.

In a previous related contribution, using only whole pictures was envisioned.

In this usage, some parts of the decoded pictures would not be used in the composite picture.

1. Patches of a composite picture belong to rectangular areas in pictures of the same access unit.
2. For each patch in a composite picture, the following information is signalled:
   1. The location of the patch in the composite picture. This information can be signalled as the position of the top-left corner of the patch in the composite picture.
   2. The size (i.e., width and height) of the patch in the composite picture.
   3. The source of the pels of the patch. This includes the following information:
      1. The layer id of the picture in the associated access unit.
      2. The location of the area in the picture that corresponds to the patch (i.e., top-left corner)
      3. The size of the area in the picture that corresponds to the patch.
3. The size of a patch in a composite picture shall be greater than 0 and the associated area (i.e., the source of the pels of the patch) in the associated picture shall be greater than 0.
4. There can be gap between two neighboring patches in a composite picture. In other words, there may be area in the composite picture that is not covered by any patch.
5. There can be overlap of patches in a composite picture. In other words, there may be area in the composite picture that is covered by two or more patches.
6. When there is overlap between two patches, the patch with higher index is rendered after the patch with lower index such that in the overlapped area between the two patches, the pels come from the patch with higher index.
7. There shall be no two patches with the same size that covered the same area in a composite picture. This is intended to avoid signalling redundant patch(es) as there is only one will be displayed / visible.
8. Two patches in a composite picture may have the same corresponding area in the same picture.
9. A patch in a composite picture shall correspond an area in a picture from an output layer only.
10. It is possible that a patch in a composite picture corresponds to an area in a picture from an output layer but the picture is not outputted by decoder (e.g., the PictureOutputFlag of the picture is derived to be equal to 0 due to certain condition). In such situation, implementer of the SEI message may skip copying the pels from the source area to the patch in the composite picture (e.g., a gray patch is shown in the composite picture).
11. A CPI SEI message describes composition for its associated OLS. The association is established by allowing CPI SEI message to be carried in a scalable nesting SEI message.
    1. When a scalable nesting SEI message contains CPI SEI message(s), the value of the value of sn\_ols\_flag shall be equal to 1.
    2. When a CPI SEI message is not present in a scalable nesting SEI message, it applies to the targetOLS (i.e., to the whole bitstream being decoded).
12. The CPI SEI message applies to AU A that contains the SEI and all AU that follow AU A in output order until one of the following is true:
    1. The end of the bitstream.
    2. The next AU contains a new CPI SEI message that applies to the same OLS.
13. There shall not be two CPI SEI messages associated with the same OLS but contain different content in an AU. This would create problem for implementer to choose which one is the correct one. Thus, it is constrained that when two CPI SEI message are present in an AU and they are associated with the same OLS, they shall have the same content.
14. When present, CPI SEI message is present in AU with TemporalId equal to 0 to avoid the SEI being removed during sublayer extraction.
15. When present in a CVS, the first of the CPI SEI message to be present in the CVSS AU.
16. Possible bit-saving for signalling:
    1. For signalling the position and size of the patch and size of the composite picture, instead of using ue(v) coding, signal the bit length for the coding and code those syntax elements using u(v).
    2. Signal a flag to specify whether scaling is allowed. When scaling is not allowed, the size of patch in the composite picture does not need to be signalled as it can be inferred to be equal to the size of the associated area in the source picture.
    3. Signal a flag to specify whether cropped picture is used for the composite picture. When the flag is equal to 0, the location of the associated area in the source picture does not need to be signalled as it can be inferred to be equal to the top-left corner of the picture.
    4. The size of the composite picture can be not signalled. This can be inferred from the position and size of patches that make up the composite picture.

A participant commented that using some externally filled cache image could also be considered.

The envisioned use would have all the pictures within one multilayer bitstream. It was commented that there can be use cases that don’t obey that assumption, and a system could provide such functionality.

Another participant commented that this is basically a scene description language as would be found in some system designs.

It was asked how to deal with extraction – e.g., if some pictures have been removed from the bitstream. There would be nesting in an OLS to prevent that.

There was also discussion of what happens if the composite picture references a non-output picture. The proponent suggested that filling such areas with flat grey could be an approach to handle that.

Multiparty videoconferencing was a described potential use.

Subpicture usage was mentioned as something that can provide a similar functionality. It was noted that subpictures do not support overlap and that gaps are more difficult using subpictures.

A proponent of a previous related proposal said that a layout based on whole pictures (first figure) could use in-place decoding to avoid needing an extra memory buffer.

Picture overlap with a precedence was proposed.

Basic questions were:

* Should such functionality be done in an SEI message or in a system layer
* If in an SEI message, how general should it be?

There was discussion of whether synchronization would be easier if this is done within a video bitstream or by a system.

It was suggested to consider this in a joint discussion with MPEG Systems and Requirements.

See also JVET-T0080 (and JVET-S2017).

See section 8.1 for joint meeting discussion.

[JVET-T0080](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10459) AHG9: Multilayered OLS decoded picture assembling SEI message [E. Thomas (TNO)]

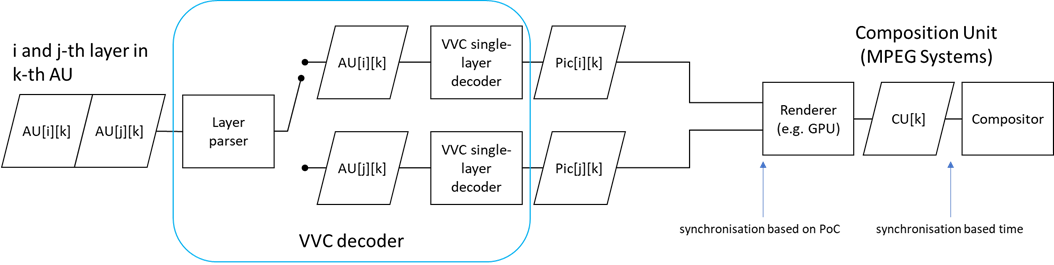
This contribution was discussed 2215 Wednesday 7 October 2020 (chaired by JRO & GJS).

This contribution presents a motivation for a combination of the decoded pictures from different layers based on PoC value as opposed to time-based combination for composition use cases.

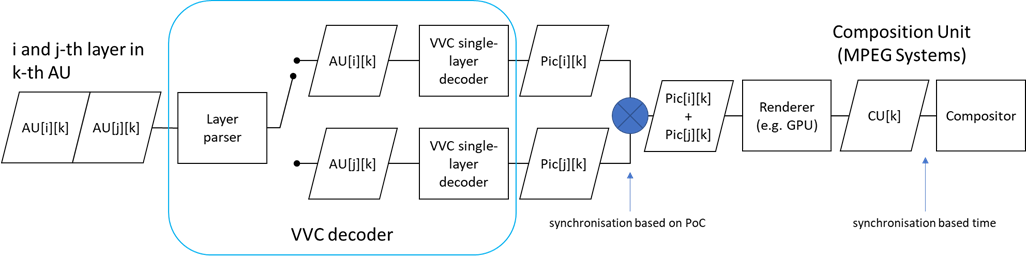
The proposal is based on the section 2 of JVET-S2017 “Technologies under consideration for VSEI” and can be summarized as follows:

1. Each layer contains the information determining its position in the assembled output picture.
2. Each layer holds the information of its neighbouring layer.
3. The position of each layer in the assembled picture is thus calculated based on the signalled information and the target OLS.
4. Decoded pictures with the same PoC are assembled into the assembled picture.
5. The existing cropping operation is applied to each decoded picture before assembled into the assembled picture.

Figures showing alternative approaches are below:



PoC-based and time-based synchronization in video pipeline



PoC-based layer combination

This does not propose using overlaps or gaps.

It was suggested to consider this in a joint discussion with MPEG Systems and Requirements.

See also JVET-T0049 (and JVET-S2017).

See section 8.1 for joint meeting discussion.

[JVET-T0053](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10432) AHG9: VVC and VSEI Annotated Regions SEI message [J. Boyce, P. Guruva reddiar (Intel)]

This contribution was discussed 2240 Wednesday 7 October 2020 (chaired by JRO & GJS).

It is proposed to add an annotated regions SEI message to VVC and VSEI, with *identical SEI message payload syntax* to that used in HEVC. Draft text and VTM software are provided.

VVC does not yet have an annotated regions SEI message.

The proposal contains only adjustment for terminology and the interface between VVC and VSEI.

Persistence was discussed, but the proposal is to keep this the same as previously specified in HEVC (and in an AVC output document).

It was noted that persistence scope is discussed in an errata report. The proponent indicated that a modified persistence would be complicated to specify, and did not propose any modification of that aspect.

It was asked whether there should be some interaction between the annotation and the subpictures concept. The proponent said that since objects do not align with subpictures, the concepts should probably not be coupled.

Decision: Remove the persistence aspect from the errata collection.

Decision: Adopt as WD, targeting v2 (when that happens).

See also T0059, which is a proposed generalization of the AR SEI message in three ways.

See also T0050, T0051, T0052.

[JVET-T0059](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10438) AHG9: Object representation SEI message [J. Chen, Y. Ye, J. Hu, K. Li, L. Hu, Y. Long (Alibaba)]

This contribution was discussed 2300 Wednesday 7 October 2020 (chaired by JRO & GJS).

This contribution proposes to provide an object representation SEI message in VSEI to represent objects detected or tracked in the pictures. The objects are represented by bounding boxes or bounding polygons with the depth in the proposed SEI message. Primary and secondary attributes of the object may also be signaled in the proposed SEI message.

In the second version, the design of secondary label present flag was simplified and some bugs in the syntax table and semantics were fixed.

HEVC supports the annotated regions SEI message, which is currently not in the VSEI spec. Though the proposed object representation SEI message is similar to the annotated regions SEI message in some aspects, its syntax supports the following capabilities that are not available in the annotated regions SEI message but are asserted to be necessary to some applications:

* Ability to represent a non-rectangular object using bounding polygons: some sophisticated object detection mechanism on the cloud can trace the object contour accurately, and such information could provide useful features (e.g. virtual background);
* Ability to attach up to two labels to a represented object: the proposed SEI can attach a primary label and a secondary label to a represented object, for example, the primary label can be “car” vs. “pedestrian” and the secondary label can be “red,” “blue”, or “yellow” for a car and “walking” or “standing” for a person;
* The depth of an object: the distance to the “camera” allows the rendering device to determine occlusion more accurately, and to select which objects to emphasize.

For the “secondary” label, a participant said that the primary/secondary content could be embedded within a label – e.g., “car, red”. For more than two degrees (primary, secondary, tertiary, etc.), the proponent said the signaling would get more complicated.

It was commented that the polygon case may need “sensibility constraints”. There are no semantics to ensure that the data that is sent is actually some polygon. There is no definition of what region is identified – which samples in the picture are within the “polygon” and which are not.

It was also commented that the number of vertices of the “polygon” is unconstrained as proposed.

The proponent of T0053 said there could be a benefit of not deviating from what is done in HEVC, by having tools and workflows that work the same way for the different contexts.

It was commented that if one scheme is a subset of the other, then some workflows could use the subset only.

It was commented that if we specify a generalization, we could also hypothetically specify a new SEI message for HEVC using the generalized form to support it in the HEVC context as well as in the VVC context.

It was commented that the use case for the existing form seems more clear than with polygons.

Among the three proposed generalizations, it was commented that the first and third seem the most potentially interesting.

Not much non-proponent support was expressed.

Further study of generalizations was encouraged, perhaps as a separate SEI message.

Stopped here in session 4.

[JVET-T0050](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10429) AHG9: HEVC Annotated Regions SEI message software update for HM [P. Guruva reddiar, J. Boyce (Intel)]

This contribution was discussed together with T0051 and T0052 at 0635 Friday 9 October 2020 (chaired by JRO & GJS).

An update is provided for the HM software encoder implementation of the Annotated Regions SEI message. This version of the software detects static regions in the input configuration file and avoids updating the per picture region parameters (position and size) signalled in the SEI message for static regions.

The contributions were appreciated. In the discussion, it was suggested to submit a merge request in GitLab for the newly provided software (both for this and the similar associated contributions), and to work with the software coordinator on exactly how to handle them.

See also T0051, T0052, T0053.

[JVET-T0051](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10430) AHG9: Object tracking information file format for use with Annotated Regions SEI messages [P. Guruva reddiar, J. Boyce (Intel)]

This contribution was discussed together with T005 and T0052 at 0635 Friday 9 October 2020 (chaired by JRO & GJS).

This contribution describes an uncompressed file format defined for describing per picture object tracking information. Example files for the JVET Class E test content is provided, based on annotations provided by Simon Frasier University.

Python scripts are provided to create HM or VTM Annotated Regions SEI config files from the object tracking information text files, and to generate the object tracking information files from the decoder’s output SEI message logs.

The Python scripts are the same for HM and VTM (assuming adoption into VTM), but a bit different for JM, as discussed in T0052.

See also T0050, T0052, T0053.

[JVET-T0052](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10431) AHG9: AVC Annotated Regions SEI message software for JM [P. Guruva reddiar, J. Boyce (Intel)]

This contribution was discussed together with T0050 and T0051 at 0635 Friday 9 October 2020 (chaired by JRO & GJS).

The Annotated Regions SEI message for AVC has been included in a draft specification output document, JCTVC-AN1006. In this contribution, software is provided for inclusion of the feature in the JM reference software. This contribution provides software for supporting the Annotated Regions SEI message in the JM reference software. It is very similar to the HEVC HM software implementation, provided in JVET-T0050, but has updated the input configuration file to align with the JM software.

See also T0050, T0051, T0053.

[JVET-T0056](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10435) AHG9: SEI manifest and SEI prefix indication SEI messages [Y.-K. Wang (Bytedance), T. Stockhammer (Qualcomm), A. M. Tourapis, D. Singer (Apple)]

This was discussed in session 13 at 0840 Tuesday 13 October (chaired by GJS & JRO).

[add summary]

There was discussion of whether the VSEI standard has the right scope to be able to specify this, or if it should be in the VVC standard instead. The VVC standard has awareness of the SEI messages and payload types that are in the VSEI standard, but not vice versa.

Decision: Adopt into an output draft for VVC; further study whether it would be possible to put this in VSEI instead.

[JVET-T0071](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10450) AHG9: SEI messages for support of cross RAP referencing based video coding [Y.-K. Wang (Bytedance)]

This was discussed in session 13 at 0900 Tuesday 13 October (chaired by GJS & JRO).

This contribution proposes the following two aspects for VVC/VSEI:

* The DRAP indication SEI message is extended by adding a random access point (RAP) picture ID, where a RAP picture is either an IRAP picture or a DRAP picture.
* A new SEI message, named type 2 DRAP indication SEI message, is added, to indicates type 2 DRAP pictures and provide RAP picture IDs of the type 2 DRAP picture itself and the list of RAP pictures it depends on, where a type 2 DRAP picture is a picture that may use the associated IRAP picture and a few other (type 1 or type 2) DRAP pictures for inter prediction reference.

DRAP (as currently specified) indicates that the current picture and all following pictures, in both decoding order and output order, can be correctly decoded without referring to any earlier pictures in decoding order except the associated IRAP picture.

It is asserted that the proposed changes enable full support of the so-called cross RAP referencing based video coding approach.

If we don’t want to extend the syntax of the existing DRAP indication SEI message, the type 2 DRAP indication could be used instead of it, although this would not be interpreted by older decoders. Both could be used at the same time.

It was commented that VVC / VSEI does not have the recovery point SEI message. GDR pictures were said to have a similar functionality to what was provided by that SEI message.

Stopped here at the end of session 13.

This was further discussed in session 14 at 1900 Tuesday 13 October (chaired by GJS & JRO).

The signalling includes a sort of duplicate DPB description. The purpose of this is not really for the video decoder, but more for a MANE.

One participant said this seems like a valuable proposal to address real-world use cases, e.g., can handle referencing across scene changes.

The proponent said this could help reduce the need for IRAP pictures and that the current DRAP picture design could result in a very long-distance reference to an IRAP picture. This would allow an intermediate update of the IRAP content in a DRAP prior to referencing by another DRAP picture.

It was commented that one could imagine a temporal sublayering structure of such DRAP usage. The concept was said to somewhat conceptually overlap with temporal sublayering.

It was asked what this can do that temporal sublayering cannot. This could indicate the use of a specific subset of “lower layer” pictures that are needed, rather than implying that all are needed (possibly with cascading dependencies). A participant supported this proposal, mentioning the cascading depencies.

It was noted that VVC does not have an indication of sub-layer non-reference pictures (alternatively, the encoder could put such a picture in the next higher layer, shifting the structure up by one temporal sublayer).

It was remarked that the “type 2” term might not be the best – e.g., the term “extended” was suggested instead.

The useful case was said to be when the encoder generates the SEI message while encoding - not by producing the SEI message by scanning the bitstream. An encoder the directly encodes into a file format would also likely not need this, as the functionality could be expressed in the file format track information.

Further study was encouraged to determine for certain whether this is a practical need.

[JVET-T0066](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10445) AHG9: IRAP only HRD SEI message [R. Skupin, Y. Sanchez, K. Suehring, T. Schierl (Fraunhofer HHI), Y.-K. Wang, L. Zhang (Bytedance), S. Deshpande (Sharp)]

This was discussed in session 14 at 1940 Tuesday 13 October (chaired by GJS & JRO).

This document proposes to add a new IRAP only HRD (IOH) SEI message to VVC that carries HRD information for an IRAP-only sub-bitstream that can be used in decoder trick play operation. The proposed IOH SEI message indicates irap\_only\_level\_idc, a maximum speedup as well as bitrate and CPB size of the IRAP-only sub-bitstream. Related changes to the scalable nesting SEI message in Annex D and extraction process and HRD timing derivation in Annex C are proposed.

There had been a related previous contribution JVET-R0068.

DASH has a mode specifically for IRAP-only play.

It proposes an indicated maximum allowed amount of speed-up of HRD operation.

It was asked whether encoders would really be likely to be designed to voluntarily provide such information (and to set the information correctly obey the expressed constraints) and whether decoders (or receiving systems) would really use it. The motivation described by the proponent is to be able to provide an expression of conforming behaviour in the syntax.

It was commented that encoders often don’t really be designed with great concern for strict HRD conformance behaviour.

Exact time speed matching was also said to not necessarily be so important for trick play.

Further study was encouraged to determine for certain whether this is a practical need.

[JVET-T0070](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10449) AHG9/AHG8: Scalability dimension SEI message [Y.-K. Wang, L. Zhang, K. Zhang, Z. Deng (Bytedance)]

This was discussed in session 21 at 2310 Thursday 15 October (chaired by GJS & JRO).

|  |  |
| --- | --- |
| scalability\_dimension( payloadSize ) { | **Descriptor** |
| **sd\_max\_layers\_minus1** | u(6) |
| **sd\_multiview\_info\_flag** | u(1) |
| **sd\_auxilary\_info\_flag** | u(1) |
| if( sd\_multiview\_info\_flag  | |  sd\_auxilary\_info\_flag ) { |  |
| if( sd\_multiview\_info\_flag ) |  |
| **sd\_view\_id\_len** | u(4) |
| for( i = 0; i  <=  vps\_max\_layers\_minus1; i++ ) { |  |
| if( sd\_multiview\_info\_flag ) |  |
| **sd\_view\_id\_val**[ i ] | u(v) |
| if( sd\_auxilary\_info\_flag ) |  |
| **sd\_aux\_id**[ i ] | u(8) |
| } |  |
| } |  |
| } |  |

This is proposed to be at the CVS level, having information for all the layers in a CVS.

Additional information would be needed for complete interpretation such as camera parameters, depth channel interpretation, alpha channel interpretation specification.

It was suggested to consider the potential use of a very large number of layers, saying that we should make sure to produce something relatively “future proof”.

It was commented that “vps\_max\_layers\_minus1” should be “sd\_max\_layers\_minus1”, although perhaps those should be required to be equal to each other.

It was commented that the SEI message approach seems like a good one for handling this sort of information.

Further study was encouraged toward development of a more complete approach.

[JVET-T0076](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10455) AHG9: ALF control SEI message [H.-B. Teo, H.-W. Sun, C.-S. Lim (Panasonic)]

[Add summary]

See section 8.1 for joint meeting discussion. TBP some use case info.

[JVET-T0090](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10469) AHG9: Digital signature SEI message [J. Xu, Y.-K. Wang, L. Zhang, K. Zhang (Bytedance)]

[Add summary]

See section 8.1 for joint meeting discussion. Further detailed presentation was not deemed necessary at this meeting.

[JVET-T0092](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10471) AHG9/AHG11: Neural network based super resolution SEI [T. Chujoh, E. Sasaki, T. Ikai (Sharp)]

[Add summary]

See also notes for this contribution in section 5.2 and also for T0096.

See section 8.1 for joint meeting discussion. Further detailed presentation was not deemed necessary at this meeting.

## Other (1)

[JVET-T0077](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10456) CICP use cases for signalling of still image graphics (chyrons) in video [Y. Syed, W. Husak, C. Seeger]

This was discussed in session 21 at 2245 Thursday 15 October (chaired by GJS & JRO).

This informational document provides a number of example use cases of the need for signalling still image graphics or chyrons that are used in different types of video experiences. Traditionally the use of chyrons are used in production of linear video feeds to composite still graphics images with the video. With the advent of IPTV technologies, increasingly the use of still image graphics is happening more downstream and closer to the consumer to create more localized or personalized experiences. With the introduction of HDR video, still image graphics needs to be closely matched with the video feed or otherwise presentation visual artefacts will manifest (e.g. flaring or dimming). It was reported that with proper signalling of the still image graphic (PQ, HLG, or SDR), these types of mismatches can be avoided. This document indicates near-term and future use case examples where these issues may become more relevant and encourages more exploration of this topic.

It was commented that these issues may be systems matters. The contributor said that ICC profiles and various file formats such as TIFF (which may be specified in SMPTE and ISO standards) could be involved. JPEG Part 4 was reported to have relevant colour profile handling aspects. The contributor suggested that this should potentially be discussed in the technical report on usage of video signal type identification (UVSTP, ITU-T H.Sup19 and ISO/IEC 23091-4). Further study was recommended, particularly including consideration of systems environment interaction, and whether is a systems matter rather than a video coding topic.

# Withdrawn (X)

Section kept for future use.

JVET-T0083 Withdrawn

JVET-T0115 Withdrawn

JVET-T0126 Withdrawn

# Plenary meetings, joint meetings, BoG reports, and summary of actions taken

## Joint meeting of JVET, VCEG, MPEG Systems, MPEG Requirements Monday 12 October 2000-2100

* New SEI proposals with potential systems impact
  + Multilayer compositing (JVET-T0049, JVET-T0080)
    - System functionality or appropriate for “in-band” signalling?
    - Proponent indicated ease of synch by in-band approach
    - Others said this seemed more like a system functionality, e.g.
      * one could have multiple ways of composing scenes from the same content or
      * could composite decoded content with the content of some other bitstream(s)
      * Another participant said that the compositing functionality is ordinarily something outside of a decoder.
      * The potential for “hints” to convey to an outside process was mentioned as a possibility, but not in-band composition.
    - Two variations were in previous TuC JVET-S2017; a third one is proposed in JVET-T0049
    - Further study may still be warranted, but no action, no output at this meeting.
  + Digital signature (JVET-T0090)
    - Similar concept was proposed at the previous meeting in JVET-S0051, with extensive discussion (see the previous meeting notes) and a conclusion of “Further study was encouraged, and should involve the question of why this should be done in the sample domain in important use cases, and should involve security experts.”
    - Question: Verification would involve decoding first, followed by verification checking.
    - What is new in the proposal since the last meeting: Adding fingerprint scheme and type indicators for digital signature type and hash type.
    - It was commented that typical integrity checking methods operate in the compressed domain (including only relevant NAL units) - not on decoded pictures. Content access (CA) control and content protection (CP) are also typically involved in security related applications.
    - The situation had not significantly changed, and further study along the lines previously identified was requested. It was further commented that more realistic application scenarios where this would be used should be provided.
* Post-processing control
  + Neural-net super-resolution (JVET-T0092)
    - Conceptually similar to post-process filtering
    - Potentially large amount of data for neural network
    - Likely high complexity of the post-processing operation itself
    - Designed for a specific resampling ratio? Would that match the need for the decoder-side display?
    - For further study - the NN processing technology doesn’t seem yet mature for action
  + ALF post-filter control (JVET-T0076)
    - Why have this if the decoder has in-loop filtering capability? Would out-of-loop be better than in-loop?
    - A proposed use case is to add post-processing for pre-encoded content without transcoding
    - Sharpening filtering is another example use described in the proposal.
    - Extra buffering capacity?
    - It was commented that this would require specific support in decoders, and thus does not “come for free”
    - For further study - benefit and implementation aspects would need further justification.
* “Traditional” SEI message proposals
  + Scalability dimension (JVET-T0070)
  + IRAP-only HRD (with level) (JVET-T0066)
  + Cross-DRAP referencing (JVET-T0071)
* Previously known topics
  + Annotated regions (JVET-T0050, JVET-T0051, JVET-T0052, JVET-T0053, JVET-T0059)
  + SEI manifest & SEI prefix (as in HEVC, AVC, V-PCC) (JVET-T0056)
    - In previous TuC JVET-S2017

## Joint meeting of JVET, VCEG (Q6/16), MPEG Requirements, and MPEG Video Tuesday 13 October 0500-0600

On future video coding work

* High bit depth / high bit rate (discussed at previous meeting) for VVC
  + Clear potential for some near-term amendment (e.g. CD April, DIS July, FDIS January 2022)
* Neural network-based video coding exploration
  + It was agreed by VCEG and WG2 of MPEG to move “DNNVC” into JVET
  + Tool level
  + End-to-end / broader investigation
  + Requirements aspects to be studied in VCEG and WG2 of MPEG while technical work is in JVET
  + No immediate action to develop an amendment or new standard
  + JPEG is planning a nearer-term “JPEG AI” effort
    - This may have some “machine vision” / “machine consumption” aspects
* Other potential work (e.g. other coding efficiency approaches)
  + MPEG Video (WG4) is continuing to work on more immersive video and dense light fields
  + Video coding for machines (under study in WG2) - still relatively early exploration - potential for a Call for Evidence to be issued soon (possibly with additional phases of work)
  + Neural network representation (NNR) as under development in WG4 may be a technology that could be used in other work that needs NN technology
  + It was agreed that JVET should study general improved coding efficiency technology (our efforts should not be limited to neural network technology)
* JVET-T0065 Conformance points for multilayer 8K [F. Bossen (Sharp)]
  + 8K enhancement layer with 4K 60 fps base layer - Level 6.2 cannot support RA config with GOP size 32; 14 buffers needed and only 8 available. (We have been using GOP size 16, which would be 12 buffers needed with only 8 available.)
    - The contributor said the penalty of reducing GOP size below 16 would probably be 5-10%. (The penalty for 16 rather than 32 is about 3%.)
  + Proposal suggests to consider doubling the DPB size for the multilayer profiles
    - On the other hand, it had been intended that scalability could “come for free” in terms of basic complexity requirements
  + Alternatively, the proposal suggests to consider defining a new level with higher DPB capacity
  + Potential performance of such a scheme was discussed.
  + The proposed operation point was said to already be practical for current implementation - mostly a matter of memory rather than anything else. The throughput requirement is already sufficient at level 6.2; only the memory capacity is a problem.
  + Some support for action was expressed, while others suggested additional study.
  + It was agreed to establish such an additional level - details to be worked out.

## Joint meeting of JVET, VCEG (Q6/16), MPEG Systems, and MPEG Video Tuesday 13 October 2030-2100

This joint meeting was held to discuss an incoming liaison statement from DASH-IF to MPEG systems.

T. Stockhammer presented the liaison statement.

With adaptive streaming becoming the dominant usage of video content, DASH-IF considers codec initialization, random access, switching, splicing, ad insertion, etc., and desires conformance bitstreams to test these functionalities.

The discussed aspects consider the boundary between a system and a video elementary bitstream. In video coding standards, we tend to consider our domain ending at an “end of bitstream NAL unit”.

Some of this can get into difficult “gray areas” where it is difficult to ensure seamless behaviour.

It was commented that perhaps some other organization might want to specify that a concatenation of bitstreams must play seamlessly; that it may be difficult or inappropriate for a video conformance specification to include concatenated transitions between substantially different configurations.

It was commented that some aspects are application specific.

However, it was also suggested that where there is some clear requirement in a video coding standard (e.g. to handle a buffer flow or transition), it is desirable for the conformance test set to test that.

In terms of schedule, VVC conformance is planned to go to CD ballot as an output of this meeting. Specific proposals for inclusion in the conformance test set (especially if accompanied by volunteers to supply the bitstreams and descriptions) are generally welcome. For EVC, a DIS ballot is planned in MPEG as an output of this meeting.

Bitstream that test profile/level/picture-size/bit-depth/colour-format transitions and non-IDR startup of a bitstream are desirable.

There are also interop tests conducted outside of our organization (e.g. in MC-IF), which can also be helpful.

Adding more tests to conformance specs later after v1 is also welcome.

MPEG Systems was planning to reply to the ILS and to provide information on the current status and plans.

There was also some ongoing discussion of whether conformance tests should be considered normative standards or not. A possibility was described by MPEG Systems for there to be a standardized “snapshot” periodically and also an informal repository.

## Joint meeting of JVET, VCEG (Q6/16), MPEG Visual Quality Assessment Wednesday 14 October 1900-1930

The status of verification test planning was reviewed and discussed, as follows:

* SDR UHD testing completed – produce output report.
* For SDR HD, plan SDR RA as well as SDR LD, as suggested at Sep. AHG meeting
  + Sequences proposed by Alibaba as conversational content appropriate for LD
  + Another proposed by Ericsson.
  + Having some gaming content would be desirable. True gaming content can sometimes have very fast action. Witcher sequence used in dry run was originally 120 Hz and was played at 60 but was still perceived as too fast.
* For HDR, see JVET-T0106 for suggested viewing configuration. Have played from MP4 files, but formal test should avoid the transcoding. HHI has a way to play raw material to HDR display. This has worked well for PQ but there has been some problem with the VLC player for HLG content.
* For 360°, checking QP settings to cover quality range. The ERP scheme seems ready for performing the testing. Blending for CMP has helped (see the notes for JVET-T0118), although some artefacts remain at mid to low rates.
  + P-CMP/GCMP padding was being done by just repeating the row/column content at the encoder side. [Reported at 2125 in Session 20: Correct notes as necessary - this is not correct.]
  + In the discussion, it was strongly suggested to study instead use the content from corresponding spatial positions in spatially neighbouring faces (at the positions that will be used for the blending).

MPEG VQA (AG5) was planning to send an LS to ITU-T SG12.

VCEG was planning to send a reply to ITU-R WP 6C.

Agreed plan: MPEG VQA will send an LS to ITU-T SG12 and ITU-R WP 6C; VCEG will send a brief reply and note it. The LS will say there will be an investigation of the suggested video test material for its potential use in our work.

## Closing plenary meeting Wednesday XX June xxxx-xxxx

## BoGs (2 - 1 revisit)

The following break-out groups were established at this meeting and produced the below-listed reports.

[JVET-T0124](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10521) BoG report on common test condition for high bit depth, high bit rate, and high frame rate coding [T. Ikai]

This BoG report was reviewed in session 18 at 0500 on 15 October 2020 (chaired by JRO & GJS).

This is about BoG on common condition for high bit-rate, high bit-depth and high frame-rate coding.

A first session took place on 13 Oct, 7:20 to 9:30 UTC with around 30 participants. Sequences, bit depth, colour format/sampling, QPs, tools and encoder configurations were recommended while evaluation criteria, spread sheet, and high frame rate were not discussed.

A second session took place on 14 Oct, 21:40 to 23:20 UTC with around 30 participants. Evaluation criteria, spread sheet, high framerate has been confirmed and an initial CE draft was quickly reviewed.

BoG recommendations for high bit depth and high bit rate common test condition (relative to JVET-T0047, Preliminary common test conditions for high bitdepth video coding test conditions from AHG):

* Keep five SVT sequences, remove FireEater2Clip from PQ, adopt four new HLG sequences, Daystreet, PeopleInShoppingCenter, NightStreet StainedGlass
* Approve RGB 4:4:4, YUV 4:4:4 and YUV 4:2:2 testing
* Approve 12 bit and 16 bit testing
* Approve QP ranges (QP= -13, -8, -3, 2, 7, 12) for 12 bit and QP ranges (-33, -28, -23, -18, -13, -8) for 16 bit
* Approve lossless test as optional
* Enable LMCS but disable LFNST (doesn’t really help at high bit rates, but adds to encoding time)
* Restrict max MTT size (a.k.a maxmtt16) as proposed in JVET-T0087, for high bit depth low QP encoder configuration
* Set extended precision flag equal to 0 for 12 bit testing and 1 for 16 bit testing
* Mandatory 12 bit RA with 97 frames. Use 50 frames for HD and 25 frames in LDB
* Choose PSNR for SDR and HLG content and both PSNR / wPSNR for PQ content as quality measures
* Test JVET-T0072, JVET-T0085, JVET-T0089 and JVET-T0105 in a CE and review CE description for rice parameter derivation
* For high frame rates, it was suggested to not establish tests for this in the upcoming meeting cycle, but encourage the identification of higher frame rate test content (beyond the current 100/120 fps but several hundred fps). It was commented that EBU may have some appropriate content.

The BoG recommendations were approved by JVET (with the CE description remaining to be reviewed).

CE coordinators were identified as A. Browne, T. Hashimoto, H.-J. Jhu, D. Rusanovskyy.

[JVET-T0130](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10527) BoG Report: Neural Network Technology [A. Segall]

This BoG report was discussed in session 18 at 0615 Thursday 15 October (chaired by GJS & JRO).

The BoG held the following meetings during the 20th JVET meeting prior to the time of JVET status review:

* October 13 - 20:20–21:20
* October 14 – 21:40–23:20

Tables of test conditions and complexity metrics with agreed “way forward” aspects were shown (QP values, bit rate targets, objective metrics, test sequences, and anchor).

The BoG recommended to reduce Class A1/A2 to 3s for LDB/LDP.

The BoG recommended to establish a mandate for an AhG to create results with temporal filtering enabled.

It was proposed to define the exploration experiment to focus on loop filter technologies. The contributions include JVET-T0057, JVET-T0069, JVET-T0079, JVET-T0088, JVET-T0094, and JVET-T0128.

It was commented that the training data is a crucial matter.

It was noted that there is no way to verify that the training set has not included the test set, unless the test set is not known in advance.

It was commented that we should not have a preference for a particular tool-level technology approach, e.g., CNN-based loop filtering rather than a more broad investigation. On the other hand, the loop filtering is what has had the most contributions and is the most straightforward test that can be performed and compared. Much of the desire is to test the test procedure, not just the technology.

There was discussion of the selection of QP values, with a participant saying that the differences in conditions with using 4 QP values and 5 QP values might be confusing and potentially cause errors in copy-pasting work. In the discussion it was clarified that there will be a single Excel spreadsheet to fill out.

It was commented that showing the evolution of “learning curve” performance over training epochs is very interesting as information toward understanding the technology and looking for potential overtraining.

It was agreed that the report of the DNNVC activity and prior review in the MPEG DNNVC AHG should be considered sufficient, such that detailed additional presentations of the MPEG contributions was not considered necessary. Further study on these was, of course, recommended.

Revisit after further BoG discussion.

## List of actions taken affecting the draft text of VVC, the VTM, and 360Lib

[Remove this section]

The following is a summary, in the form of a brief list, of the actions taken at the meeting that affect the text of the VVC draft text, VTM or 360Lib description. Both technical and editorial issues are included. This list is provided only as a summary – details of specific actions are noted elsewhere in this report and the list provided here may not be complete and correct. The listing of a document number only indicates that the document is related, not that it was adopted in whole or in part. The description given in the “Tool” column is a best effort for the sake of understanding but may not precisely reflect the functionality of the tool. It is also noted that in cases where several contributions proposed the same method, usually only one of the is listed as adoption below; refer to the meeting notes about the adoption to see which other contributions are related.

[Add actions of Friday 24 April.]

[This is just a reflection of what has already been recorded.]

# Project planning

## Core experiment planning

No CEs were planned at this meeting.

## 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 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 XXday XX Apr 2020.

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 during the opening plenary on January 9 that those rules which had been set up or refined during the 12th meeting should be observed. In particular, for some CEs, results were available late, and some changes in the experimental setup (particularly in CE4) were not 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 the draft standard by the next meeting.
* 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; the method for members to obtain the credentials is TBA on the reflector).
* 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 are described in the output document JVET-N1010.

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.

[Add info on software access.]

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 marked as “withdrawn”.

T2 = Test model software release + 2 weeks or X XX, whichever is earlier: 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 – 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 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.

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.

## Software development and anchor generation (update)

Proponents of adopted aspects were asked to inform the AHG3 chairs to clarify whether software action is needed.

The planned timeline for software releases was established as follows:

* VTM8.0 will be released by 2020-02-17 including all adoptions necessary for CTC. VTM8.1 with non-CTC adoptions will be released 2020-03-16. Further versions of VTM may be released for additional bug fixing, as appropriate.
* Preparation of the VTM software will include immediate removal of macros that were added in the previous meeting cycle. The software coordinator has the discretion to retain some such macros.
* 360lib software is to be revised for the modified generalized cubemap, which was requested by 2019-02-28
* No change of HDRTools software was noted in response to meeting.

# 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)).

Initial review in session 21 at 2145 Thursday.

+Meeting plans for ad hocs

|  |  |  |
| --- | --- | --- |
| **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]. * Collect reports of errata for the VVC, VSEI, HEVC, AVC, CICP, the codepoint usage TR specification and the published HDR-related technical reports. * Produce and finalize JVET-T2002 VVC Test Model 11 (VTM 11) Algorithm and Encoder Description. * Propose improvements to the JCTVC-AN1002 HEVC Test Model 16 (HM 16) Update 14 of Encoder Description * 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, V. Drugeon, S. Kim, S. Liu, J.‑R. Ohm, K. Sharman, 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 model (VTM, HM, SCM, SHM, HTM, MFC, MFCD, JM, JSVM, JMVM, 3DV-ATM, 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. * Coordinate with AHG6 for integration with 360lib software. | F. Bossen, X. Li, K. Sühring (co-chairs), 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 draft verification test plan JVET-T2009 and develop proposed improvements for verification testing of VVC capability. * Maintain the video sequence test material database for testing the VVC and HEVC standards and potential future extensions. * 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. * 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, particularly including Rep. ITU-R BT.2245. * Prepare availability of viewing equipment and facilities arrangements for future meetings. | V. Baroncini, T. Suzuki, M. Wien (co-chairs), E. François, A. Norkin, A. Segall, P. Topiwala, S. Wenger, Y. Ye (vice-chairs) | Tel.  2 weeks notice |
| **Conformance testing (AHG5)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Produce the JVET-T2008 draft conformance testing specification and develop proposed 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, K. Sühring, X. Xu (vice-chairs) | N |
| **360° video coding, software and test conditions (AHG6)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the effect on compression and subjective quality of different projections formats, resolutions, and packing layouts. * Solicit additional test sequences, and evaluate suitability of test sequences on head-mounted displays and normal 2D displays. * Study the effect of viewport resolution, field of view, and viewport speed/direction on visual comfort. * Prepare and deliver the 360Lib-11 software version and common test condition configuration files according to JVET-L1012. * Generate CTC anchors and PERP results for the VTM according to JVET-L1012 within two weeks of availability of SDR CTC anchors. * Coordinate with AHG4 in preparation for verification testing for 360° video content. * Produce documentation of 360° software usage for distribution with the software. | J. Boyce and Y. He (co-chairs), K. Choi, J.-L. Lin, Y. Ye (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. * Generate CTC anchors for the VTM according to JVET-T2011. * Study the luma/chroma bit allocation in the HDR CTC, especially for HLG content. * Coordinate implementation of HDR anchor aspects in the test model software with AHG3. * Coordinate with AHG4 in preparation for verification testing for HDR video content. * Study additional aspects of coding HDR/WCG content. | A. Segall (chair), E. François, W. Husak, S. Iwamura, D. Rusanovskyy (vice-chairs) | N |
| **Layered coding and resolution adaptivity? (AHG8)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study approaches for support of layered scalable coding and adaptive-resolution coding, including spatial, temporal, quality, view, subpicture, and region-of-interest aspects; and analyse their coding efficiency and complexity characteristics. * Consider 360° viewport-dependent streaming and real-time communication applications of layered coding and resolution adaptivity. * Coordinate with AHG2 and AHG3 for text drafting and software development for the layered coding and resolution adaptivity aspects of the VVC design. * Study and develop improvements of the JVET-Q2015 functionality testing condition description. * Propose common test conditions for layered coding and resolution adaptivity. | S. Wenger and A. Segall (co-chairs), M. M. Hannuksela, Hendry, S. McCarthy, Y.-C. Sun, P. Topiwala, Y.-K. Wang (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 SEI showcase and usage information for SEI messages, including encoder and decoder implementations and bitstreams for demonstration and testing. * Identify potential needs for additional SEI messages. * Study SEI messages defined in HEVC and AVC for potential use in the VVC context. * Coordinate with AHG3 for software support of SEI messages. | S. McCarthy, J. Boyce (co-chair), 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 GOP structures, GDR, LMCS and perceptually optimized adaptive quantization for encoder optimization. * Study encoding techniques of optimization for objective quality metrics and their relationship to subjective quality. * Study the impact of adaptive quantization. * 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. | A. Duenas, A. Tourapis (co-chairs), A. Norkin, R. Sjöberg, J. Le Tanou (vice-chairs) | N |
| **Neural network-based video coding (AHG11)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study potential extensions of VVC with NN-based coding tools for video coding, such as intra or inter prediction modes, partitioning, transforms, and in-loop or post filtering. * Study NN-based encoding optimization for VVC. * Study the impact of training on the performance of candidate technology. * Analyse complexity characteristics and perform complexity analysis of candidate technology. * Identify video test materials, training set materials, and testing methods for assessment of the effectiveness and complexity of considered tools. * Develop reporting templates for test results and analysis of candidate technology. * Coordinate with relevant activities of the parent bodies. | E. Alshina, S. Liu, J. Pfaff, M. Wien, P. Wu, Y. Ye (co-chairs) | Tel.  2 weeks notice |
| **High bit depth, high bit rate, and high frame rate coding (AHG12)**  ([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. * Identify potential needs for future extension of VVC to support such application usage. * Define testing conditions and test sequences for high bit depth, high bit rate, and high frame rate coding in coordination with AHG 4. * Solicit/study HFR test sequences … | A. Browne, T. Ikai, X. Xiu (co-chairs) | Tel.  2 weeks notice |
| **Tool reporting procedure and testing (AHG13)**  **Discontinue – establish AHG on study new coding tools?**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Prepare output document JVET-S2005, which describes the methodology of tool-off testing and a list of tools to be tested by identified testers, including non-CTC configurations as appropriate. * Produce, study and develop improvements of the JVET-R2013 testing condition description for non-4:2:0 colour format coding. * Provide configurations files, bitstreams, and results of tool-on/tool-off testing. * Develop and collect test results for additional testing of VVC capabilities. * Maintain VTM software aspects for memory bandwidth analysis in coordination with AHG3. * Use the tool usage counts and memory bandwidth usage to study the decoder complexity of features in on/off testing. * Prepare a report with results of the tests. | W.-J. Chien, J. Boyce (co-chairs), Y.-W. Chen, R. Chernyak, K. Choi, R. Hashimoto, Y.**-**W. Huang, H. Jang, R.-L. Liao, S. Liu (vice-chairs) | N |

*Former JCT-VC AHGs (change reflector, and merge some)*

|  |  |  |
| --- | --- | --- |
| **Title and Email Reflector** | **Chairs** | **Mtg** |
| **JCT-VC project management (AHG1) merge**  ([jct-vc@lists.rwth-aachen.de](mailto:jct-vc@lists.rwth-aachen.de))   * Coordinate overall JCT-VC interim efforts. * Report on project status to JCT-VC reflector. * Provide a report to next meeting on project coordination status. | G. J. Sullivan, J.-R. Ohm (co‑chairs) | N |
| **Test model editing and errata reporting (AHG2)**  ([jct-vc@lists.rwth-aachen.de](mailto:jct-vc@lists.rwth-aachen.de))   * Propose improvements to the JCTVC-AL1002 HEVC Test Model 16 (HM 16) Update 13 of Encoder Description * Collect reports of errata for the HEVC, AVC, CICP, the codepoint usage TR specification and the published HDR-related technical reports. * Gather and address comments for refinement of these documents. * Coordinate with AHG3 on software development and software technical evaluation to address issues relating to mismatches between software and text. | B. Bross, C. Rosewarne (co‑chairs), J.‑R. Ohm, K. Sharman, G. J. Sullivan, A. Tourapis, Y.‑K. Wang (vice‑chairs) | N |
| **Software development and software technical evaluation (AHG3)**  ([jct-vc@lists.rwth-aachen.de](mailto:jct-vc@lists.rwth-aachen.de))   * Coordinate development of the HM, SCM, SHM, HTM, MFC, MFCD, JM, JSVM, JMVM, 3DV-ATM, and HDRTools software and their distribution. * Enable software support for recently standardized additional SEI messages. * Produce documentation of software usage for distribution with the software. * Prepare and deliver results, reporting templates, and anchor test results according to JCT-VC common 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 AHG2 on HEVC and AVC test model editing and errata reporting to identify any mismatches between software and text. | K. Sühring (chair), B. Li, K. Sharman, V. Seregin, G. Tech, A. Tourapis (vice‑chairs) | N |
| **Supplemental enhancement information (AHG4) merge**  ([jct-vc@lists.rwth-aachen.de](mailto:jct-vc@lists.rwth-aachen.de))   * Study the status of text and potential needs for SEI messages for AVC and HEVC. * Consider proposals for additional SEI message data and associated syntax and semantics specification. * Develop usage scenario descriptions and showcase demonstrations. * Coordinate with AHG3 for software support of SEI messages. | J. Boyce (chair), C. Fogg, S. McCarthy, G. J. Sullivan, Y.-K. Wang (vice‑chairs) | N |
| **Test sequence material (AHG5) merge**  ([jct-vc@lists.rwth-aachen.de](mailto:jct-vc@lists.rwth-aachen.de))   * Maintain the video sequence test material database for development of HEVC and its RExt, SHVC and SCC extensions. * Identify, collect, and make available a variety of video sequence test material, especially focusing on new needs for HDR/WCG test material and corresponding SDR test material. * Collect information about test sequences that have been made available by other organizations. * Study coding performance and characteristics in relation to video test materials. * Identify and recommend appropriate test materials and corresponding test conditions for use in development of HEVC and its extensions. * Coordinate with the activities in AHG3 and AHG6 regarding HDR/WCG testing. | T. Suzuki, V. Baroncini (co‑chairs), E. François, P. Topiwala, S. Wenger (vice‑chairs) | N |

# 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 MPEG output document, a separate version under the MPEG document header should be generated. This version should be sent to GJS and JRO for upload.

[JVET-T1000](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9674) Meeting Report of the 20th JVET Meeting [G. J. Sullivan, J.-R. Ohm] (2020-11-13)

Initial versions of the meeting notes (d0 … dA) 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: [JCTVC-AN1002](http://phenix.int-evry.fr/jct/doc_end_user/current_document.php?id=11000) High Efficiency Video Coding (HEVC) Test Model 16 (HM 16) Encoder Description Update 14 [C. Rosewarne (primary editor), K. Sharman, R. Sjöberg, G. J. Sullivan (co-editors)] (WG 11 N 19473) [2020-10-05] (near next meeting)

JVET-T1003 Draft revised coding-independent code points for video signal type identification [G. Sullivan, T. Suzuki, A. Tourapis] [2020-10-30]

DIS N 12 in ISO/IEC with DoCR N 11.

Ballot comments 54912

* Code value 22 in Table 3 - remove the informative reference
* Remove “(historical)” from ST 240
* Reference the 2019 edition of SMPTE 428-1 (assuming that its content is the same in regard to the referenced aspects) and the 2018 edition of ARIB STD-B67 to the latest editions (BT.470-7 and STD-B67. Update other references if identified.
* Include the year in SMPTE references.
* Remove colour from figures.
* Add history commentary

[JVET-T1004](http://phenix.int-evry.fr/jct/doc_end_user/current_document.php?id=11006) Errata report items for HEVC, AVC, Video CICP, and CP usage TR [C. Rosewarne, G. J. Sullivan, Y. Syed, Y.-K. Wang] [2020-12-15] (near next meeting)

(C. Rosewarne added.)

[JVET-T1005](http://phenix.int-evry.fr/jct/doc_end_user/current_document.php?id=10997) Shutter interval information SEI message for HEVC (Draft 3) [S. McCarthy, G. J. Sullivan, Y.-K. Wang] (FDAM N 8) [2020-10-30]

DoCR N 7.

JVET-T1006 Annotated regions and shutter interval information SEI messages for AVC (Draft 2) [J. Boyce, S. McCarthy, Y.-K. Wang] [2020-12-15]

Request N 15 and WD N 16, include errata. The specification will reference VSEI. [Check title]

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 (editors)] (WG 11 [N 15778](http://phenix.it-sudparis.eu/mpeg/doc_end_user/current_document.php?id=53941&id_meeting=165))

JVET-T1008 Usage of video signal type code points (Draft 2 for version 3) [W. Husak, G. J. Sullivan, Y. Syed, A. Tourapis (editors)] [2020-11-13]

TR in MPEG N 14 with DoCR N 13

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 (editors)]

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 (editors)]

No output: JVET-T1011

No output: JVET-T1012

No output: 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 (editors)] (WG 11 [N 16049](http://phenix.it-sudparis.eu/mpeg/doc_end_user/current_document.php?id=54889&id_meeting=166))

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 (editors)] [2017-02-17]

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 (editors)] [2017-02-17] (1 month)

Remains valid for HM – not updated: [JCTVC-AF1100](http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=10693) Common Test Conditions for HM Video Coding Experiments [K. Sharman, K. Sühring (editors)] [2018-09-14]

[JVET-T2001](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9675) Versatile Video Coding Editorial Refinements on Draft 10 [B. Bross, J. Chen, S. Liu, Y.-K. Wang] [2020-10-30]

[JVET-T2002](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9676) Algorithm description for Versatile Video Coding and Test Model 11 (VTM 11) [J. Chen, Y. Ye, S. Kim] [] (2020-12-15, near next meeting)

N 23.

To include GOP structures (GOP 32 & LD) and MCTF description.qq

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] (2019-04-01)

[JVET-T2004](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9677) Algorithm descriptions of projection format conversion and video quality metrics in 360Lib (Version 12) [Y. Ye, J. Boyce] (2020-12-15, near next meeting)

Including cross-boundary blending for GCMP. If confirmed feasible and appropriate in interim study, to pad the encoder side with adjacent content rather than boundary extension.

It was noted that this includes some “stale” formats that are no longer subjects of active investigation.

Remains valid for VTM – not updated: [JVET-S2005](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9678) Methodology and reporting template for coding tool testing [W.-J. Chien and J. Boyce] (2020-08-15)

[JVET-T2006](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=5758) Common Test Conditions and evaluation procedures for neural network-based video coding technology [S. Liu] (2020-10-xx)

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] [WG 11 N 19472] (2020-07-29)

[JVET-T2008](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9680) Conformance testing for versatile video coding (Draft 5) [J. Boyce, E. Alshina, F. Bossen, K. Kawamura, I. Moccagatta, W. Wan] [] (2020-10-30)

CD N 11.

[JVET-T2009](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9681) VVC verification test plan (Draft 4) [M. Wien, V. Baroncini, A. Segall, Y. Ye] [] (2020-12-15)

N 22.

The plan is to finalize the VT plan for HD, HDR and 360 in January and conduct the testing and report the results by April.

[JVET-T2010](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=6643) VTM common test conditions and software reference configurations for SDR video [F. Bossen, J. Boyce, X. Li, V. Seregin, K. Sühring] (2020-10-xx)

Updated to reflect the GOP structure change (disable MCTF for category F, any scene cut issues?).

[JVET-T2011](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=8862) VTM common test conditions and evaluation procedures for HDR/WCG video [A. Segall, E. François, W. Husak, S. Iwamura, D. Rusanovskyy] (2020-10-xx)

Updated to reflect the GOP structure change (MCTF enabled - may not need to mention in document).

Remains valid – not updated: [JVET-L1012](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=4840) JVET common test conditions and evaluation procedures for 360° video [P. Hanhart, J. Boyce, K. Choi, J.-L. Lin] (2018-10-26)

No change was needed for the GOP structure issue (MCTF usage?).

[JVET-T2013](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9682) 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] (2020-10-xx)

Updated to reflect the GOP structure change (does MCTF operate for non-4:2:0 cases?).

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] (2020-03-02)

Change needed for the GOP structure issue? (MCTF usage?)

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] (2020-03-02)

[JVET-T2016](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9673) Summary information on BD-rate experiment evaluation practices [K. Andersson, F. Bossen, J.-R. Ohm, A. Segall, R. Sjöberg, J. Ström, G. J. Sullivan, A. Tourapis] [] (2020-11-06)

N 6 with DoCR N 5.

JVET-T2017 Additional SEI messages for VSEI (Draft 1) [J. Boyce, Y.-K. Wang] (2020-11-13)

Annotated regions

Request N 17, WD N 18 (target April CDAM).

JVET-T2018 Common Test Conditions for High Bit Depth and High Bit Rate Coding [A. Browne, T. Ikai, D. Rusanovskyy, X. Xiu]

JVET-T2019 New level and additional SEI messages for VVC (Draft 1) [F. Bossen, Y.-K. Wang] (2020-11-13)

Request N 19 “Operation range extensions”, WD N 20 (target April CDAM).

New level, SEI manifest, SEI prefix, interface to annotated regions.

JVET-T2020 VVC verification test report for UHD SDR video content [V. Baroncini, M. Wien] (2020-11-13)

N 21.

[JVET-T2021](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9680) Reference software for versatile video coding (Draft 1) [F. Bossen, K. Suehring, X. Li] [] (2020-10-30)

CD N 12.

JVET-T2022 CE on Entropy Coding for High Bit Depth and High Bit Rate Coding [A. Browne, T. Hashimoto, H.-J. Jhu, D. Rusanovskyy] (2020-10-30)

N 24.

JVET-T2023 EE on Neural Network-based Video Coding [TBD] (2020-1x-xx)

N 25.

# 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/WG 11 auspices (or its successor) when it meets (ordinarily starting meetings on the Friday prior to such meetings and closing it at lunchtime on the last day of the WG 11 meeting – a total of 7.5 meeting days).

In cases where an exceptionally high workload is expected for a meeting, an earlier starting date may be defined.

Some specific future meeting plans (to be confirmed) were established as follows:

* Wed. 7 – Fri. 9 and Mon 12 – Friday 16 October 2020, 20th meeting under WG 11 auspices as a teleconference (7.5 meeting days).
* Fri. 8 – Fri. 15 January 2021, 21st meeting under WG 11 auspices in Capetown, ZA.
* Wed. 21 – Wed. 28 April 2021, 22nd meeting under ITU-T auspices in Geneva, CH.
* Fri. 9 – Fri. 16 July 2021, 23rd meeting under WG 11 auspices in Prague, CZ.

The agreed document deadline for the 20th JVET meeting was planned to be Wednesday 30 Sep. 2020.

Vittorio Baroncini, Andrew Segall, Mathias Wien, and Yan Ye were thanked for their efforts in further developing the VVC verification test plan and procedure. GBTech was thanked for kindly providing resources and conducting the dry run for SDR UHD test cases with non-expert viewers. Individuals who participated in the remote experts viewing, and experts who prepared bitstreams were also thanked.

NHK was thanked providing additional 4K HLG test sequences.

Thanks were also expressed to the editors, software coordinators, and all contributors to the development of the VVC and VSEI standards.

The 19h JVET meeting was closed at approximately 2100 hours UTC on Wednesday 1 July 2020.

# Annex A to JVET report: List of documents

# Annex B to JVET report: List of meeting participants

The participants of the nineteenth meeting of the JVET, according to an attendance sheet circulated during the meeting sessions (approximately XXX people in total), were as follows: