|  |  |
| --- | --- |
| **Joint Video Experts Team (JVET)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  11th Meeting: Ljubljana, SI, 10–18 July 2018 | Document: JVET-K\_Notes\_d8 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Meeting Report of the 11th meeting of the Joint Video Experts Team (JVET), Ljubljana, SI, 10–18 July 2018** | | |
| *Status:* | Report document from 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/ WG 11 held its eleventh meeting during 10–18 July 2018 at the GR – Ljubljana Exhibition and Convention Centre (Dunajska cesta 18, 1000 Ljubljana, Slovenia). 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 had modified it when entering the phase of formal development of a new standard by the previous meeting. The name Versatile Video Coding (VVC) was chosen as the informal nickname for the new standard.

The JVET meeting began at approximately 1300 hours on Tuesday 10 July 2018. Meeting sessions were held on all days (including weekend days) until the meeting was closed at approximately XXXX hours on Wednesday 18 July 2018. Approximately XXX people attended the JVET meeting, and approximately XXX input documents and 13 AHG reports were discussed. The meeting took place in a collocated fashion with a meeting of SG16 – one of the two parent bodies of the JVET. The subject matter of the JVET meeting activities consisted of developing video coding technology with a compression capability that significantly exceeds that of the current HEVC standard, or otherwise gives better support regarding the requirements of future application domains of video coding. As a primary goal, the JVET meeting reviewed the work that was performed in the interim period since the tenth JVET meeting in producing a first draft of the VVC standard and the first version of the associated VVC test model (VTM). Further important goals were reviewing the results of 13 Core Experiments (CE), reviewing other technical input on novel aspects of video coding technology, and producing the next versions of draft text and VTM, and plan next steps for further investigation of candidate technology towards the formal standard development.

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

* JVET-J1001 Versatile Video Coding specification text (Draft 1)
* JVET-J1002 Algorithm description for Versatile Video Coding and Test Model 1 (VTM 1)
* JVET-J1003 Report of results from the Call for Proposals on Video Compression with Capability beyond HEVC
* JVET-J1005 Methodology and reporting template for tool testing
* JVET-J1010, JVET-J1011, and JVET-J1012 JVET common test conditions and software reference configurations for SDR, HDR/WCG, and 360° video
* JVET-J1021 through JVET-J1033, Description of Core Experiments 1 through 13

For the organization and planning of its future work, the JVET established XX “ad hoc groups” (AHGs) to progress the work on particular subject areas. At this meeting, XX Core Experiments (CE) were defined. The next four JVET meetings were planned for 3–12 October 2018 under WG11 auspices in Macao, CN, during 9–18 January 2019 under WG11 auspices in Marrakesh, MA, during 19–27 March 2019 under ITU-T auspices in Geneva, CH, and during 4–12 July 2019 under WG11 auspices in Gothenburg, SE .

The document distribution site <http://phenix.it-sudparis.eu/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://mailman.rwth-aachen.de/mailman/listinfo/jvet>.

# 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 11 Moving Picture Experts Group (MPEG). The parent bodies of the JVET are ITU-T WP3/16 and ISO/IEC JTC 1/SC 29/WG 11.

The Joint Video Experts Team (JVET) of ITU-T WP3/16 and ISO/IEC JTC 1/ SC 29/ WG 11 held its eleventh meeting during 10–18 July 2018 at the GR – Ljubljana Exhibition and Convention Centre (Dunajska cesta 18, 1000 Ljubljana, Slovenia). 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 had modified it when entering the phase of formal development of a new standard by the previous meeting. The name Versatile Video Coding (VVC) was chosen as the informal nickname for the new standard.

## 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/2019_07_K_Ljubljana/>.

## Primary goals

As a primary goal, the JVET meeting reviewed the work that was performed in the interim period since the tenth JVET meeting in producing a first draft of the VVC standard and the first version of the associated VVC test model (VTM). Further important goals were reviewing the results of 13 Core Experiments (CE), reviewing other technical input on novel aspects of video coding technology, and producing the next versions of draft text and VTM, and plan next steps for further investigation of candidate technology towards the formal 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.it-sudparis.eu/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 JEM software but have no normative effect are marked by the string “Decision (SW):”.
* Decisions that fix a “bug” in the JEM description (an error, oversight, or messiness) or in the software are marked by the string “Decision (BF):”.

This meeting report is based primarily on notes taken by the responsible leaders. 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 Monday, 2 July 2018. Any documents uploaded after 1159 hours Paris/Geneva time on Tuesday 3 July were considered “officially late”, giving a grace period of 12 hours to accommodate those living in different time zones of the world.

All contribution documents with registration numbers JVET-K0385 and higher were registered after the “officially late” deadline (and therefore were also uploaded late). However, some documents in the “K0385+” range might include break-out activity reports that were generated during the meeting, and are therefore better considered as report documents rather than as late contributions.

In many cases, contributions were also revised after the initial version was uploaded. The contribution document archive website retains publicly-accessible prior versions in such cases. The timing of late document availability for contributions is generally noted in the section discussing each contribution in this report.

One suggestion to assist with the issue of late submissions was to require the submitters of late contributions and late revisions to describe the characteristics of the late or revised (or missing) material at the beginning of discussion of the contribution. This was agreed to be a helpful approach to be followed at the meeting.

There were no technical design proposal contributions that were registered on time but uploaded late for the current meeting.

The following technical design proposal contributions were registered and/or uploaded late:

* JVET-K0XXX (a proposal on … ), uploaded 07-XX.
* …

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

* JVET-K0XXX (an information document on …), uploaded 07-XX.
* …

The following cross-verification reports were registered and uploaded late: JVET-K0XXX [uploaded 07-XX], … .

The following contribution(s) registration were later cancelled, withdrawn, never provided, were cross-checks of a withdrawn contribution, or were registered in error: JVET-K0XXX, ….

“Placeholder” contribution documents that were basically empty of content, with perhaps only a brief abstract and some expression of an intent to provide a more complete submission as a revision, had been agreed to be considered unacceptable and rejected in the document management system. There were no initial uploads of contribution documents that were rejected as “placeholders” at the current meeting.

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, EE 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

The output documents of the previous meeting, particularly the meeting report JVET-K1000, the Versatile Video Coding specification text (Draft 1) JVET-J1001, the Algorithm description for Versatile Video Coding and Test Model 1 (VTM 1) JVET-J1002, the Report of results from the Call for Proposals on Video Compression with Capability beyond HEVC JVET-J1003, the Methodology and reporting template for tool testing JVET-J1005, the JVET common test conditions and software reference configurations for SDR, HDR/WCG, and 360° video (JVET-J1010, JVET-J1011, and JVET-J1012), and the Description of Core Experiments 1 through 13 (JVET-J1021 through JVET-J1033), were approved. The software implementations of VTM (versions 1.0 and 1.1), BMS (versions 1.0 and 1.1), and the 360Lib software implementation (version 6.0) were also approved.

The group had initially been asked to review the meeting report of the previous meeting for finalization. The meeting report was later approved without modification.

All 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.

## Agenda

The agenda for the meeting was as follows:

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

## 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 N10791 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 and HDRtools as well.

## Communication practices

The documents for the meeting can be found at <http://phenix.it-sudparis.eu/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://mailman.rwth-aachen.de/mailman/options/jvet>, 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 was 928.

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

Some terminology used in this report is explained below:

* **ACT**: Adaptive colour transform.
* **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**: Adaptive multi-core transform.
* **AMVR**: (Locally) adaptive motion vector resolution.
* **APS**: Active parameter sets.
* **ARC**: Adaptive resolution conversion (synonymous with DRC, and a form of RPR).
* **ARSS**: Adaptive reference sample smoothing.
* **ATMVP**: 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.
* **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).
* **BIO**: Bi-directional optical flow.
* **BL**: Base layer.
* **BoG**: Break-out group.
* **BR**: Bit rate.
* **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.
* **CCLM**: Cross-component linear model.
* **CCP**: Cross-component prediction.
* **CG**: Coefficient group.
* **CGS**: Colour gamut scalability (historically, coarse-grained scalability).
* **CL-RAS**: Cross-layer random-access skip.
* **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.
* **CTC**: Common test conditions.
* **CVS**: Coded video sequence.
* **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.
* **DRC**: Dynamic resolution conversion (synonymous with ARC, and a form of RPR).
* **DT**: Decoding time.
* **ECS**: Entropy coding synchronization (typically synonymous with WPP).
* **EE**: Exploration Experiment – a coordinated experiment conducted toward assessment of coding technology.
* **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.
* **ET**: Encoding time.
* **FRUC**: Frame rate up conversion (pattern matched motion vector derivation).
* **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).
* **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).
* **IPCM**: Intra pulse-code modulation (similar in spirit to IPCM in AVC and HEVC).
* **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.
* **LIC**: Local illumination compensation.
* **LM**: Linear model.
* **LP** or **LDP**: Low-delay P – the variant of the LD conditions that uses P frames.
* **LUT**: Look-up table.
* **LTRP**: Long-term reference pictures.
* **MC**: Motion compensation.
* **MCP**: Motion compensated prediction.
* **MDNSST**: Mode dependent non-separable secondary transform.
* **MMLM**: Multi-model (cross component) linear mode.
* **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).
* **MV**: Motion vector.
* **MVD**: Motion vector difference.
* **NAL**: Network abstraction layer (as in AVC and HEVC).
* **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.
* **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).
* **PDPC**: Position dependent (intra) prediction combination.
* **PMMVD**: Pattern-matched motion vector derivation.
* **POC**: Picture order count.
* **PoR**: Plan of record.
* **PPS**: Picture parameter set (as in AVC and HEVC).
* **QM**: Quantization matrix (as in AVC and HEVC).
* **QP**: Quantization parameter (as in AVC and HEVC, sometimes confused with quantization step size).
* **QT**: Quadtree.
* **BT**: Binary tree.
* **TT**: Ternary tree.
* **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.
* **RASL**: Random-access skipped leading.
* **R-D**: Rate-distortion.
* **RDO**: Rate-distortion optimization.
* **RDOQ**: Rate-distortion optimized quantization.
* **ROT**: Rotation operation for low-frequency transform coefficients.
* **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.
* **SD**: Slice data; alternatively, standard-definition.
* **SDT**: Signal dependent transform.
* **SEI**: Supplemental enhancement information (as in AVC and HEVC).
* **SH**: Slice header.
* **SHM**: Scalable HM.
* **SHVC**: Scalable high efficiency video coding.
* **SIMD**: Single instruction, multiple data.
* **SPS**: Sequence parameter set (as in AVC and HEVC).
* **STMVP**: Spatial-temporal motion vector prediction.
* **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.
* **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.
* **VTM**: VVC Test Model.
* **VVC**: Versatile Video Coding, the standardization project developed by JVET.
* **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 JEM:
  + **CTB**: Coding tree block (luma or chroma) – there are three CTBs per CTU in P/B slice, and one CTB per luma CTU and two CTBs per chroma CTU in I slice.
  + **CTU**: Coding tree unit (synonymous with LCU, containing both luma and chroma in P/B slice, containing only luma or chroma in I slice), 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 to a CU.
  + **TB**: Transform block, a luma or chroma block of a TU.
  + **TU**: Transform unit, has the same size to a CU.

## Opening remarks

Activities on the opening cross-checking day of the meeting 1300 Tuesday 10 April (chaired by GJS and JRO) were as follows.

* The meeting logistics, agenda, working practices, policies, and document allocation were reviewed.
* The results of the previous meeting were reviewed.
* The primary goal of the meeting was to review the results of CEs, and identify promising technology directions, and potential adoptions to the draft and VTM.
* Due to high number of inputs, parallelization and breakout work will be needed
* Principles of standards development were discussed.

## Scheduling of discussions

Scheduling: Generally meeting time was scheduled during 0900–2000+ hours, with coffee and lunch breaks as convenient. Ongoing scheduling refinements were announced on the group email reflector as needed. Some particular scheduling notes are shown below, although not necessarily 100% accurate or complete:

* Tue. 10 July, 1st day
  + 1300–1720 Opening, AHG reports (chaired by GJS & JRO)
  + 1720–2020 CE1 review (overall report, detailed discussion Sub-CE1&2) (chaired by JRO)
* Wed. 11 July, 2nd day
  + 0900–1100 CE1 review cont. (Sub-CE3..8) (chaired by JRO)
  + 1130–1300 Track B: CE2.1/2.2 (chaired by JRO)
  + 1330–1900 Track B: CE2.3-2.5 (chaired by JRO)
  + 1900–2100 Track B: CE4.1, aspects of affine motion comp (chaired by JRO)
  + 1140–1340, 1500-1800 Track A (chaired by GJS) CE3 Intra prediction and mode coding
  + 1820–2000 Track A (chaired by GJS) CE5 Arithmetic coding engine
* Thu. 12 July, 3rd day
  + 0900-1200 BoG (chaired by L. Zhang) on CE2 ALF, BoG (chaired by B. Bross) on CE1.1 (Partitioning), and BoG (chaired by K. Misra) on CE1.2 (Partition boundaries)
  + 1400–1500 Track A (chaired by GJS) CE5 Arithmetic coding engine
  + 1500–1920 Track A (chaired by GJS) CE6 Transforms and transform signalling
  + 1920–2000 Track A (chaired by GJS) CE7: Quantization and coefficient coding
  + 1300-2100 Track A (chaired by JRO) CE4 Inter prediction and MV coding
  + 2000-2200 BoG (chaired by B. Bross) on CE1.1 (Partitioning)
* Fri. 13 July, 4th day
  + 0900-1040 Track B (chaired by JRO) on CE4 Inter prediction and MV coding
  + 1100-1230 Track B (chaired by JRO) on CE9 Decoder side MV derivation
  + 1600-2030 Track B (chaired by JRO) CE10, CE11 and CE11 related
  + 0900–1100 Track A (chaired by GJS) CE7: Quantization and coefficient coding
  + 1000 BoG (chaired by J. Boyce) on 360° CE13 and non-CE
  + 1400 plenary (chaired by GJS & JRO)
  + 1600 Track A (chaired by GJS) CE5 related arith coding engine
  + 1700 Track A (chaired by GJS) CE6 related transforms and transform signalling
* Sat. 14 July, 5th day
  + 0900-1030 Track A (chaired by GJS) CE6 & CE6 related transforms and transform signalling
  + 1045 Track A (chaired by GJS) CE7 related – Quantization and coefficient coding
  + 1145 Track A (chaired by GJS) CE8 related – Current picture referencing
  + 1215 Track A (chaired by GJS) CE6 & CE6 related transforms and transform signalling
  + 0900-1040 Track B (chaired by JRO) on CE4 revisits
  + 1115-1300 and 1415-1615 Track B (chaired by JRO) on CE9 related contributions
  + 1630-2120 Track B (chaired by JRO) on CE2 revisits
  + 2120-2230 Track B (chaired by JRO) on CE10 related contributions
  + 0930-1300 BoG on ALF (chaired by L. Zhang)
  + 1100-1300 and 1400-1900 BoG on CE4 related contributions (chaired by H. Yang)
* Sun. 15 July, 6th day
  + 0900-1100 Track B (chaired by JRO) CE2 related: Loop filters
  + 1130-1345 Track B (chaired by JRO) CE4 related: Inter prediction and MV coding, BoG reviews
  + 1430-1800 JVET plenary: Reports from tracks, planning, complexity (8), HLS (7.17)
* Mon. 16 July, 7th day
  + 1400-1640 Track B (chaired by JRO) Revisits, remaining docs, NN technology (7.14)
  + 1400-1640 Track A (chaired by GJS) Revisits, remaining docs
  + 1650-1800 Joint meeting with parent bodies on project development
  + 1815-2030 JVET plenary: Setup of CEs, further planning
* Tue. 17 July, 8th day
  + 0900-1100 Track B (chaired by JRO) Review coding studies (4.3) and “other” (7.18/7.19); review BoG reports
  + 0900-1100 Track A (chaired by GJS) Revisits, remaining docs
  + 1100-1345 JVET plenary (chaired by GJS and JRO): Review tracks, BoG 360, High level syntax
  + 1445-1610 Track B (chaired by JRO) Review remaining documents from CE3 related; review specification text of ALF and affine
  + 1610-1800 JVET plenary (chaired by GJS and JRO): Setup AHGs, plan output docs

## Contribution topic overview

The approximate subject categories and quantity of contributions per category for the meeting were summarized as follows:

* AHG reports (13) (section 3) (Plenary)
* Project development (12) (section 4) (Plenary)
* Test material (2) (section 5) (Plenary)
* Core Experiments (247) (section 6) with subtopics
  + CE1: Partitioning (32) (section 6.1) (Track A) - BoG report
  + CE2: Loop filters (30) (section 6.2) (Track B)
  + CE3: Intra prediction and mode coding (38) (section 6.3) (Track A)
  + CE4: Inter prediction and motion vector coding (35) (section 6.4) (Track B)
  + CE5: Arithmetic coding engine (9) (section 6.5) (Track A)
  + CE6: Transforms and transform signalling (26) (section 6.6) (Track A)
  + CE7: Quantization and coefficient coding (12) (section 6.7) (Track A)
  + CE8: Current picture referencing (5) (section 6.8) (Track A)
  + CE9: Decoder side motion vector derivation (25) (section 6.9) (Track B)
  + CE10: Combined and multi-hypothesis prediction (9) (section 6.10) (Track B)
  + CE11: Composite reference pictures (4) (section 6.11) (Track B)
  + CE12: Mapping for HDR content (4) (section 6.12) (BoG)
  + CE13: Projection formats (8) (section 6.13) (BoG)
* Non-CE technology proposals (164) (section 7) with subtopics
  + CE1 related – Partitioning (9) (section 7.1) (Track A)
  + CE2 related – Loop filters (15) (section 7.2) (Track B)
  + CE3 related – Intra prediction and mode coding (26) (section 7.3) (BoG/Track A) [defer K0348/K0276]
  + CE4 related – Inter prediction and motion vector coding (41) (section 7.4) (BoG/Track B)
  + CE5 related – Arithmetic coding engine (3) (section 7.5) (Track A)
  + CE6 related – Transforms and transform signalling (19) (section 7.6) (Track A)
  + CE7 related – Quantization and coefficient coding (7) (section 7.7) (Track A)
  + CE8 related – Current picture referencing (3) (section 7.8) (Track A)
  + CE9 related – Decoder side motion vector derivation (16) (section 7.9) (Track B)
  + CE10 related – Combined and multi-hypothesis prediction (5) (section 7.10) (Track B)
  + CE11 related – Composite reference pictures (3) (section 7.11) (Track B)
  + CE12 related – Mapping for HDR content (1) (section 7.12) (BoG)
  + CE13 related – Projection formats (1) (section 7.13) (BoG)
  + NN technology related (5) (section 7.14) (Track B)
  + 360° video related (5) (section 7.15) (BoG)
  + Extended colour volume related (0) (section 7.16) (BoG)
  + HL syntax (section 7.17) (5) (Plenary)
  + Resilient intra refresh (section 7.18) (Track A)
  + Palette mode (section 7.19) (Track A)
* Complexity analysis and reduction (7) (section 8) (Plenary)
* Encoder optimization (1) (section 9) (Plenary)
* Metrics and evaluation criteria (1) (section 10) (Plenary)
* Withdrawn (8) (section 11)
* Joint meetings, plenary discussions, BoG reports, Summary of actions (section 12)
* Project planning (section 13)
* Establishment of AHGs (section 14)
* Output documents (section 15)
* Future meeting plans and concluding remarks (section 16)

# AHG reports (13)

These reports were discussed Tuesday 10 July 1400–1700 (chaired by GJS and JRO).

[JVET-K0001](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3782) 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.

In the interim period since the 10th JVET meeting, work towards finalizing the following (13) documents had been performed:

* JVET-J1001 Versatile Video Coding specification text (Draft 1)
* JVET-J1002 Algorithm description for Versatile Video Coding and Test Model 1 (VTM 1)
* JVET-J1003 Report of results from the Call for Proposals on Video Compression with Capability beyond HEVC
* JVET-J1005 Methodology and reporting template for tool testing
* JVET-J1010, JVET-J1011, and JVET-J1012 JVET common test conditions and software reference configurations for SDR, HDR/WCG, and 360° video
* JVET-J1021 through JVET-J1033, Description of Core Experiments 1 through 13

The work of the JVET overall had proceeded well in the interim period with a huge number of input documents submitted to the current meeting. Intense discussion had been carried out on the group email reflector, and all but one output documents from the preceding meeting had been produced.

Except as noted below, output documents from the preceding meeting had been made available at the "Phenix" site (<http://phenix.it-sudparis.eu/jvet/>) or the ITU-based JCT-VC site (<http://wftp3.itu.int/av-arch/jvet-site/2018_04_J_SanDiego/>), particularly including the following:

* The meeting report (JVET-J1000) [Posted 2018-07-10]
* Versatile Video Coding (Draft 1) (JVET-J1001) [Posted 2018-05-09]
* Algorithm description for Versatile Video Coding and Test Model 1 (VTM 1) (JVET-J1002) [Posted 2018-05-09]
* Report of results from the Call for Proposals on Video Compression with Capability beyond HEVC (JVET-J1003) [Posted 2018-07-XX]
* Methodology and reporting template for tool testing (JVET-J1010) [Posted 2018-04-25]
* JVET common test conditions and software reference configurations (JVET-J1010) [Posted 2018-04-24]
* JVET common test conditions and evaluation procedures for HDR/WCG video (JVET-J1011) [Posted 2018-05-02]
* JVET common test conditions and evaluation procedures for 360° video (JVET-J1012) [Posted 2018-04-28]
* Description of CE 1..13 (JVET-J1021..33) [Posted 2018-04-20]

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

Software integration of the VTM and BMS was finalized approximately according to the plan.

Various problem reports relating to asserted bugs in the software, draft specification text, and reference encoder description had been submitted to an informal "bug tracking" system. That system is not intended as a replacement of our ordinary contribution submission process. However, the bug tracking system was considered to have been helpful to the software coordinators and text editors. The bug tracker reports had been automatically forwarded to the group email reflector, where the issues were discussed – and this is reported to have been helpful.

It is foreseen to migrate software distribution as well as bug tracking to GitLab after the current meeting. So far, SVN had been used in the interest of making the software available in a timely fashion.

More than 400 input contributions to the current meeting (not counting the AHG reports) had been registered for consideration at the meeting. Most of these relate to Core Experiments.

Based on studies done in AHG 3, a basic summary of the bit-rate savings (PSNR-based CTC BD-Rate) achieved in the work on VVC, relative to the HM16.18 HEVC reference software (with 10 bit encoding) is as shown in the table below for the random access case:

**Random Access Configuration Comparison**

|  |  |  |
| --- | --- | --- |
| **vs HM16.18** | **VTM** | **BMS** |
| **4k UHD** | 10% | 28% |
| **1080p** | 8% | 22% |
| **WVGA** | 6% | 19% |
| **Average** | **8%** | **23%** |
| **Decode time** | 0.8× | 2× |
| **Encode time** | 2× | 9× |

It is noted that there have been various discussions on how to manage the project work and structure its results. Since the VVC project was just formally launched at the previous meeting, the current meeting may be a good opportunity to establish and refine working practices for the effort.

* We note that a multi-company contribution JVET-K0263 advocates for the document text structure to be somewhat different from what was previously done for AVC and HEVC, by splitting the design of the standard into multiple documents that are developed together.
* Another multi-company contribution JVET-K0311 advocates for an approach to interoperability point signalling for VVC with finer granularity than in past efforts, based on subsets of profiles & levels defined by JVET in the VVC specification and using a non-normative user-registered sub-profile indicator.
* Another aspect of project management that became evident in the interim work on core experiments is how to manage the need to use two different reference configurations in experiment comparisons, i.e., the VTM and BMS.
* Potential approaches to project management that have been discussed recently in the parent bodies and among the management have included external profiling, switchable fall-back modes, the priority of the proponent of an adopted feature in the work on subsequent refinement of the feature design, coordinated development of multiple “modular” standards for aspects previously all specified in the same document, multiple phases of profile development, avoiding inheritance of technology “by default” across generations, and voluntary expressions of licensing timelines in proposal rights declarations.

These and other aspects of project management are suggested to be considered in the parent bodies and JVET, as appropriate, at the current meeting.

A preliminary basis for the document subject allocation and meeting notes for the 11th meeting had been circulated to the participants by being uploaded to the ITU-hosted ftp site that is routinely used for that purpose.

[JVET-K0002](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3789) JVET AHG report: Draft text and test model algorithm description editing (AHG2) [E. Alshina, B. Bross, J. Chen]

This document reports the work of the JVET ad hoc group on draft text and test model algorithm description editing (AHG2) between the 10th Meeting in San Diego, US (10–20 Apr 2018) and the 11th meeting in Ljubljana, SI (10–18 July 2018).

At the 10th JVET meeting, JVET defined the first draft of Versatile Video Coding (VVC) (JVET-J1001) and the VVC Test Model 1 (VTM1) encoding method (JVET-J1002). It was decided to include a quadtree with nested multi-type tree using binary and ternary splits coding block structure as the initial new coding feature of VVC. Draft reference software to implement the VVC decoding process and VTM1 encoding method has also been developed.

The normative decoding process for Versatile Video Coding is specified in the VVC draft 1 text specification document. This VVC Test Model 1 (VTM 1) Algorithm and Encoder Description document provides an algorithm description as well as an encoder-side description of the VVC Test Model 1, which serves as a tutorial for the algorithm and encoding model implemented in the VTM1.0 software.

Two versions of JVET-J1001 and two versions of JVET-J1002 were published by the Editing AHG between the 10th Meeting in San Diego (10–20 Apr 2018) and the 11th meeting (10–18 July 2018).

JVET-J1001 has been established from scratch and now contains the following:

* Basic definitions, abbreviations and conventions
* A basic high-level syntax (HLS) with NAL units, SPS, PPS and slice header.
* Block partitioning by a quadtree with nested multi-type tree using binary and ternary splits with
  + CU leaf nodes
  + Prediction at CU level
  + Transform at CU level
  + Minimum CU size with 4x4 luma coding block and corresponding chroma coding blocks (2x2 for 4:2:0)
  + Maximum TU size with 64x64 luma transform block and corresponding chroma transform blocks (32x32 for 4:2:0)
  + Minimum TU size with 4x4 luma transform block and corresponding chroma transform blocks (2x2 for 4:2:0)
  + Single tree for luma and chroma

JVET-J1002 has also been established from scratch. The document generally describes the basic coding architecture, the partitioning of the picture into CTUs, and the partitioning of the CTUs using a quadtree with nested multi-type tree.

For initial testing purposes of the aspects of the design that have not yet been determined, the test model software uses syntax, semantics, and decoding processes that correspond to those in prior well-known video coding designs. However, these aspects are considered only to be “placeholders” for specific design details yet to be determined. The exact details of the binary/ternary/quaternary segmentation tree structure to be used are also yet to be determined. This document may contain a description of some such details that should not be considered completely agreed upon.

As agreed in the 10th JVET meeting, the following features that are found in HEVC are not included in the initial VVC test model.

* Special strong boundary smoothing for 32×32 luma block intra prediction
* Boundary smoothing across edges for intra prediction (a horizontal filter for vertical prediction and vice versa, and the first row and column with DC prediction)
* DST-VII style transform in 4×4 intra blocks
* Mode-dependent scan for intra blocks
* Quantization weighting matrices
* Residual sign bit hiding
* VPS and VPS VUI
* Dependent slices
* Tiles
* Wavefronts (entropy coding sync)

In terms of the impact of this on specific elements of the design, this includes removal of the following features (and some others):

* Partitioning of a CU into multiple PUs (including asymmetric partitionings)
* Partitioning of a CU into multiple luma blocks for intra prediction (i.e., signalling of multiple luma intra prediction modes for a CU), except for implicit splits when the CU size is too large for the maximum transform size
* The coding unit syntax element part\_mode
* Partitioning of a CU into multiple TUs, except for implicit splits when the CU size is too large for the maximum transform size
* Transforms that are applied across prediction block boundaries
* The syntax element split\_transform\_flag
* Non-aligned luma and chroma transform blocks
* All VPS and VPS VUI syntax
* SPS syntax elements
  + log2\_min\_luma\_transform\_block\_size\_minus2 (always use 4x4 luma and corresponding chroma)
  + log2\_diff\_max\_min\_luma\_transform\_block\_size
  + max\_transform\_hierarchy\_depth\_inter
  + max\_transform\_hierarchy\_depth\_intra
  + amp\_enabled\_flag

The AHG recommended to:

Approve the edited JVET-J1001 and JVET-J1002 documents as the JVET outputs

* Continue to edit the VVC WD and Test Model documents to ensure that all agreed elements of VVC are fully described
* Compare the VVC documents with the VVC software and resolve any discrepancies that may exist, in collaboration with the Software AHG
* Continue to improve the editorial consistency of VVC WD and Test Model documents
* Ensure that, when considering the addition of new feature to VVC, properly drafted text for addition to the VVC Test Model and/or the VVC Working Draft is made available in a timely manner

[JVET-K0003](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3989) 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 10th and 11th JVET meetings.

Initial versions of VTM and BMS were checked into SVN repositories. Versions 1.0 and 1.1 of both were released. JEM 7.2 was released, with a corresponding version of BMS (BMS-0), that can cross-decode bitstreams. Software development guidelines are proposed in JVET-K0461. Moving the software development from SVN to git (GitLab) was proposed.

The VTM software can be found at

https://jvet.hhi.fraunhofer.de/svn/svn\_VVCSoftware\_VTM/

The BMS software can be found at:

https://jvet.hhi.fraunhofer.de/svn/svn\_VVCSoftware\_BMS/

After three release candidates, VTM 1.0 and BMS 1.0 were tagged on May 17, 2018. This version reflects all meeting decisions regarding tool integration. Tools were removed by moving the code into macros, which are disabled. The BMS software still contains all disabled code. VTM 1.0 was derived from BMS 1.0 by stripping the JEM\_TOOLS macro.

VTM 1.1 and BMS 1.1 were tagged on June 1, 2018, with the following changes:

* WPSNR for HDR
* A fix for the SIMD config setting being ignored at the decoder
* A fix for ALF decoding with low QP values
* A fix for adaptive luma QP
* Fixes for config files

The following shows VTM 1.0 performance over HM 16.18:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **All Intra Main10** | | | | |
|  | **Over HM 16.18** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | −5.06% | −9.60% | −10.38% | 860% | 102% |
| Class A2 | −4.75% | −11.83% | −11.88% | 920% | 109% |
| Class B | −3.38% | −12.53% | −13.55% | 982% | 113% |
| Class C | −3.32% | −11.90% | −12.03% | 1139% | 119% |
| Class E | −5.29% | −12.96% | −12.87% | 862% | 103% |
| **Overall** | −4.19% | −11.86% | −12.29% | 961% | 110% |
| Class D | −2.60% | −10.85% | −11.22% | 1104% | 110% |
|  |  |  |  |  |  |
|  | **Random Access Main 10** | | | | |
|  | **Over HM 16.18** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 | −10.44% | −14.37% | −15.90% | 253% | 78% |
| Class A2 | −10.43% | −18.14% | −17.05% | 224% | 78% |
| Class B | −7.53% | −17.60% | −16.91% | 206% | 76% |
| Class C | −6.49% | −14.69% | −15.18% | 233% | 92% |
| Class E |  |  |  |  |  |
| **Overall** | −8.42% | −16.28% | −16.28% | 225% | 81% |
| Class D | −5.19% | −14.11% | −14.03% | 206% | 90% |
|  |  |  |  |  |  |
|  | **Low delay B Main10** | | | | |
|  | **Over HM 16.18** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | −8.02% | −17.30% | −18.11% | 179% | 82% |
| Class C | −7.18% | −12.84% | −14.36% | 202% | 97% |
| Class E | −10.77% | −22.53% | −22.95% | 89% | 72% |
| **Overall** | −8.43% | −17.12% | −18.07% | 156% | 84% |
| Class D | −6.01% | −9.64% | −9.99% | 195% | 101% |
|  |  |  |  |  |  |
|  | **Low delay P Main10** | | | | |
|  | **Over HM 16.18** | | | | |
|  | Y | U | V | EncT | DecT |
| Class A1 |  |  |  |  |  |
| Class A2 |  |  |  |  |  |
| Class B | −8.90% | −20.17% | −20.51% | 175% | 89% |
| Class C | −7.44% | −14.02% | −15.18% | 191% | 102% |
| Class E | −11.70% | −25.27% | −25.58% | 91% | 79% |
| **Overall** | −9.11% | −19.40% | −20.00% | 153% | 90% |
| Class D | −6.18% | −10.46% | −11.12% | 181% | 108% |

Full results for the VTM and BMS are attached to this AHG report as Excel files.

JEM and the NextSoftware were modified to allow bitstream cross-decoding. The resulting versions were released as JEM 7.2 and BMS-0, which was the initial check-in of BMS:

https://jvet.hhi.fraunhofer.de/svn/svn\_HMJEMSoftware/tags/HM-16.6-JEM-7.2/

https://jvet.hhi.fraunhofer.de/svn/svn\_VVCSoftware\_BMS/tags/BMS-0

An area was created in the BMS repository with restricted read and write access. Approximately 70 company accounts were created to allow access for CE participants. Base directories were created for each CE, in which the CE coordinators could create the appropriate branches for sub-CEs and test. 288 branches were created for CE software development.

Guidelines for software development are proposed in JVET-K0461. These guidelines are derived from guidelines previously used for HM software development, but contain several changes detailed in the document.

To improve the software development process, it is desirable to switch from subversion to git. Especially the GitLab environment provides many features that are helpful for CE software coordination:

* Personal accounts (instead of company account)
* Cloning of repositories for CEs and software submissions
* Merge requests
* Multiple level of access right management, i.e. CE coordinators can assign access rights themselves.

A GitLab server was set up by Fraunhofer HHI that can host the git repositories in the future. It allows user registration for contributing software after adoption, or for CEs.

The JEM bug tracker was extended to also allow filing bugs for VTM, BMS and specification text. It is now located at:

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

The old URLs will continue to work and are forwarding to the new location.

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 recommends to:

* Continue to develop the VTM reference software
* 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 bit-streams/test cases that trigger bugs in VTM.
* Adopt the proposed guidelines for software development
* Switch to a git server for software development

It was remarked that for RA, the VTM decoder is actually faster than the HM decoder. This seemed to be the result of using larger block sizes, and possibly some difference in SIMD optimization.

[JVET-K0004](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3932) JVET AHG report: Test material and visual assessment (AHG4) [V. Baroncini, R. Chernyak, P. Hanhart, A. Norkin, T. Suzuki, J. Ye]

The test sequences used for CfP (JVET-H1002) are available on ftp://jvet@ftp.ient.rwth-aachen.de in directory “/jvet-cfp” (accredited members of JCT-VC may contact the JCT-VC 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).

The test sequences were provided for JVET standardization purposes. The AHG recommended reminding the participants of the copyright terms of each of the test sequences. JVET members must not use JVET test sequences for purposes that are not allowed under the associated copyright release.

A contribution JVET-K0409 had been submitted regarding the copyright status of some test sequences.

New test sequences were offered by Tencent in JVET-K0294.

The AHG recommended:

* To review all related contribution
* To remind JVET members on copyright and usage of JVET test sequences
* To continue to collect new test sequences available for JVET with licensing statement

It was remarked that the CTC document describes where to find the test sequences for the CTC. It was requested to provide information in a revision of the AHG report about where to find the CTC sequences – perhaps just referring to the CTC document for where to get the information.

[JVET-K0005](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3987) JVET AHG Report: Memory bandwidth consumption of coding tools (AHG5) [R. Hashimoto, E. Alshina, T. Ikai, H. Yang, M. Zhou]

The document summarizes activities of AhG on memory bandwidth consumption of coding tools between the 10th and the 11th JVET meetings.

There was no related email discussion on the JVET reflector during this meeting cycle.

Contributions to this meeting are as follows.

* JVET-K0451 “AHG5: How to use the software to evaluate memory bandwidth”, R. Hashimoto, S. Mochizuki (Renesas)
* JVET-K0452 “AHG5: Proposal of template for comparing memory bandwidth”, R. Hashimoto, S. Mochizuki (Renesas)

The AHG recommended to review the related contributions.

A participant requested for an analysis to be performed of the BMS and VTM memory bandwidth relative to the HM.

Another participant requested that the analysis consider cache blocks rather than cache lines. The presenter said this may be available soon.

[JVET-K0006](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3948) JVET AHG report: 360 video conversion software development (AHG6) [Y. He, K. Choi]

The document summarizes activities on 360-degree video content conversion software development between the 10th (10 – 20 Apr. 2018) and the 11th (10 – 18 Jul. 2018) JVET meetings.

The 360Lib-6.0 software package included following changes:

* Software:
  + 360Lib interfaces changes for VTM/BMS integration;
  + The patch for VTM/BMS for the integration;
* 360Lib-6.0 related releases:
  + Viewport FOV size and resolution in viewport configurations are changed according to 360 CTC(JVET-J1012);
* Configurations:
  + 360Lib-6.0 related release:
  + 360Lib-6.0rc1 with support of VTM-1.0rc4 and BMS-1.0rc3 was released on May 15, 2018;
  + 360Lib-6.0 with support of VTM-1.0 and BMS-1.0 was released on May 21, 2018;

One bug (#58) was reported by Intel, and the bug is related to PERP full resolution coding and does not impact on CTC coding; The bug fix was checked in 360Lib-6.1-dev branch;

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

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

The released version of 360Lib-6.0 can be found at:

https://jvet.hhi.fraunhofer.de/svn/svn\_360Lib/tags/360Lib-6.0/

360Lib-6.0 testing results can be found at:

ftp.ient.rwth-aachen.de/testresults/360Lib-6.0

360Lib bug tracker

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

The tables below are for the projection formats comparison using VTM-1.0 and BMS-1.0 according to 360o video CTC (JVET-J1012). The first lists the VTM-1.0 CMP coding performance compared to PERP coding. The second compares the BMS-1.0 CMP coding with BMS-1.0 PERP coding. The third and fourth are for codec comparison under PERP and CMP projection formats.

**VTM-1.0 CMP vs PERP (VTM-1.0 PERP coding as anchor)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **CMP over PERP (VTM-1.0)** | | | | | |
|  | **End-to-end WS-PSNR** | | | **End-to-end S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | -5.48% | -2.80% | -2.72% | -5.64% | -2.82% | -2.73% |
| Class S2 | -0.50% | 1.28% | 1.07% | -0.50% | 1.38% | 1.14% |
| **Overall** | -3.49% | -1.17% | -1.21% | -3.59% | -1.14% | -1.18% |

**BMS-1.0 CMP vs PERP (BMS-1.0 PERP coding as anchor)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **CMP over PERP (BMS-1.0)** | | | | | |
|  | **End-to-end WS-PSNR** | | | **End-to-end S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | -5.31% | -1.19% | -2.71% | -5.47% | -1.25% | -2.77% |
| Class S2 | -0.50% | 1.99% | 1.54% | -0.53% | 2.07% | 1.63% |
| **Overall** | -3.39% | 0.08% | -1.01% | -3.49% | 0.08% | -1.01% |

**BMS-1.0 PERP vs VTM-1.0 PERP (VTM-1.0 PERP coding as anchor)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **PERP - Over TM-1.0** | | | | | |
|  | **End-to-end WS-PSNR** | | | **End-to-end S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | -11.17% | -17.02% | -17.46% | -11.18% | -17.03% | -17.43% |
| Class S2 | -15.20% | -18.29% | -18.20% | -15.19% | -18.28% | -18.17% |
| **Overall** | -12.78% | -17.53% | -17.75% | -12.78% | -17.53% | -17.73% |

**BMS-1.0 CMP vs VTM-1.0 CMP (VTM-1.0 CMP coding as anchor)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **CMP - Over TM-1.0** | | | | | |
|  | **End-to-end WS-PSNR** | | | **End-to-end S-PSNR-NN** | | |
|  | Y | U | V | Y | U | V |
| Class S1 | -10.58% | -15.47% | -17.06% | -10.59% | -15.50% | -17.07% |
| Class S2 | -14.45% | -16.97% | -16.99% | -14.45% | -16.98% | -16.97% |
| **Overall** | -12.13% | -16.07% | -17.03% | -12.13% | -16.09% | -17.03% |

The AHG recommended:

* To coordinate with AHG3 to integrate those 360Lib interface related changes to VTM-2.0;
* To continue software development of the 360Lib software package.

[JVET-K0007](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3980) JVET AHG report: Coding of HDR/WCG material (AHG7) [A. Segall, E. François, D. Rusanovskyy]

This document summarizes the activity of AHG7: Coding of HDR/WCG Material between the 10th meeting in San Diego, US (10 – 20 Apr. 2018) and the 11th meeting in Ljubljana, SI (10 – 18 July 2018).

The AHG used the main JVET reflector, jvet@lists.rwth-aachen.de, with an [AHG7] indication on message headers. The primary activity of the AhG was related to the mandates of (i) coordinating the implantation of HDR anchor aspects in the test software and (ii) studying and evaluating available HDR/WCG test content with the goal of reducing the number of frames in the HLG sequences. This work is described in the following subsections.

During the AHG study period, it was reported that versions 1.0 of the VTM and BMS software encoder generated significant visual artifacts for PQ content. After study, it was determined that the artifact could be removed by defining the WCG\_EXT macro in the software. This macro had the additional benefit in providing improved wPSNR performance, as it activates the weighted RDO using weights corresponding to the wPSNR metrics calculation.

Also during the AhG study period, a patch was provided for the VTM and BMS software. This patch is commit id 38 in the repository, and it was included in version 1.1 of the VTM and BMS software. After study, it was determined that the patch did not change the encoder performance when the WCG\_EXT macro was enabled. Furthermore, it was determined that the enabling the WCG\_EXT macro provided more coding gain in the wPSNR metrics than did enabling commit id 38.

As a result of the study, the CTC was updated to clarify that the WCG\_EXT macro should be enabled when testing HDR content, and anchors were generated using this macro setting.

It was further noted that it may be desirable to have this configuration supported in a future version of the VTM and BMS software.

The group also had the mandate to study and evaluate available HDR test content, with one goal being the reduction of the number of frames used for testing in the HLG sequences. To make progress on the mandate, the AhG used the GammutTest software that is part of the HDRTools repository to compute the dynamic range of each sequence as a function of time. The results are provided as part of the AhG report.

To assist in the selection of frames included in the test conditions, plots of the dynamic range were also computed.

There are 3 contributions related to HDR video coding:

JVET-K0032 m43229 CE12: Summary report on HDR coding E. Francois, D. Rusanovskyy, P. Yin

JVET-K0298 m43231 CE12: Report of Dynamic Range Adaptation (DRA) and DRA refinement E. Francois (Technicolor), D. Rusanovskyy (Qualcomm)

JVET-K0308 m43241 CE12: HDR In-loop Reshaping (CE12-5, 12-6, 12-7 and 12-8) T. Lu, F. Pu, P. Yin, W. Husak, S. McCarthy, T. Chen (Dolby)

The AHG recommended the following:

* Enable the WCG\_EXT macro in the next version of the VTM/BMS software
* Select 300 frame versions of the HLG sequences during the meeting
* Review all input contributions

[JVET-K0008](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3988) JVET AHG report: 360° video coding tools and test conditions (AHG8) [J. Boyce, G. v. d. Auwera, K. Choi, P. Hanhart]

This document summarizes the activity of AHG8: 360° video coding tools and test conditions between the between the 10th Meeting in San Diego, US (10 – 20 Apr 2018) and the 11th meeting in Ljubljana, SI (0–18 July 2018).

There was no AHG email activity on the main jvet reflector

There are 5 non-CE contributions related to 360° video coding, which are listed below. In addition, CE13 on projection formats is related to 360° video coding, and has 7 contributions, and one additional CE13-related contribution.

Most related contributions are directed at handling of projection face boundaries, with many claiming subjective quality improvements.

JVET-K0141

AHG8: 360°-based inter/intra prediction for cubemap projection C.-H. Shih, J.-L. Lin, H.-C. Lin, S.-K. Chang, C.-C. Ju (MediaTek)

JVET-K0142

AHG8: 360°-based in-loop filters for cubemap projection S.-Y. Lin, L. Liu, C.-H. Shih, J.-L. Lin, H.-C. Lin, S.-K. Chang, C.-C. Ju (MediaTek)

JVET-K0183

AHG8：Face boundary filtering for 360° video X. Huangfu, Y. Sun, B. Wang, L. Yu (Zhejiang Univ.)

JVET-K0333 AHG8: Horizontal geometry padding for PERP P. Hanhart, Y. He, Y. Ye (InterDigital)

JVET-K0404 AHG8: Selective In-loop filtering for 360 Video Compression C. Pujara, S.N. Akula, A. Singh, R. Narayana, W. Choi (Samsung)

The AHG recommends the following:

* Review input contributions
* Conduct informal subjective viewing of contributions
* Review common test conditions for 360° video, including objective metrics and viewports
* Review 360° video test material, and consider adding or replacing test sequences for common test conditions

[JVET-K0009](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3963) JVET AHG report: Neural Networks in Video Coding (AHG9) [S. Liu, B. Choi, K. Kawamura, Y. Li, L. Wang, P. Wu, H. Yang]

This document summarizes the activity of AHG9: Neural network in video coding between the 10th meeting in San Diego, US (10 – 20 April 2018) and the 11th meeting in Ljubljana, SI (10 – 18 July 2018).

There was no relevant email exchange on the main reflector

Some input documents (technical proposals) related to AHG9 are summarized as follows.

* JVET-K0158 “AHG9: Separable Convolutional Neural Network Filter with Squeeze-and-Excitation block”, T. Hashimoto, E. Sasaki, T. Ikai (Sharp).
* JVET-K0222 “AHG9: Convolution neural network loop filter”, Y.-L. Hsiao, T.-D. Chuang, C.-Y. Chen, C.-W. Hsu, Y.-W. Huang, S.-M. Lei (MediaTek).
* JVET-K0391 “AHG9: Dense Residual Convolutional Neural Network based In-Loop Filter”, Y. Wang, Z. Chen, Y. Li (Wuhan Univ.), L. Zhao (Tencent).
* JVET-K0266 “CE3: Non-linear weighted intra prediction (Test 6.1.1)”, P. Merkle, J. Pfaff, P. Helle, R. Rischke, M. Schäfer, B. Stallenberger, H. Schwarz, D. Marpe, T. Wiegand (Fraunhofer HHI)

The AHG recommended:

* To review all related contributions
* To continue discussions about methodologies and measurements for evaluating neural network related video coding tools

It was hoped that the level of interest on this would increase, as there was substantial related material submitted in response.

[JVET-K0010](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3997) JVET AHG report: Encoding algorithm optimizations (AHG10) [R. Sjöberg, E. Alshina, C. Helmrich, S. Ikonin, A. Norkin]

The document summarizes the activities of the AHG on Encoding algorithm optimizations between the 10th meeting in San Diego, USA (April 10-20, 2018) and the 11th meeting in Ljubljana, SI (10-18, July 2018).

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

* JVET-K0108: Decoding-Energy-Rate-Distortion Optimization by Friedrich-Alexander University Erlangen-Nürnberg (FAU)
  + This contribution proposes to include Decoding-Energy-Rate-Distortion Optimization (DERDO) into the encoder reference software. DERDO can be used to control and minimize the decoding energy. Potential savings have been analyzed for intra only coding and shows 5.74% energy savings at the expense of a bit rate increase of 7.57%. The authors claim that energy savings of a higher magnitude at lower rate increases can be expected for inter prediction coding tools.
* JVET-K0154: On encoding distortion evaluation of VTM/BMS software by Sharp Corporation
  + In this contribution, it is proposed to use 10-bit distortion evaluation instead of 8-bit distortion for the next VVC software and CTC. For VTM AI/RA/LDB/LDP, the contribution reports the following BD-rate numbers for the proposal: -0.04%/-0.02%/-0.01%/-0.03%
  + The contribution also reports a bug related to bit-depth for the BMS ALF tool.
  + In addition, the contribution also reports that “There seem to be several bugs of tools with non-BMS configuration; however, we will not discuss them at this time.”
* JVET-K0206: AHG10: Improved perceptually optimized QP adaptation and associated distortion measure by Fraunhofer HHI
  + This input document describes seven improvements to the QP adaptation scheme that is present in the VVC software. The document reports that the subjective merit of these modifications has been demonstrated via formal comparative subjective evaluation. The contribution suggests that the proposed improvements to the QPA design is adopted into in the next version of the VTM/BMS software.
* JVET-K0390: Rate Control for VVC by Tencent and Wuhan University
  + The HM software contains a rate control algorithm known as the R-lambda rate control algorithm. This contribution presents four modifications to the R-lambda algorithm. The contribution reports BD-rate difference of -9.03% and -2.94% for RA and LD for VTM-1.1 when using VTM-1.1 with a straightforward implementation of the HM R-lambda algorithm as anchor. It can be noted that no corresponding input contribution to JCT-VC was submitted.

The AHG recommended that the related input contributions are reviewed and to continue to study Encoding algorithm optimizations in JVET.

A participant remarked that some encoding optimization work had also been done in CE11 on optimizations for long-term referencing. [Check new version for content added about this.]

[JVET-K0011](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3964) JVET AHG report: Screen Content Coding (AHG11) [S. Liu, J. Boyce, Y.-C. Sun]

This document summarizes the activity of AHG11: Screen Content Coding between the 10th meeting in San Diego, US (10 – 20 April 2018) and the 11th meeting in Ljubljana, SI (10 – 18 July 2018).

There was little or no email exchange on the main reflector about the work of this AHG.

The first table below shows the test sequences in current CTC class F which have been used by the community for a number of years. However, it was advised by several groups of experts that these sequences may be outdated such as in resolution, frame rate, or the contents themselves. One suggestion was to replace class F by the SCC 420 TGM class sequences which were used for developing HEVC SCC. These sequences are shown in the next table. There was also some interest in adopting one or two eSports sequences to represent a type of content which is widely distributed nowadays. Some eSports sequences are proposed in JVET-K0294 (refined versions of JVET-J0052.)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Category | Resolution | Sequence name | fps | Bit depth | Frames to be encoded |
| Class F | 832x480  1024x768  1280x720  1280x720 | BasketballDrillText  ChinaSpeed  SlideEditing  SlideShow | 50  30  30  20 | 8  8  8  8 | 0-499  0-499  0-299  0-499 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Category | Resolution | Sequence name | fps | Bit depth | Frames to be encoded |
| 4:2:0 TGM | 1920x1080  1920x1080  1920x1080  1920x1080 | FlyingGraphics\*  Desktop  Console  ChineseEditing | 60  60  60  60 | 8  8  8  8 | 0-299\*  0-599  0-599  0-599 |

On top of updating Class F testing sequences, some experts suggested to make Class F mandatory instead of the current optional test.

Input documents related to AHG11 are summarized as follows.

* JVET-K0048 “CE8: Intra Region-based Template Matching (Test 8.1)”, G. Venugopal, K. Müller, H. Schwarz, D. Marpe, T. Wiegand (HHI).
* JVET-K0050 “CE8 related: Intra Region-based Template Matching for luma and chroma”, G. Venugopal, K. Müller, H. Schwarz, D. Marpe, T. Wiegand (HHI).
* JVET-K0075 “CE8-2.1: Current picture referencing using block level flag signalling”, X. Xu, X. Li, S. Liu (Tencent).
* JVET-K0075 “CE8-2.2: Current picture referencing using reference index signalling”, X. Xu, X. Li, S. Liu (Tencent).
* JVET-K0411 “AHG11: Palette mode”, Y.-C. Sun, J. An, J. Lou (Alibaba).
* JVET-K0294 “Tencent test sequences, and Class F test set restructure”, J. Ye, S. Wenger, et al. (Tencent).

It was also reported that line-based intra prediction (JVET-J0014, JVET-K0049) may help improve screen content coding efficiency.

The AHG recommended:

* To review all related contributions
* To evaluate new test materials
* To consider updating the Class F sequences
* To discuss SCC test conditions

A participant commented that metrics for screen content coding would be helpful. It was also commented that sometimes the most benefit for SCC is seen with test sequences that can already be coded with very low bit rate. Some gain was reportedly obtained with gaming content. Generally, more work on SCC was encouraged.

It was noted that 4:4:4 becomes desirable for screen content.

[JVET-K0012](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3953) JVET AHG report: High-level parallelism [T. Ikai, M. Coban, H. M. Jang, R. Skupin, Y.-K. Wang]

This document summarizes activities of the AHG on high-level parallelism between the 10th and the 11th JVET meetings.

There was no email discussions in the main reflector

Three contributions, listed below, are on tiles and related to this AHG. All the three contributions contain aspects related to the first mandate. Some of the three contributions also contain aspects related to the third mandate.

The three contributions are:

JVET-K0155 AHG12: Flexible Tile Partitioning Y. Yasugi, T. Ikai (Sharp)

JVET-K0260 Flexible Tiles R. Sjöberg, M. Damghanian, M. Pettersson, J. Enhorn (Ericsson)

JVET-K0300 m43233 Design goals for tiles M. M. Hannuksela, A. Zare, M. Homayouni, R. Ghaznavi-Youvalari, A. Aminlou (Nokia)

Below is a summary of the three contributions:

* JVET-K0155 proposes flexible size tile, which enables more flexible tiling segmentation than in HEVC, such that the tile size is in units of an explicitly signalled tile size unit, which can be other than in the units of CTUs, and provides experimental results for conventional and proposed tile in BMS (VTS config).
* JVET-K0260 proposes flexible tile segmentation that allows
  + rectangular tiles of varying sizes (wherein the number of tiles and their sizes are explicitly signalled); and
  + efficient tile structure signalling with tile sizes in subtile units and “use a previous tile size” flag.
* JVET-K0300 proposes some design goals to be used in evaluating merits of technical contributions on tiles and to be included as mandates of an appropriate JVET AHG covering tiling aspects. The proposed design goals are to make VVC tiles more suited for viewport-dependent 360° video streaming, including:
  + Encoding of motion-constrained tile sets (MCTSs) that are more efficient than HEVC MCTSs in terms of rate-distortion penalty;
  + Avoiding visible MCTS boundaries with as small processing cost as possible;
  + Intra block copy across tiles for enabling prediction from one constituent frame to another for frame-packed stereoscopic video, provided that intra block copy is adopted as a tool in VVC;
  + Extracting VCL NAL units of a subset of MCTSs from one VVC bitstream and reposition them to another VVC bitstream without VCL NAL unit modifications.

Recommendations

* Review the related contributions.
* Continue to study high-level parallelism techniques.

[JVET-K0013](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3911) JVET AHG report: Tool reporting procedure (AHG13) [W.-J. Chen, J. Boyce, E. Alshina, J. Chen, E. Francois, Y. He, Y. W. Huang]

This document summarizes the activity of AHG13: “Tool reporting procedure” between the 10th Meeting in San Diego, US (10–20 Apr 2018) and the 11th meeting in Ljubljana, SI (10–18 July 2018). Tool on/off experimental results vs. VTM and BMS anchors are provided for the tools specified in JVET-J1005, which include the BMS tools and the HEVC tools not included in the VTM.

The initial version of JVET-J1005 “Methodology and reporting template for tool testing” was provided on April 25, with updates provided on May 8 and May 29. The document contained a reporting template.

All tests described in JVET-J1005 were conducted. The tested tools, testers, and cross-checkers are listed in the tables below.

List of tools included in BMS but not included in VTM

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Tool Name** | **Abbrev. Name** | **Document reference(s)** | **VTM anchor, tool on/off** | **BMS anchor, tool on/off** | **AI** | **RA** | **LD** | **Tester** | **Crosscheck** |
| 65 intra prediction mode | 65IPM | JVET-C0055 | on | off | X | X |  | W.-J. Chien (Qualcomm) | K. Choi (Samsung) |
| AMT+4x4 NSST | TRM | JVET-D0120 | on | off | X | X |  | T. Tsukuba (Sony) | K. Choi (Samsung) |
| Affine motion | AFF | ITU-T SG16 Doc. COM16–C1016 | on | off |  | X | X | H. Yang (Huawei) | Y. He (InterDigital) |
| Adaptive loop filter | ALF | JVET-E0104 | on | off | X | X |  | W.-J. Chien (Qualcomm) | R. Chernyak (Huawei) |
| Adaptive motion vector precision | AMVR | JVET-E0076 | on | off |  | X | X | W.-J. Chien (Qualcomm) | R. Chernyak (Huawei) |
| Coefficient coding | CFC | ITU-T SG16 Doc. COM16-C806 | on | off | X | X |  | W.-J. Chien (Qualcomm) | R. Chernyak (Huawei) |
| Decoder motion vector refinement | DMVR | JVET-E0052 | on | off |  | X | X | S. Esenlik (Huawei) | Y. He  (InterDigital) |
| LM Chroma mode | LMC | JVET-E0077 | on | off | X | X |  | W.-J. Chien (Qualcomm) | R. Chernyak (Huawei) |
| Subblock merge candidate (ATMVP) | ATMVP | JVET-C0035 | on | off |  | X | X | W.-J. Chien (Qualcomm) | R. Chernyak (Huawei) |

Table 2 List of tools included in HEVC but not included in VTM or BMS

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Tool Name** | **Abbrev. Name** | **Document reference(s)** | **VTM anchor, tool on/off** | **BMS anchor, tool on/off** | **AI** | **RA** | **LD** | **Tester** | **Crosscheck** |
| Strong intra smoothing | SIS | JCTVC-K0139 | on | on | X | X |  | F. Racape (Technicolor) | K. Choi (Samsung) |
| Boundary smoothing | BDS | JCTVC-G0457 | on | on | X | X |  | T.-D. Chuang  (MediaTek) | K. Choi (Samsung) |
| DST-VII for 4x4 intra block | DST | JCTVC-E125 | on | on | X | X |  | Y. He  (InterDigital) | K. Choi (Samsung) |
| Mode dependent coefficient scan | MDCS | JCTVC-G0232 | on | on | X | X |  | F. Racape (Technicolor) | K. Choi (Samsung) |
| Sign data hiding | SDH | JCTVC-H0481 | on | on | X | X |  | T. Tsukuba (Sony) | K. Choi (Samsung)/T.-D. Chuang  (MediaTek) /Y. Peng (Dolby) |

The results of the tests are included in the AHG report. The attached spreadsheet provides additional data. Scatter plots 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-1.0/

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. The largest runtime differences were found for TRM (AMT+4x4 NSST), where the tester uses GCC 6.3.0 and SIMD=SSE42and the crosschecker uses GCC 4.8.3 and SIMD=AVX.

Experiment test results were provided.

Sign data hiding had the largest gain among the things we disabled from HEVC %0.7.

The AHG recommends the following:

* Consider the reported tool test results during tool adoption decision making
* Review related contributions
* Refine list of tested tools and test methodology for the next meeting cycle
  + Consider the reported tool test results as a benchmark for CE tests
  + Consider including reporting of compute system information for testers and cross-checkers
  + Consider additional performance or complexity metrics

Among the BMS tools, 65 direction intra prediction modes, LM chroma, DMVR, and ALF had significant gains. Various techniques seemed to fall along the same line of complexity versus performance.

K0410 discusses the reliability of runtimes - e.g., on different encoders.

A participant remarked that running the decoder more than once was suggested as a way to make the runtimes more stable.

K0312 discusses the percentage of the video for which a tool is used.

It was remarked that the VTM has 4x4 biprediction and 2x2 chroma (intra and bipred) prediction, which are difficult, and it was suggested to measure the impact of these (which would require a software change).

We need to define what should be done in an AHG versus a CE.

# Project development (12)

Contributions in this category were discussed XXday XX July XXXX–XXXX (chaired by XXX).

## Text and general standard development (2)

[JVET-K0263](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3773) VVC document structure [[T. Amara (Amazon)](mailto:Amatarek@amazon.com), [D. Singer (Apple)](mailto:singer@apple.com), [A. Duenas (ARM)](mailto:Alberto.Duenas@arm.com), [G. Martin-Cocher (BlackBerry)](mailto:gmartincocher@blackberry.com), [A. Hinds (CableLabs)](mailto:A.Hinds@cablelabs.com), [T. Davies (Cisco)](mailto:thdavies@cisco.com), [P. Pahalawatta (DirectTV)](mailto:pp2960@att.com), [J. Samuelsson (Divideon)](mailto:jonatan.samuelsson@divideon.com), [X. Ducloux (Harmonic)](mailto:Xavier.Ducloux@harmonicinc.com), [J. Chen (Huawei)](mailto:jianle.chen@huawei.com), [J. Boyce (Intel)](mailto:jill.boyce@intel.com), [A. Norkin (Netflix)](mailto:anorkin@netflix.com), [G. Teniou (Orange)](mailto:gilles.teniou@orange.com), [M. Karczewicz (Qualcomm)](mailto:martak@qti.qualcomm.com), [J. Song (Samsung)](mailto:jy_song@samsung.com), [T. Suzuki (Sony)](mailto:teruhiko.s@sony.com), [D. Gibellino (Telecom Italia)](mailto:diego.gibellino@telecomitalia.it), [D. Nicholson (Vitec)](mailto:didier.nicholson@vitec.com)]

[JVET-K0311](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3825) Interoperability point signalling for VVC [J. Boyce (Intel), X. Ducloux (Harmonic), A. Hinds (CableLabs), S. Wenger (Tencent)]

## Software development and CTC (10)

Allocated to BoG (coord by F. Bossen)

[JVET-K0054](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3555) Unification of PSNR calculation for JVET CTC [R. Chernyak, T. Solovyev, S. Ikonin, J. Chen (Huawei)]

[JVET-K0149](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3656) Reference software extension for coding block statistics [J. Sauer, J. Schneider, M. Bläser, M. Wien (RWTH Aachen)]

[JVET-K0154](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3662) On encoding distortion evaluation of VTM/BMS software [T. Chujoh, T. Ikai (Sharp)]

[JVET-K0496](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4026) Crosscheck report of JVET-K0154 on encoding distortion evaluation [S. Iwamura, S. Nemoto, A. Ichigaya (NHK)] [late]

[JVET-K0433](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3956) Crosscheck of JVET-K0154 on On encoding distortion evaluation of VTM/BMS software [X. Li (Tencent)] [late]

[JVET-K0261](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3771) AHG3: VVC Software Cleanup [A. Wieckowski, K. Sühring, H. Schwarz]

[JVET-K0312](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3826) AHG13: Reporting of adjusted decoder runtimes in tool on/off tests [J. Boyce, D. Gurulev, V. Aristarkhov (Intel), A. Tourapis (Apple)]

[JVET-K0389](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3908) AHG3: Proposed software management for VTM [E. Thomas, A. Gabriel (TNO)] [late]

[JVET-K0410](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3930) Comments on timing measurement variations for JVET experiments [F. Glapin, T. Poirier, F. Le Léannec, E. François (Technicolor)] [late]

[JVET-K0461](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3990) AHG3: Guidelines for VVC reference software development [F. Bossen, X. Li, K. Sühring (AHG chairs)] [late]

## Coding studies (1)

[JVET-K0445](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3970) Compression performance report of VTM/BMS for 8K test sequences [S. Iwamura, S. Nemoto, A. Ichigaya (NHK)]

This contribution reports on compression performance of BMS-1.1 associated with 8K test sequences. The clips span a wide range of spatial temporal characteristics and portray content typical of broadcasting programs. Three ITU-R BT.2020 SDR test sequences have been encoded with HEVC test model (HM-16.16) and VVC test model (BMS-1.1 with VTM/BMS configurations). The simulation results show similar tendency of bit rate saving and encoding/decoding complexity for 8K test sequences to those for CTC sequences. It is confirmed that current BMS-1.1 software works for 8K sequences as expected.

Presented in track B Tuesday 17th 10:00

The 8K test sequences used for this test are parts of the test materials “Ultra-high definition/wide-color-gamut standard test sequences – Series B” distributed by Japanese academic society, the Institute of Image Information and Television Engineers (ITE). NOTE that these 8K test sequences are only provided to the licensee and not allowed to redistribute. They can therefore not be used by JVET for standard development.

Results with RA configuration (only 33 frames, i.e. 2 IRAPs were encoded at beginning and end, which approximately corresponds to ¼ s RA period. Gain compared to HM is somewhat less than what we currently see for 4K, which may however be due to the fact that more intra coding is used.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **BMS-1.1 with VTM config over HM-16.16** | | | | |
|  | Y | U | V | EncT | DecT |
| b01\_WaterPolo-Goal\_8K\_from00380\_to00412 | -6.61% | -9.38% | -21.89% | 380% | 100% |
| b05\_HorseRace-Finish\_8K\_from00252\_to00284 | -9.22% | -23.36% | -31.60% | 140% | 88% |
| b09\_Marathon-PanningDown\_8K\_from00764\_to00796 | -5.26% | -9.88% | -13.22% | 266% | 104% |
| All | -7.03% | -14.21% | -22.23% | 242% | 97% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **BMS-1.1 with BMS config over HM-16.16** | | | | |
|  | Y | U | V | EncT | DecT |
| b01\_WaterPolo-Goal\_8K\_from00380\_to00412 | -18.75% | -24.00% | -93.16% | 1906% | 223% |
| b05\_HorseRace-Finish\_8K\_from00252\_to00284 | -29.82% | -57.83% | -105.45% | 821% | 258% |
| b09\_Marathon-PanningDown\_8K\_from00764\_to00796 | -23.32% | -22.57% | -44.60% | 1273% | 235% |
| All | -23.97% | -34.80% | -81.07% | 1258% | 238% |

Interesting information – contributors annouce that they will conduct this test from time to time to report how future versions VTM/BMS perform on 8K material.

# Test material (2)

Contributions in this category were discussed XXday XX July XXXX–XXXX (chaired by XXX).

[JVET-K0294](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3807) Tencent test sequences, and Class F test set restructure [J. Ye, S. Wenger, X. Li, S. Liu, L. Wu, C. Xie, K. Liu, B. Wang, P. Liu, K. Dong, Y. Kuang, W. Feng (Tencent)]

Discussed in the CTC BoG.

[JVET-K0409](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3929) AHG4: On copyright of HLG test sequences [T. Suzuki (Sony), A. Ichigaya (NHK)] [late]

Discussed in the context of the AHG4 report.

# Core Experiments

## CE1: Partitioning (32)

Contributions in this category were discussed Tuesday 10 July 1720–2020 and Wednesday 11 July 900-1100 (chaired by JRO).

[JVET-K0021](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3618) CE1: Summary report on partitioning [J. Ma, F. Le Léannec, M. W. Park]

This was discussed Tuesday 10 July 1710 (JRO)

This document evaluates CE1: Partitioning (JVET-J1021). In total there are 59 tests each for VTM and BMS that have been cross-checked by at least one cross-checker. 5 additional tests have been withdrawn from the original CE description. Out of the 59 tests that have been cross-checked, the cross-checkers reported in some cases mismatches in timings, otherwise there were no significant mismatches in BD-rates. Some reported small mismatches around 3 or 4 digits after decimal point were noted and were most likely due to parallel processing.

**SubCE1: Partitioning structure (replace J numbers by K numbers)**

The experiments conducted in this SubCE can be categorized as follows. Different proposed partitioning methods are tested in

* SubCE1.0.1-1.0.4 (different configurations)
* SubCE1.0.5-1.0.9 (different configurations)
* SubCE 1.0.12-1.0.15 (different configurations)

Different context models for VTM are tested in

* SubCE1.0.11
* SubCE1.0.16
* SubCE1.0.17
* SubCE1.0.19.

See also JVET-K0220.

ABT (JVET-J022, JVET-J0075)

ABT is an alternative partitioner to QTBT+TT proposed by Technicolor. The ABT partitioner allows additional split modes 1/4 and 3/4 in addition to the quad split and the 1/2 split both from QTBT. Further, 1/3 and 2/3 split modes are allowed in general. If not further specified below, the following sizes were used

* Minimum and maximum CTU size: 4 and 128
* Minimum and maximum TU size: 4 and 64
* Minimum and maximum PU size: 4 and 128
* Additional transform types and sizes: 12, 24, 48

Different split possibilities and transform sizes are tested as follows:

* SubCE1.0.1: CU Sizes 12, 24, and 48 (luma) are allowed, as well as transforms for these sizes. The additional split ratios 1/3 and 2/3 are disabled.
* SubCE1.0.2: CU Sizes 12 and 24 (luma) are allowed, as well as transforms for these sizes. The additional split ratios 1/3 and 2/3 are disabled.
* SubCE1.0.3: CU Sizes 12 and 24 (luma) are allowed, as well as transforms for these sizes. The additional split ratios 1/3 and 2/3 are enabled.
* SubCE1.0.4: CU Sizes 12, 24 and 48 (luma) are allowed. Transforms with sizes not equal to power 2 are disabled in this test. CUs with non-power of 2 sizes are either coded with null residual or divided into power of 2 transform blocks.

QTBTS (JVET-J0035)

QTBTS is an alternative partitioner to QTBT+TT proposed by Fraunhofer HHI. The QTBTS partitioner allows additional split modes 1/4, 3/4, 3/8, and 5/8 in addition tot he quad split and the 1/2 split both from QTBT. Further, 1/3, 2/3, 1/5, 2/5, 3/5, and 4/5 split modes are allowed in general. If not further specified below, the following sizes were used:

* Minimum and maximum CTU size: 4 and 128
* Minimum and maximum TU size: 4 and 64
* Minimum and maximum PU size: 4 and 128
* Additional transform types and sizes: 6, 10, 12, 20, 24, 40, 48
* Other structural differences to VTM software: The affine motion vectors were moved from the motion buffer to the prediction units.

Different split possibilities and transform sizes are tested as follows:

* SubCE1.0.5: 1/4, 3/4, 1/3, and 2/3 split modes are allowed. CU Sizes 12, 24, and 48 (luma) are allowed, as well as transforms for these sizes.
* SubCE1.0.6: 1/4 is allowed but 2/3 must follow. Leading to a final partitioning structure that consists of blocks with power of 2 sizes.
* SubCE1.0.7: 1/4, 3/4, 3/8, 5/8, 1/3, 2/3, 1/5, 2/5, 3/5, and 4/5 split modes are allowed. CU Sizes 12, 24, 20 and 48 (luma) are allowed, as well as transforms for these sizes.
* SubCE1.0.8: 1/4 is allowed but 2/3 must follow. 5/8 is allowed but either 2/5 or 4/5 must follow. Leading to a final partitioning structure that consists of blocks with power of 2 sizes.

Context modeling (JVET-J0021)

* SubCE1.0.9:
  + Minimum and maximum CTU size: 4 and 128
  + Minimum and maximum TU size: 4 and 64
  + Minimum and maximum PU size: 4 and 128
  + Additional transform types and sizes: None.
  + Other structural differences to VTM software: None.
  + SubCE1.0.9 uses the QTBT+TT partitioner provided by VTM but uses a different context model for the split decision tree.

Signalling (JVET-J0026)

* SubCE1.0.11:
  + Minimum and maximum CTU size: 4 and 128
  + Minimum and maximum TU size: 4 and 64
  + Minimum and maximum PU size: 4 and 128
  + Additional transform types and sizes: None.
  + Other structural differences to VTM software: None.
  + In SubCE1.0.11 proponents removed bins for BT offset.

(QT)BTT (JVET-J0024)

* If not further specified below, the following sizes were used:
  + Minimum and maximum CTU size: 4 and 128
  + Minimum and maximum TU size: 4 and 64
  + Minimum and maximum PU size: 4 and 128
  + Additional transform types and sizes: None.
* (QT)BTT is a recursive partitioner that uses binary and ternary splits on top of a quad split at CTU level. The split availability is controlled by CU ratios and sizes.
* SubCE1.0.12 (Configuration: (QT)BTT): The following CU ratios are allowed: 1:1, 1:2, 1:4. The max CU size for 1:4 and TT splits is 64. Further the maximum binary tree depth is 8, and the maximum ternary tree depth is 6.
* SubCE1.0.13 (Configuration: (QT)BTT):
  + The following CU ratios are allowed: 1:1, 1:2, 1:4. The max CU size for 1:4 and TT splits is 64.
  + Further the maximum binary tree depth is 10, and the maximum ternary tree depth is 8.
* SubCE1.0.14 (Configuration: QTBT+TT): The following CU ratios are allowed: 1:1, 1:2, 1:4, 1:8.
* SubCE1.0.15 (Configuration: QTBT+TT):
  + The following CU ratios are allowed: 1:1, 1:2, 1:4, 1:8. The max CU size for 1:4 and TT splits is 64.
  + Context modeling (JVET-J0017)
* SubCE1.0.16:
  + Minimum and maximum CTU size: 4 and 128
  + Minimum and maximum TU size: 4 and 64
  + Minimum and maximum PU size: 4 and 128
  + Additional transform types and sizes: None.
  + Other structural differences to VTM software: None.
  + SubCE1.0.16 uses a different context model for the split decision tree.

Context modeling (JVET-J0024)

* SubCE1.0.17:
  + Minimum and maximum CTU size: 4 and 128
  + Minimum and maximum TU size: 4 and 64
  + Minimum and maximum PU size: 4 and 128
  + Additional transform types and sizes: None.
  + Other structural differences to VTM software: None.
  + SubCE1.0.17 uses a different context model for the split decision tree.

Context modeling (JVET-J0018)

* SubCE1.0.19:
  + Minimum and maximum CTU size: 4 and 128
  + Minimum and maximum TU size: 4 and 64
  + Minimum and maximum PU size: 4 and 128
  + Additional transform types and sizes: None.
  + Other structural differences to VTM software: None.
* SubCE1.0.9 uses a different context model for the split decision tree.

CTC overall results relative to the VTM are



CTC overall results relative to the VTM are



Sub-CE1 notes from discussion:

1. Alternative partitioning structures:   
   - Asymmetric binary tree (ABT) also including encoder opt in some conf (Sub-CE8), which typically comes with heavy reduction of encoder runtime. This raises the general question what would be possible with QT/TT/BT if a more complex encoder would be used.   
   - QTBTS in different configurations, including case 1.0.6/1.0.8 which restrict to dyadic transform sizes and by two subsequent splits can mimic TT; also reduces encoder runtime, but has increase in decoder runtime  
   - QTBTT imposes some restrictions to the syntax, and gives gain for AI (at the expense of 1.5x encoder run time), and

The main questions remaining were:

- How much of the gain comes from encoder optimization, and how much from syntax changes and additional tools?

- If gain comes from changes in the decoder, what is the impact on formulating it, and also implement e.g. restrictions in the decoding process

This was further investigated. In particular, it was noted that for fair comparison against the VTM an encoder with similar complexity should be used. A BoG (B. Bross) was requested to discuss options of further VTM encoder runtime and performance optimization, and asked to suggest solutions how the questions above could be answered, taking examples and experiences from Sub-CE1. See notes for the BoG report K0528 and section 7.1.

Compared to BMS, somewhat less but similar gain is observed.

1. Context modelling (only for QT/TT/BT)
2. Results in this aspect are partially targeting complexity reduction (e.g. reducing number of contexts), or improvement of compression (typically small, <0.1%). These proposals rather fall into category of fine-tuning, which may be premature at the current status where further modifications of partitioning are still an option for the future.

Note: As a side-remark, Excel sheets documenting the cross-checks should be uploaded with the CE summaries

**SubCE2: Picture boundary handling (replace J numbers by K numbers)**

The experiments conducted in this SubCE are all on picture boundary handling methods applied to different partitioning methods that were tested in SubCE2.

*Summary of proposals*

If not further specified below, the following sizes were used.

* Minimum and maximum CTU size: 4 and 128
* Minimum and maximum TU size: 4 and 64
* Minimum and maximum PU size: 4 and 128
* Additional transform types and sizes: None.

The different picture boundary handling methods are tested as follows:

*SubCE 2.0.1 (JVET-J0022, Configuration: QTBT+TT):*

If a CU partially lies outside the coded picture, then the symmetric BT split (i.e split ratio ½) is always allowed for current CU, in the orientation of the concerned picture border. If the bottom-right corner of current CU is outside the picture, then only QT split is allowed. Moreover the TT split mode in concerned orientiation is also allowed, if one of the 2 split boundaries issued from TT split perfectly matches the picture border. A rate distortion choice between all allowed split modes is perormed. The non-inferred split mode signalling is coded and parsed.

*SubCE 2.0.2 (JVET-J0022, Configuration: QTBT+ABT)*

Same process as in SubCE2.0.1, where the ABT replaces TT in the proposed policy. Hence, when a block partially lies outside the picture and the ABT split (1/3,3/4) or (3/4,1/4) in considered direction provides a SubCU that perfectly matched the picture border, then it becomes a split mode candidate in the rate distortion choice of current block partitioning.

Additional transform types and sizes: 6, 12, 24, 48

*SubCE 2.0.3 (JVET-J0035, Configuration : QTBT)*

This approach makes a rate distortion optimal choice between the QT split and the symmetric ½ binary split in the direction parallel to the concerned picture border. A bin signals the use of quad split or the BT split. If the QT split is not allowed because the current CU has a binary tree depth higher than zero, then the only allowed split is the symmetric binary split. Since it is inferred, it is called the implicit BT split mode in that case.

*SubCE 2.0.4 (JVET-J0035, Configuration: QTBT)*

Same process as in VTM, only QT is allowed at frame boundaries.

*SubCE 2.0.5 (JVET-J0035, Configuration QTBT+TT)*Same process as in SubCE2.0.3, but with QTBT + TT configuration.  
This approach makes a rate distortion optimal choice between the QT split and the symmetric ½ binary split in the direction parallel to the concerned picture border. A bin signals the use of quad split or the BT split. If the QT split is not allowed because the current CU has a binary tree depth higher than zero, then the only allowed split is the symmetric binary split. Since it is inferred, it is called the implicit BT split mode in that case.

*SubCE 2.0.6 (JVET-J0035, Configuration QTBT+TT)*This is VTM-1.0, QTBT + TT configuration, only QT allowed at frame boundaries.

*SubCE 2.0.7 (JVET-J0035, Configuration QTBTS)*  
Same process as in SubCE2.0.3 and SubCE2.0.5, but with QTBT + BTS configuration.  
This approach makes a rate distortion optimal choice between the QT split and the symmetric ½ binary split in the direction parallel to the concerned picture border. A bin signals the use of quad split or the BT split. If the QT split is not allowed because the current CU has a binary tree depth higher than zero, then the only allowed split is the symmetric binary split. Since it is inferred, it is called the implicit BT split mode in that case.

*SubCE 2.0.8 (JVET-J0035, Configuration: QTBTS)*  
Same process as in VTM, only QT is allowed at frame boundaries.

*SubCE 2.0.9 (JVET-J0032, Configuration: QTBT+TT)*

This method consists in allowing BT split for boundary CTU, and adapting the maximum allowed QT depth and BT depth, for concerned CUs. A ForceSplitLevel and ForceQTSplitLevel are deptermine to ensure that CUs resulting from BT split within the coded picture. Next an BT depth extension is used to provide a certain level of freedom in the BT/TT partitioning of CUs near the picture border.

*SubCE 2.0.10 (JVET-J0018, Configuration: QTBT+TT)*

This method allows QT and BT/TT split for CU across the picture border. Force CU overlapping the bottom-right picture corner, QT split is inferred. If the maximum QT depth is reached or current BT/TT depth is higher than 0, then symmetric BT split is inferred. No implicit BT depth. No RDO between QT and BT if BT split is inferred.

*SubCE 2.0.11 (JVET-J0024, Configuration: QTBT+TT)*

QT or horizontal BT for bottom boundary, QT or vertical BT for right boundary, for bottom-right corner, only QT is used. First N level of QT is performed, the level N is adaptive selected.

*SubCE 2.0.12 (JVET-J0024, Configuration: QTBT+TT)*QT inferred for bottom-right boundary, BT for bottom or right boundary, no RDO between QT and BT at frame boundary.

*SubCE 2.0.14 (JVET-J0026, Configuration: QTBT+TT)*   
An implicit recursive partitioning is applied until all leaves of the tree are inside the picture. First quad-tree splitting is performed based on a minimum number of quad-tree splits for the CTU. Then implicit binary split is performed, by selecting the split that provided an edge closest to the picture border.   
Add high level syntax to indicate, if QT is inferred at frame boundary, if currDepth is below a given threshold. Compute a number of additional BT depth to reach the frame boundary, this number increment the maxBTDepth.

*SubCE 2.0.15 (JVET-J0026, Configuration QTBT+TT)*Same process as in SubCE2.0.14, but use implicit TT if it gets closest to the picture boundary.

**SubCE2 – BMS results. For full results see the attached Excel files.**

*CTC overall results compared to VTM*



*CTC overall results compared to BMS*



Sub-CE2 notes from discussion:

- some gain for RA/LB 0.3-0.5%

- It is noted that typically the gain is largest for HD

The current WD and VTM enforce an implicit QT split in boundary CTUs (as inherited by HEVC), which makes any subsequent TT/BT split inefficient, as it would need to start from small square blocks. It is obvious that this causes some loss in compression efficiency, and a solution should be sought that has a reasonable tradeoff between compression and imposing very specific dependencies of the syntax decoding from the position in the picture.

BoG (K. Misra) to further study the different proposals and suggest the most viable solution (including CE related proposals K0320, K0366). Revisit.

Sub-CE3: Split restrictions

- very small gain, neither bit rate nor enc/dec time

- some proposals impose more restrictions, other proposals remove restrictions that are currently in the VVC draft syntax (6 restrictions are there). Further, there are different restictrictions on intra and inter slices

- Agreed that at the current stage of development we should rather target simplifying, i.e. remove restrictions, and unify intra and inter cases.

Using multiple split modes for partitioning normally leads to partitioning redundancy. To avoid redundant partitioning several restriction methods are proposed. If not further specified below, the following sizes were used.

* Minimum and maximum CTU size: 4 and 128
* Minimum and maximum TU size: 4 and 64
* Minimum and maximum PU size: 4 and 128
* Additional transform types and sizes: None.

Different prevention methods are tested as follows: (replace J numbers by K numbers)

* *SubCE3.0.1 (JVET-J0021, JVET-J0022, Configuration: QTBT+TT):*  
  This test is based on QTBT+TT structure and the following restrictions are tested.  
  - Normative prevention of emulating BT splits with TT  
  - Restriction to reach a quaternary block division throught at most one succession of BT/TT splits  
  - unify intra and inter, such that restrictions of inter are also applied for intra, but keep all restrictions
* *SubCE3.0.2 (JVET-J0021, JVET-J0022, Configuration: QTBT+ABT):*  
  This test is based on QTBT+ABT structure and the following restrictions are tested.  
  - Normative prevention of emulating BT splits with ABT  
  - Restriction to reach a ternary block division throught at most one series of splits  
  - Restriction to reach a quaternary block division throught at most one succession of BT/ABT splits
* *SubCE3.0.3 (JVET-J0020, Configuration: QTBT+TT):*  
  This test is based on QTBT+TT structure and the following restrictions are tested.  
  - Normative prevention of emulating BT split with TT  
  - Restriction to reach very narrow block for TT
* *SubCE3.0.4 (JVET-J0017, Configuration: QTBT+TT):*  
  This test is based on QTBT+TT structure and the following restrictions are tested.  
  - Normative prevention of emulating QT splits with BT  
  - Normative prevention of redundant partitioning with BT
* *SubCE3.0.6 (JVET-J0024, Configuration: QTBT+TT):*  
  This test is based on QTBT+TT structure and the following restriction is tested.  
  - Normative prevention of redundant partitioning with BT
* *SubCE3.0.7 (Configuration: QTBT+TT):*  
  This test is based on QTBT+TT structure and the restrictions in VTM are tested. This test proposes encoder only prevention of redundant split partitioning in VTM and includes four sub-test cases as follows:  
  a) Remove QTBT+TT restriction in VTM at both encoder and decoder  
  b) Remove QTBT+TT restriction in VTM at decoder, restriction is used for encoder speed-up  
  c) Keep only the TT restriction (preventing binary split with same orientation in center partition of the ternary split) in VTM at both encoder and decoder  
  d) Keep only the TT restriction (preventing binary split with same orientation in center partition of the ternary split) in VTM at decoder, restriction is used for encoder speed-up

To control complexity of VTM several methods for changing or modifying configuration of the segmentation tree are proposed. If not further specified below, the following sizes were used.

* Minimum and maximum CTU size: 4 and 128
* Minimum and maximum TU size: 4 and 64
* Minimum and maximum PU size: 4 and 128
* Additional transform types and sizes: None.

Different configurations and restrictions are tested as follows:

* *SubCE3.0.8 (JVET-J0024, Configuration: QTBT+TT):*  
  This test is based on QTBT+TT structure. The split restriction on a node with an aspect ratio 1:N is tested. The node with a certain aspect ratio disallows split modes which results in child nodes of aspect ratio higher than the aspect ratio. This test provides two sub-test cases as follows:  
  a) Split restriction on a node the aspect ratio 1:4 (N=4)  
  b) Split restriction on a node the aspect ratio 1:8 (N=8)
* *SubCE3.0.9 (JVET-J0072, Configuration: QTBT+TT):*  
  This test is based on QTBT+TT structure. The restriction, which TT split is only allowed at leaf nodes, is tested.
* *SubCE3.0.10 (JVET-J0018, Configuration: QTBT+TT):*  
  This test is based on QTBT+TT structure. A threshold, which is signalled in SPS, controlling the maximum TT split depth is proposed. The maximum TT split depth is set to MaxBTTDepth – threshold. This test provides two sub-test cases as follows:  
  a) threshold=2 (i.e., the maximum TT split depth = MaxBTTDepth – 2)  
  b) threshold=1 (i.e., the maximum TT split depth = MaxBTTDepth – 1)

Comparison vs. VTM



Comparison vs. BMS



From the criteria above, the proposal in SubCE3.0.7c (from JVET-K0351) is the most appropriate solution. Decision (VTM/WD): Adopt JVET-K0351 (test c): Keep only the TT restriction (preventing binary split with same orientation in center partition of the ternary split) in VTM at both encoder and decoder

Sub-CE4: Split unit coding order

Split Unit Coding Order (SUCO) in JVET-J0024 enables a more flexible coding order, such as left to right (L2R) and right to left (R2L), to allow intra prediction from right reference pixels and inter prediction with right motion vector predictors. If a SU is partitioned vertically (vertical splitting), a flag is signalled to indicate L2R or R2L coding order of partition units. If a SU is partitioned by quad tree structure, a flag is shared for above two units and bottom two units. If no flag is signalled for coding order of an SU, the coding order follows its parent’s SU coding order. Due to more flexible coding order in SU level, the neighboring availability of a leaf CU become more diverse than common left and above neighbors in HEVC. There are four different availability cases if only left and right neighboring blocks are considered, i.e., LR\_10, LR\_01, LR\_11 and LR\_00. Above block is always available unless the current CU locates on the top boundary of a slice. Availability of the above left or above right corner block depends on the corresponding left or right neighbor availability. Intra prediction and most probable mode list is modified accordingly based on the left and right availability. Derivation for motion vector predictor in inter prediction is also modified accordingly based on the left and right availability.

Gain is 0.08% in RA, 0.23% in AI. The change in intra prediction and motion vector prediction is quite substantial. No action was taken on this, as a technology change like this should be justified by substantial compression gain.

Sub-CE5: Separate trees luma and chroma

If not further specified below, the following sizes were used.

* Minimum and maximum CTU size: 4 and 128
* Minimum and maximum TU size: 4 and 64
* Minimum and maximum PU size: 4 and 128
* Additional transform types and sizes: None.

The following tests were conducted in this subCE: (replace J numbers by K numbers)

* *SubCE5.1.1 (JVET-*J0018*, Configuration: QTBT+TT)*:  
  This test applies the separate trees for I slices that is implemented in VTM to QTBT+TT by enabling the config parameter DualITree.
* *SubCE5.1.2 (JVET-J0018, Configuration: QTBT+TT):*  
  In this test the separate trees are applied to QTBT+TT as in SubCE5.1.1 but with additional larger transforms for chroma (TU sizes 64) and multi-DM is also enabled.
* *SubCE5.2.1 (JVET-J0035, Configuration: QTBT+TT):*  
  Same as SubCE5.1.1. Both tests have equal results.
* *SubCE5.2.2 (JVET-J0035, Configuration: QTBTS):*  
  The separate trees are applied to QTBTS.
* *SubCE5.2.3(JVET-J0022, Configuration: ABT):*  
  The separate trees are applied to ABT.
* *SubCE5.2.4 (JVET-J0021, Configuration: QTBT+TT):*  
  Technology description is identical to SubCE5.1.1.
* *SubCE5.2.5 (JVET-J0021, Configuration: QTBT+TT):*  
  Technology description is identical to SubCE5.1.2 without larger chroma transforms.
* *SubCE5.3.1: (JVET-J0026, Configuration: QTBT):*  
  In this test the proponents use an adaptive switching between shared and separate trees between luma and chroma for intra-slices and inter slices. This technology is applied to the QTBT partitioner.
* *SubCE5.3.2: (JVET-J0026, Configuration: QTBT+TT):*In this test the proponents use an adaptive switching between shared and separate trees between luma and chroma for intra-slices and inter slices. This technology is applied to the QTBT+TT partitioner.
* *SubCE5.4.1: (JVET-J0026, Configuration: QTBT):*In this test the proponents use an adaptive switching between shared and separate trees between luma and chroma for inter slices. For intra-slices the separate trees are always used. This technology is applied to the QTBT partitioner.
* *SubCE5.4.2: (JVET-J0026, Configuration: QTBT+TT):*In this test the proponents use an adaptive switching between shared and separate trees between luma and chroma for inter slices. For intra-slices the separate trees are always used. This technology is applied to the QTBT+TT partitioner.

Results compared to VTM:



Results on top of BMS:



Additional results for 5.2.5 compared to BMS (not yet in v3 of summary report):

AI: -0.79/-9.76/-9.77% RA: -0.21/-5.64/-5.51%, encoder/decoder run times comparable to 5.1.1

Otherwise, this is identical to 5.1.1, but multi-DM is also enabled

- Separate partitioning for luma/chroma was already in JEM for I slices

- 5.1.x and 5.2.x apply only for I slices, 5.3.x and 5.4.x for intra and inter slices

- Benefit on top of VTM is small, benefit with BMS is larger (reason for this could be that it has some benefit in combination with LM chroma mode)

- Separate trees for luma and chroma have some impact on implementation, e.g. LM chroma mode would require some synchronization in the decoding of luma and chroma blocks of unequal size.

- It was initially planned to take no action with regard to draft text / VTM, but potentially consider including separate tree structures for luma and chroma into the BMS (if it still would exist). Candidates would be either 5.1.1 or 5.2.5. The outcome on this topic is recorded in section 7.1.

Sub-CE6: Large CTU handling

JVET-K0227: Two methods for processing CUs larger than the max transform size proposed in JVET-J0018 were tested as follows: (JVET-K0227)

* *SubCE6.1.1 (JVET-J0018, Configuration: QTBT+TT):*  
  This test is based on QTBT+TT structure, and proposes that when a CU with width or height greater than the maximum transform size (64) the CU is implicitly partitioned by quad tree split mode. The following sizes were used for this test.  
  - Minimum and maximum CTU size: 4 and 64 (VTM: 4 and 128)  
  - Minimum and maximum TU size: 4 and 64  
  - Minimum and maximum PU size: 4 and 64 (VTM: 4 and 128)  
  - Additional transform types and sizes: None.
* *SubCE6.1.2 (JVET-J0018, Configuration: QTBT+TT):*  
  This test is based on QTBT+TT structure, and proposes that when a CU with width or height greater than the maximum transform size (64), for inter slices the CU is coded with SKIP or AMVP modes, but not intra mode, and CBF values of the CU are inferred to be equal to zero, and for intra slices the CU is implicitly partitioned by quad tree split mode. The following sizes were used for this test.  
  - Minimum and maximum CTU size: 4 and 128  
  - Minimum and maximum TU size: 4 and 64  
  - Minimum and maximum PU size: 4 and 128  
  - Additional transform types and sizes: None.

JVET-K0120: A method for processing CUs larger than the max transform size proposed in JVET-J0028 was tested. The following sizes were used for this test.

* Minimum and maximum CTU size: 4 and 128
* Minimum and maximum TU size: 4 and 64
* Minimum and maximum PU size: 4 and 128
* Additional transform types and sizes: None.

One aspect was tested as follows:

* *SubCE6.2.1 (JVET-J0028, Configuration: QTBT+TT):*  
  This test is based on QTBT+TT structure, and proposes that when a CU with width or height greater than the maximum transform size (64), the CU is coded with zero residual or in skip mode and an inter prediction.

Results vs. VTM:



Results vs. BMS:



It is further noted that the loss of 6.1.2 is larger for class A1 than it is on average.

The results are not providing evidence that any of the methods gives significant advantage, and none of them would need some specific definition for large CTUs. No action.

Sub-CE7: Partitioning configuration parameters

The following test was conducted in this SubCE:

* *SubCE7.0.1 (JVET-J0018, Configuration: QTBT+TT):*

This proposal tests adaptive maximum CU sizes for BT and TT splits based on the statistics of the previously coded slices.

Results vs. VTM:



Results vs. BMS:



The proposal requires a syntax change (signalling of max CU size at slice header). Main advantage is the decrease in encoder run time, however it also slightly loses performance.

The current spec already has this syntax element, but it is only applied for CUs that are originating from BT split, and this syntax element is also used in the current VTM encoder.

Encoder-only option would be more desirable. Further study should be performed to check if the current syntax element is necessary at all (removing it and get the speedup just by omitting certain CU size checks from past statistics).

Sub-CE8: Encoder optimization for ABT- see above under Sub-CE1, Sub-CE8 was included in the table.

[JVET-K0528](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4058) BoG report on partitioning structures (CE1 SubCE1) [B. Bross]

Discussed Saturday 1430 (GJS)

The BoG met on July12, 2018 8:00pm-9:40pm to discuss options of encoder runtime and performance optimization of partitioning structures studied in SubCE1 of CE1.

The following partitioning schemes have been tested in SubCE1:

* QT+ABT - asymmetric binary tree - adding non-power-of-two transforms
* QT+BTS - “quadtree with binary tree and shifting of the binary split position” and a “closing” of the splits for power-of-two transforms
  + If only power-of-two transforms are used, the shifted binary splits are not really just binary splits - they become ternary splits or asymmetric 4-way splits.
* QT+BTT - The current VTM with some restrictions of CU ratio and CU size

No consensus was reached in the BoG. There was interest expressed to look at additional data. For the subsequent resolution of the topic, see section 7.1.

[JVET-K0559](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4089) Report of BoG on Picture Boundary Partitioning [K. Misra]

This BoG report was discussed Monday 16 July 1400 (chaired by GJS).

The BoG met on July 12, 2018, July 13, 2018 and July 15, 2018 to further study the different picture boundary partitioning proposals and suggest the most viable solution that has a reasonable tradeoff between compression and imposing very specific dependencies of the syntax decoding from the position in the picture.

The related contributions (including non-CE contributions) were reviewed and surveyed.

The BoG recommended:

* That CE1-2.0.5 (with a fix to ensure that the minQTSize constraint is obeyed) be included in VTM/BMS and VVC draft. Draft text was provided in JVET-K0554. (The fix does not affect the CTC.)
* Further study in a CE was recommended for JVET-K0366, JVET-K0320, CE1-2.0.10

minQTSize is the limit of the block size at which no further quadtree splits are allowed. It is set to 8 in the CTC, so the minimum QT leaf node size is 8x8 in the CTC, so there is no signal for whether to split an 8x8 block with a QT split – subsequent splits are binary splits. (It is still possible to use 3 binary splits to produce four 4x4 blocks.)

It was commented that although it was intended that the minQTSize not be violated, the BMS software that has been available can violate that at boundaries for pictures that are not multiples of 8 in width or height (although this case is not encountered in the CTC, since all CTC test sequences are multiples of 8 in size).

For the adopted approach, there is about 0.0%/0.3%/0.5% improvement for AI/RA/LB in the CTC. The effect is larger with large QPs.

The previous scheme just did implicit QT splits at the picture boundary, tending to result in an excessive use of small blocks. For HD, there is a 56x128 shape to contend with, which is an awkward shape to handle, and the proposed approach gives a more logical segmentation result that avoids a lot of small blocks at the bottom and right sides. Decision (BF/cleanup): Adopted as recommended by BoG (for both VTM and BMS).

[JVET-K0078](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3581) CE1: Partitioning signalling and split restriction (Test 1.0.16 and 3.0.4) [J. Nam, J. Lim, S. Kim (LGE)]

[JVET-K0109](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3613) CE1: Partitioning Structure in JVET-J0024 (Tests 1.0.12, 1.0.13, 1.0.14, and 1.0.15) [M. W. Park, M. Park, C. Kim (Samsung)]

[JVET-K0111](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3615) CE1: Picture Boundary Split in JVET-J0024 (Test 2.0.12) [M. W. Park, M. Park, C. Kim (Samsung)]

[JVET-K0120](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3626) CE1: Processing to support large CUs (Test 6.2.1) [K. Kondo, T. Suzuki (Sony)]

[JVET-K0133](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3639) CE1.4: Split Unit Coding Order [Y. Piao, J. Chen, X. Ouyang, A. Tamse, M. Park, C. Kim (Samsung)]

[JVET-K0134](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3640) CE1: Context modeling of MTT split modes (Test 1.0.17) [Y. Zhao, H. Yang, J. Chen (Huawei)]

[JVET-K0136](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3642) CE1: Redundant partition prevention with redundancy existence check (Test 3.0.6) [[Y. Zhao](mailto:yin.zhao@huawei.com), [H. Yang](mailto:haitao.yang@huawei.com), [J. Chen (Huawei)](mailto:jianle.chen@huawei.com)]

[JVET-K0137](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3643) CE1: On configuration of the MTT (Test 3.0.8 and Test 3.0.9) [Y. Zhao, H. Yang, J. Chen (Huawei)]

[JVET-K0150](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3657) CE1: Split restriction for narrow TT block (Test 3.0.3) [H. B. Teo, C. S. Lim (Panasonic)]

[JVET-K0197](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3706) CE1: Asymmetric Binary Tree (tests 1.0.1, 1.0.2, 1.0.3, 1.0.4, 8.0.1 and 8.0.2) [F. Le Léannec, T. Poirier, F. Galpin (Technicolor)]

[JVET-K0205](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3714) CE1: Dual Tree in I Slices in ABT configuration (test 5.2.3) [F. Le Léannec, T. Poirier, F. Galpin (Technicolor)]

[JVET-K0209](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3718) CE1: tests 2.0.1 and 2.0.2 [T. Poirier, F. Le Léannec (Technicolor)]

[JVET-K0210](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3719) CE1: tests 3.0.1 and 3.0.2 [T. Poirier, F. Le Léannec (Technicolor)]

[JVET-K0223](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3733) CE1.1.0.19: Context modeling for coding CU split decisions [S.-T. Hsiang, S.-M. Lei (MediaTek)]

[JVET-K0224](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3734) CE1.2.0.10: CU partitioning along picture boundaries [S.-T. Hsiang, S.-M. Lei (MediaTek)]

[JVET-K0225](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3735) CE1.3.5: Maximum TT split depth restriction [C.-M. Tsai, C.-W. Hsu, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-K0226](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3736) CE1.5.1: Separate coding tree partitioning for luma and chroma in I slices [S.-T. Hsiang, T.-D. Chuang, S.-M. Lei (MediaTek)]

[JVET-K0227](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3737) CE1.6.1: Coding large size CUs [S.-T. Hsiang, C.-Y. Chen, C.-Y. Lai, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-K0229](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3739) CE1.7.0.1: Signalling maximum CU size for BT/TT split [S.-T. Hsiang, S.-M. Lei (MediaTek)]

[JVET-K0280](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3793) CE1: QTBTS partitioning and boundary handling (1.0.5, 1.0.6, 1.0.7, 1.0.8, 2.0.7) [J. Ma, A. Wieckowski, H. Schwarz, D. Marpe, T. Wiegand (HHI)]

[JVET-K0287](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3800) CE1-2.0.11: Picture Boundary Handling [H. Gao, S. Esenlik, Z. Zhao, A. M. Kotra, J. Chen (Huawei)]

[JVET-K0418](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3939) Cross-check of JVET-K0287: CE1-2.0.11: Picture Boundary Handling [A. Wieckowski (HHI)] [late]

[JVET-K0314](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3828) CE1 Partition Signalling Context Reduction (Test 1.0.11) [J. Zhao, W. Zhu, K. Misra, A. Segall (Sharp)]

[JVET-K0316](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3830) CE1: Implicit QT,BT and MTT Partitions on Picture Boundary (Test 2.0.15) [W. Zhu, K. Misra, A. Segall (Sharp)]

[JVET-K0317](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3831) CE1: Implicit QT and BT Partitions on Picture Boundary (Test 2.0.14) [W. Zhu, K. Misra, A. Segall (Sharp)]

[JVET-K0326](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3840) CE1: Context modelling for QT/BT/TT decision tree (Test 1.0.9) [N. Hu, M. Karczewicz (Qualcomm)]

[JVET-K0340](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3857) CE1-5.2.4-5: Separate trees for luma and chroma in I slice (Test 5.2.4) and Multiple Direct Modes (MDM) method (Test 5.2.5) [A. K. Ramasubramonian, L. Pham Van, V. Seregin, W.-J. Chien, M. Karczewicz (Qualcomm)]

[JVET-K0351](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3868) CE1-3.0.7: removing prevention of redundant split partitioning [W.-J. Chien, V. Seregin, M. Karczewicz, N. Shlyakhov (Qualcomm)]

[JVET-K0353](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3871) CE1: Shared-separate partition tree in QT+BT configuration (Tests 5.3.1 and 5.4.1) [K. Misra, W. Zhu, A. Segall (Sharp)]

[JVET-K0354](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3872) CE1: Shared-separate partition tree in QT+BT+TT configuration (Tests 5.3.2 and 5.4.2) [K. Misra, W. Zhu, A. Segall (Sharp)]

[JVET-K0376](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3895) CE1-2.0.9: Picture Boundary Handling [Y. Li, D. Liu (USTC)]

## CE2: Loop filters (30)

Contributions in this category were discussed Wednesday 11 July in Track B 1130–1300 and 1430-1900 (chaired by JRO).

[JVET-K0022](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3609) CE2: Summary Report on In-Loop Filters [L. Zhang, K. Andersson, C.-Y. Chen]

This contribution provides a summary report of Core Experiment 2 on in-loop filters. Five categories, including 1) bilateral filter, 2) deblocking filters, 3) sample adaptive offset (SAO) filter, 4) Adaptive Loop Filters (ALF), and 5) non-local filters are covered by this CE.

Test conditions are specified for each category. The corresponding coding performance of each coding tool under evaluated in CE2 are summarized in this contribution. In addition, answers to questions mentioned in [1] and crosschecking results are also integrated in this contribution.

**CE2.1: Bilateral Filter**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| 2.1.1 | Same as BMS/JEM version |  |
| 2.1.2 | Bilateral filter - spatial filter strength adjustment | [JVET-K0231](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3741) |
| 2.1.3 | In-loop bilateral filter (also operated after block reconstruction,  i.e. affecting subsequent intra prediction) | [JVET-K0384](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3903) |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Test#** | **# of filter taps** | **Sample difference calculation** | **Parallel friendly (each sample can be filtered independently from other samples)** | **Table Size** |
| 2.1.1 | 5 | 1×1 | Y | 2778 bytes |
| 2.1.2 | 5 | 1×1 | Y | 2778 bytes |
| 2.1.3 | 5 | 1×1 for intra  3×3 for inter | Y | 30000 18bits |

Results vs. VTM:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **AI** | | | | | **RA** | | | | | **LDB** | | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 2.1.1 | -0.11% | -0.09% | -0.12% | 108% | 107% | -0.41% | -0.14% | -0.29% | 105% | 102% | -0.37% | 0.26% | 0.31% | 105% | 101% |
| 2.1.2 | -0.10% | -0.10% | -0.15% | 103% | 109% | -0.40% | -0.18% | -0.29% | 102% | 103% | -0.37% | 0.12% | 0.14% | 101% | 101% |
| 2.1.3 | -0.29% | -0.08% | -0.15% | 105% | 109% | -0.79% | -0.24% | -0.28% | 105% | 107% | -0.63% | 0.30% | 0.42% | 105% | 109% |

Results vs. BMS:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **AI** | | | | | **RA** | | | | | **LDB** | | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 2.1.1 | -0.37% | 0.26% | 0.31% | 105% | 104% | -0.41% | -0.38% | -0.47% | 107% | 102% | -0.51% | -0.12% | 0.08% | 102% | 102% |
| 2.1.2 | -0.26% | -0.39% | -0.39% | 103% | 104% | -0.44% | -0.45% | -0.52% | 104% | 100% | -0.46% | 0.06% | 0.06% | 103% | 101% |
| 2.1.3 | -0.33% | -0.25% | -0.24% | 105% | 106% | -0.54% | -0.27% | -0.28% | 107% | 103% | -0.62% | 0.20% | 0.38% | 106% | 104% |

The complexity impact should be further studied, in particular regarding

- the pipelining aspects with small intra prediction blocks

- the table size for solution 2.1.3

- further simplification by re-using difference computations

Some of these aspects are reported to be touched in CE related documents

Further study should be performed (continuation of CE).

**CE2.2: Deblocking filter**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| 2.2.1.1.a | Long deblocking filters and fixes (only luma) | [JVET-K0307](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3821) |
| 2.2.1.1.b | Long deblocking filters and fixes (version which only applies fixes for luma,  no long deblocking filter) | [JVET-K0307](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3821) |
| 2.2.1.2 | Extended Deblocking Filter (only luma) | [JVET-K0393](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3913) |
| 2.2.1.3 | Long deblocking filters (only luma) | [JVET-K0232](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3742) |
| 2.2.1.4 | Tests on long deblocking (only long for luma also filtering  chroma when long filters are used for luma) | [JVET-K0334](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3849) |
| 2.2.1.5 | Long-tap deblocking filter (only luma) | [JVET-K0112](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3616) |
| 2.2.1.6.a | Long-tap deblocking filter for luma | [JVET-K0152](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3660) |
| 2.2.1.6.b | Long-tap deblocking filter for chroma | [JVET-K0152](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3660) |
| 2.2.1.6.c | Long-tap deblocking filter for luma and chroma | [JVET-K0152](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3660) |
| 2.2.1.7 | Deblocking Improvements for Large CUs both luma and chroma | [JVET-K0315](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3829) |
| 2.2.2.1 | Deblocking filter with asymmetric weighting (weak filter modification) | [JVET-K0129](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3634) |
| 2.2.2.2 | Luma-adaptive deblocking filter (qp offset change based on luma level) | [JVET-K0386](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3905) |

An analysis of design aspects of the different proposals is included in the CE report (v3), but not fully agreed among participants. This should be resolved offline, insert tables from section 2.3 when confirmed

Performance vs. VTM (very similar for BMS)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **AI** | | | | | **RA** | | | | | **LDB** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 2.2.1.1.a | -0.20% | 0.00% | 0.00% | 104%  \*\* | 103% | -0.17% | -0.05% | -0.01% | 103%  \*\* | 103% | -0.05% | 0.04% | -0.15% | 99%  \*\* | 102% |
| 2.2.1.1.b | -0.19% | 0.00% | 0.00% | 103%  \*\* | 101% | -0.12% | -0.01% | 0.06% | 104%  \*\* | 101% | -0.11% | 0.12% | 0.02% | 102%  \*\* | 100% |
| 2.2.1.2 | 0.32% | -0.01% | -0.01% | 100% | 102% | 0.14% | 0.01% | 0.03% | 100% | 101% | 0.04% | 0.08% | -0.01% | 100% | 101% |
| 2.2.1.3 | 0.00% | 0.00% | 0.00% | 100% | 104% | -0.02% | -0.02% | -0.01% | 100% | 103% | 0.01% | 0.06% | 0.05% | 100% | 103% |
| 2.2.1.4 | 0.01% | 0.01% | -0.01% | 103%  \*\* | 105% | -0.09% | -0.39% | -0.44% | 105%  \*\* | 104% | 0.02% | -0.35% | -0.47% | 112%  \*\* | 106% |
| 2.2.1.5 | 0.03% | 0.00% | 0.00% | 98% | 100% | -0.02% | -0.03% | 0.03% | 99% | 105% | 0.10% | 0.12% | 0.17% | 100% | 98% |
| 2.2.1.6.a | -0.03% | 0.00% | 0.00% | 105%  \*\* | 102% | -0.08% | -0.03% | 0.04% | 95%  \*\* | 101% | 0.01% | -0.04% | 0.03% | 104%  \*\* | 101% |
| 2.2.1.6.b | 0.00% | -1.06% | -0.83% | 100%  \*\* | 102% | -0.04% | -1.93% | -1.88% | 102%  \*\* | 100% | -0.07% | -1.65% | -1.79% | 104%  \*\* | 101% |
| 2.2.1.6.c | -0.04% | -1.06% | -0.83% | 113%  \*\* | 104% | -0.10% | -1.95% | -1.89% | 115%  \*\* | 103% | -0.04% | -1.66% | -1.76% | 124%  \*\* | 103% |
| 2.2.1.7 | -0.01% | 0.33% | 0.33% | 105% | 106% | -0.08% | 0.22% | 0.26% | 105% | 107% | 0.02% | 0.20% | 0.20% | 104% | 90% |
| 2.2.2.1 | -0.11% | 0.00% | 0.01% | 99% | 99% | 0.08% | 0.02% | 0.05% | 100% | 100% | 0.13% | 0.15% | 0.16% | 100% | 101% |
| 2.2.2.2 | 0.00% | 0.00% | 0.00% | 100% | 101% | 0.01% | -0.04% | 0.01% | 100% | 100% | 0.03% | 0.02% | 0.03% | 100% | 101% |

For deblocking, subjective viewing is needed, PSNR does not provide evidence for case of deblocking.

Subjective viewing with QPs 32+37, compare to VTM

Candidate sequences:

UHD: Food market, Campfire, Tango

HD: Ritual Dance, Kristen and Sara

RA conf. for UHD, LD for HD sequences

From 2.2.1.1, a and b should be tested (b is BF only, no long filter)

From 2.2.1.6, only c should be tested

This was further discussed Saturday 14th 1715 after the viewing. A report was given as follows:

A decision was taken during the JVET meeting to perform a expert subjective assessment to evaluate the performances of the participants to the CE 2.2.

The Test Chair was asked to design a test trying to assess the Anchor (VTM 1.0) vs all the received submissions.

The test was performed with the participation of 15 JVET experts (5 more participated as informal viewers).

The CE experts asked to perform a visual assessment comparing the Anchor with each Submission, at the UHD resolution, using three test sequences coded at two QP rates.

A total of 9 submissions were considered and labelled with the P-codes from P10 to P19 (P18 was not considered, being not available the data), two QP were considered (QP32 and QP37) and three test sequences were encoded (Campfire, Market, Tango).

The test site was reasonably acceptable, being made of a room isolated from visual and audible external noise; light was dimmable from full 100 candles peak light down to a complete dark; no light was hitting the surface of the monitor.

The monitor was a mid-low consumer 55” TV set; all local post processing features were disables and light and brightness was put and the top values to allow a better vision of artefacts.

Five viewers were seated in front of the monitor at 2H and arranged inside a 60° angle from the screen centre. An analysis of the collected data showed no significant difference between including or excluding the two viewers seated on the external sides.

The A vs. B test was done presenting on the screen alternatively the Anchor and the sequence under test; presentation order was randomised trying to equally distribute the content and the quality across the viewing sessions.

The Basic Test Cell (BTC) of this test was made presenting the label “A” on the screen followed by the Anchor and the letter “B” and a coded clip; a message “vote N” was shown for four seconds on the screen to allow the viewers to fill out the scoring sheet. The viewers were told that the sequence was random, i.e. they did not know that “A” was the anchor.

Having to assess a total of 54 cases, and being each BTC 25 seconds long, the total test time was 23’ and 20”; this lead to the design of two test sessions each including the evaluation of 27 coded clips; to examine the behaviour of the viewers six dummy cases were inserted in the test comparing (for each test sequence and QP) the anchor vs. itself.

All scores were collected on paper scoring sheets; the viewers were asked to score 1 when they retained the sequence “A” (the Anchor) was better than “B” (the coded clips) and to score 2 when “A” was worst than “B”; when “A” was equal to “B” the viewers had to score 0.

The MUP player was used together with a high speed PC to provide a smooth flow of UHD content.

The coded video clips were all made of 300 frames. This led to a viewing time of 10 seconds for the sequence “Tango” and of only 5 seconds for the sequences “FoodMarket4” and “Tango”.

It was also noted that:

* in general the compression ratio was rather low and general quality was rather high,
* 5 seconds was a time too short to assess some impairments watching “FoodMarket4” and “Tango”.

For the above reasons it decided (together with many CE2.2 experts) to decrease the frame rate of “FoodMarket4” and “Tango” from 60 to 30 fps, getting a sequence length of 10 seconds allowing a better detection of any possible difference between the Anchor and the coded video clips.

The test sessions were conducted from 11:15 am to 2:20 pm of Saturday July 14 2018, in the test room.

Results of the test are shown in table below.

All the scores “2” (i.e. B better than A) were converted in the score “1”.

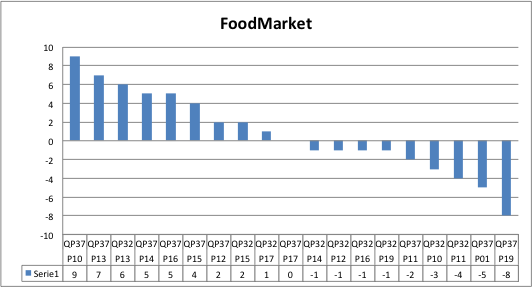
Then all the scores for each test cases were added to get a quality index.

The values of the indexes were ranging from 9 to -8 for the sequence “FoodMarket4”, from 9 to -10 for the sequence “Campfire” and from 6 to -8 for the sequence “Tango”.

The difficulty and the relative low reliability of this testing procedure (for which the Test Chair was discouraging the Experts to proceed) is shown by the “trap” cases inserted in the test, asking the viewers to compare Anchors vs. Anchor (i.e. itself).

A score of 0 (or at least close to 0) was expected for all the six “trap” cases; only in two case the “trap” got 0 and 1 while the other four cases the traps got scores of -6 and -5 (two times).

Here below are reported the graphs with the results.





P01 was the result of comparing the anchor against itself, i.e. identical sequences. The result of this might be judged to be the uncertainty of the test, e.g. in the case of Campfire any result less than +/-6 seem to be random.

The results of the visual test do not allow to draw reasonable conclusions. It was suggested by the test chair to iterate the test by the next meeting (or better before the meeting).

The CE description already mentioned that results should be prepared with QP42, however most participants did not provide them. During the current meeting, some results with QP42 should be assessed together with Vittorio, to judge if that would be a better operation point for comparison.

It was also mentioned that for the next round of viewing, 10s sequences should be used rather than slowing 5s down.

Preferred sequences would be Campfire, Food Market, Park Running.

No conclusion possible – continue CE, AND PLEASE READ THE CE DESCRIPTION CAREFULLY.

Decision (VTM/BMS): Apply the following fixes suggested in JVET-K0307, JVET-K0237, JVET-K0369, JVET-K0232, JVET-K0315:

* Perform deblocking at boundaries of TUs with any size >=64.

There is also the suggestion to avoid duplicate filtering at 4x4 CU boundaries by reducing the deblocking to only 1 sample at the boundary; the current VTM software (the draft text does not specify deblocking) just applies filtering at CU boundaries (which could have a minimum size of 4x4), whereas the original deblocking of HEVC was at 8x8 boundaries. It is not clear if the deblocking at a 4x4 grid is necessary, as already HEVC had 4x4 TUs and PUs, and they were never deblocked. 4x4 deblocking quadruples the worst case complexity, and also has impact on parallelism.

Decision (SW): It is suggested in this context to implement VTM/BMS SW as the original HEVC deblocking, filtering on an 8x8 grid as minimum size. Was discussed in JVET plenary on Sunday and agreed. Kenneth will provide the SW update.

Another fix suggested is related to BMS, where it is suggested to apply deblocking at subblock boundaries as well.

Further study is necessary on the latter two aspects. Include this in CE2

**CE2.3: SAO**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| AhG13 | Coding gain of SAO in VTM | [JVET-K0013](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3911) |
| 2.3.1.a | SAO with EO sign constraints removal | [JVET-K0233](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3743) |
| 2.3.1.b | SAO with EO sign constraints removal and more EO patterns | [JVET-K0233](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3743) |
| 2.3.2.a | SAO Merge with 128×128 unit | [JVET-K0324](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3838) |
| 2.3.2.b | SAO Merge with 128×128 or 256×256 unit | [JVET-K0324](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3838) |
| 2.3.2.c | SAO Merge with 64×64 or 128×128 unit | [JVET-K0324](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3838) |
| 2.3.3.a | SAO Palette results and discussion: Test a (fixed block size, as LCU) | [JVET-K0192](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3701) |
| 2.3.3.b | SAO Palette results and discussion: Test b (adaptive block size) | [JVET-K0192](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3701) |
| 2.3.4.a | SAO Modification: EO (modified threshold) | [JVET-K0176](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3684) |
| 2.3.4.b | SAO Modification: BO offset quantization | [JVET-K0176](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3684) |
| 2.3.4.c | SAO Modification: BO offset band (fixed instead of being signalled) | [JVET-K0176](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3684) |
| 2.3.4.d | SAO Modification: shifted region | [JVET-K0176](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3684) |
| 2.3.4.e | SAO Modification: EO and BO offset range | [JVET-K0176](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3684) |
| 2.3.4.f | SAO Modification: All | [JVET-K0176](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3684) |
| 2.3.4.g | SAO Modification: a + b + d | [JVET-K0176](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3684) |
| 2.3.5 | CTU adaptive sample adaptive offset (modified band position and range) | [JVET-K0153](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3661) |

Results vs. VTM:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **AI** | | | | | **RA** | | | | | **LDB** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| AhG13 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2.3.1.a | 0.00% | 0.03% | 0.03% | 100% | 100% | -0.04% | -0.10% | -0.07% | 100% | 101% | -0.17% | 0.06% | -0.04% | 100% | 99% |
| 2.3.1.b | -0.04% | -0.06% | -0.07% | 101% | 100% | -0.13% | -0.21% | -0.11% | 100% | 100% | -0.15% | -0.32% | -0.21% | 100% | 100% |
| 2.3.2.a | -0.12% | -0.86% | -0.92% | 99% | 104% | -0.25% | -2.00% | -1.82% | 101% | 109% | -0.28% | -2.99% | -3.68% | 101% | 104% |
| 2.3.2.b | -0.12% | -0.84% | -0.91% | 99% | 103% | -0.26% | -2.02% | -1.85% | 100% | 107% | -0.25% | -2.80% | -3.58% | 100% | 105% |
| 2.3.2.c | -0.11% | -0.97% | -1.12% | 93% | 97% | -0.30% | -2.10% | -2.10% | 100% | 100% | -0.43% | -3.28% | -4.30% | 96% | 96% |
| 2.3.3.a | -0.17% | -0.43% | -0.43% | 100% | 92% | -0.29% | -0.90% | -0.72% | 100% | 87% | -0.37% | -1.60% | -1.87% | 96% | 92% |
| 2.3.3.b | -0.17% | -0.33% | -0.41% | 102% | 93% | -0.38% | -1.16% | -1.03% | 97% | 100% | -0.52% | -2.07% | -2.75% | 98% | 111% |
| 2.3.4.a | -0.08% | -0.15% | -0.18% | 99% | 96% | -0.12% | 0.08% | 0.14% | 97% | 94% | 0.03% | 0.51% | 0.66% | 98% | 95% |
| 2.3.4.b | -0.04% | 0.39% | 0.51% | 99% | 97% | -0.04% | 0.43% | 0.51% | 97% | 97% | -0.05% | 0.06% | 0.38% | 100% | 103% |
| 2.3.4.c | -0.06% | 0.86% | 1.17% | 100% | 98% | 0.13% | 0.97% | 1.29% | 98% | 101% | 0.49% | 1.59% | 2.82% | 100% | 101% |
| 2.3.4.d | 0.00% | -0.13% | -0.15% | 99% | 96% | -0.01% | -0.15% | -0.14% | 97% | 94% | 0.02% | -0.20% | -0.49% | 98% | 95% |
| 2.3.4.e | 0.00% | 0.00% | 0.01% | 100% | 100% | -0.01% | 0.00% | 0.03% | 98% | 96% | 0.02% | 0.03% | -0.06% | 101% | 101% |
| 2.3.4.f | -0.15% | 0.63% | 0.90% | 100% | 96% | -0.02% | 1.06% | 1.44% | 97% | 94% | 0.46% | 2.30% | 3.71% | 98% | 95% |
| 2.3.4.g | -0.13% | 0.17% | 0.23% | 100% | 98% | -0.17% | 0.46% | 0.62% | 98% | 98% | -0.06% | 0.58% | 1.14% | 99% | 97% |
| 2.3.5 | 0.06% | -0.75% | -0.84% | 100% | 101% | -0.06% | -0.92% | -0.94% | 100% | 100% | -0.15% | -1.52% | -2.03% | 100% | 101% |

Results vs. BMS:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **AI** | | | | | **RA** | | | | | **LDB** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| AhG13 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2.3.1.a | 0.00% | -0.01% | -0.07% | 100% | 100% | 0.02% | -0.08% | -0.06% | 100% | 100% | -0.07% | -0.07% | -0.04% | 100% | 100% |
| 2.3.1.b | -0.03% | -0.03% | -0.12% | 100% | 100% | -0.03% | -0.24% | -0.17% | 100% | 100% | -0.14% | -0.57% | -0.30% | 101% | 100% |
| 2.3.2.a | -0.09% | -0.74% | -0.69% | 99% | 102% | -0.15% | -2.34% | -2.19% | 101% | 105% | -0.37% | -3.37% | -3.99% | 100% | 103% |
| 2.3.2.b | -0.10% | -0.70% | -0.67% | 98% | 101% | -0.17% | -2.36% | -2.20% | 99% | 104% | -0.42% | -3.53% | -4.27% | 99% | 103% |
| 2.3.2.c | -0.08% | -0.87% | -0.92% | 98% | 98% | -0.18% | -2.43% | -2.40% | 101% | 103% | -0.42% | -3.77% | -4.60% | 99% | 101% |
| 2.3.3.a | -0.15% | -0.13% | 0.00% | 99% | 94% | -0.18% | -1.10% | -0.87% | 101% | 92% | -0.47% | -1.99% | -2.43% | 101% | 97% |
| 2.3.3.b | -0.14% | -0.02% | 0.07% | 101% | 96% | -0.24% | -1.38% | -1.23% | 100% | 95% | -0.56% | -2.60% | -3.15% | 105% | 97% |
| 2.3.4.a | 0.00% | 0.11% | 0.02% | 100% | 98% | 0.01% | 0.10% | 0.05% | 98% | 98% | -0.05% | 0.21% | 0.71% | 101% | 97% |
| 2.3.4.b | -0.04% | 0.63% | 0.70% | 101% | 105% | 0.02% | 0.50% | 0.59% | 100% | 101% | -0.11% | 0.03% | 0.30% | 101% | 103% |
| 2.3.4.c | -0.07% | 1.22% | 1.43% | 102% | 106% | 0.08% | 0.83% | 1.05% | 100% | 102% | 0.09% | 1.26% | 2.18% | 101% | 105% |
| 2.3.4.d | 0.00% | -0.13% | -0.13% | 100% | 98% | 0.01% | -0.17% | -0.13% | 98% | 98% | -0.07% | -0.41% | -0.19% | 99% | 97% |
| 2.3.4.e | 0.00% | 0.00% | 0.00% | 100% | 100% | 0.02% | -0.01% | 0.01% | 99% | 99% | -0.09% | 0.01% | 0.07% | 100% | 100% |
| 2.3.4.f | -0.08% | 1.37% | 1.49% | 101% | 100% | 0.04% | 0.99% | 1.14% | 97% | 99% | 0.09% | 1.80% | 3.39% | 98% | 97% |
| 2.3.4.g |  |  |  |  |  | -0.01% | 0.59% | 0.67% | 98% | 99% |  |  |  |  |  |
| 2.3.5 | 0.07% | -0.90% | -1.02% | 100% | 102% | 0.02% | -0.88% | -0.96% | 100% | 100% | -0.13% | -1.62% | -1.99% | 100% | 101% |

2.3.2.x and 2.3.3.x have picture-level optimization and give gain up to 0.4% in RA compared to VTM, and It was remarked that the same effect could be achieved by a smarter encoder e.g. performing lookahead

The methods that do not use picture level optimization provide small gain (<0.2% for RA compared to VTM)

Compared to BMS, the methods without picture level opt. do not provide gain any more in RA, and the gain of the picture-level methods reduces to <0.25%

Generally, the gain in LDB is slightly higher (around 0.5% vs VTM and BMS for picture level opt.), where however this would introduce one frame delay at the encoder.

The results do not suggest any action at this moment. a) No evidence is currently available that there is subjective improvement; b) when combined with ALF, most gain is lost; c) more smart encoders likely could get similar gain by picture-level optimization (at the expense of additional encoding delay) without syntax change.

**CE2.4 Adaptive Loop Filters**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| AhG13 | Coding gain of ALF in BMS | [JVET-K0013](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3911) |
| 2.4.1.1.a | Subsampled sum-modified-Laplacian with 4×4 level classification | [JVET-K0164](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3672) |
| 2.4.1.1.b | Subsampled sum-modified-Laplacian with 2×2 level classification | [JVET-K0164](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3672) |
| 2.4.1.2 | Adaptive Loop Filter Simplification | [JVET-K0327](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3841) |
| 2.4.1.3 | ALF with Multiplication Replaced by Bit-Shifting | [JVET-K0215](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3725) |
| 2.4.1.4.a | luma 7×7, classifier 2×2, chroma 5×5 | [JVET-K0371](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3889) |
| 2.4.1.4.b | luma 7×7, classifier 2×2, chroma 7×7 | [JVET-K0371](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3889) |
| 2.4.1.4.c | luma 7×7, classifier 4×4, chroma 5×5 | [JVET-K0371](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3889) |
| 2.4.1.4.d | luma 7×7, classifier 4×4, chroma 7×7 | [JVET-K0371](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3889) |
| 2.4.1.4.e | luma 9×9, classifier 2×2, chroma 5×5 | [JVET-K0371](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3889) |
| 2.4.1.4.f | luma 9×9, classifier 2×2, chroma 7×7 | [JVET-K0371](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3889) |
| 2.4.1.4.g | luma 9×9, classifier 4×4, chroma 5×5 | [JVET-K0371](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3889) |
| 2.4.1.4.h | luma 9×9, classifier 4×4, chroma 7×7 | [JVET-K0371](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3889) |
| 2.4.1.4.i | deblocking TC offset setting (-6 for AI, and TC offset -2 for RA and LB) | [JVET-K0371](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3889) |
| 2.4.2.1.a | Multiple-feature based adaptive loop filter (MCALF)  with 4 classifiers (whole proposal) | [JVET-K0285](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3798) |
| 2.4.2.1.b | MCALF with two classifiers (GALF classification) and  (sample-based classification, like BO) | [JVET-K0285](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3798) |
| 2.4.2.1.c | MCALF with two classifiers*and*  (ranking with sample-based classification, 3×3 pattern)  (counting number of samples that are larger or smaller) | [JVET-K0285](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3798) |
| 2.4.2.1.d | MCALF with two classifiers *and*  **(**ranking with local variation-based, cross-sign pattern)  (checking difference in neighborhood) | [JVET-K0285](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3798) |
| 2.4.2.2.a | CTB-based ALF with slice filter sets, 3 classification methods (whole proposal)  ( IntensitySA (like BO), SimilaritySA(5x5 diamond),  and GeometricBA (GALF classification)) | [JVET-K0235](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3745) |
| 2.4.2.2.b | CTB-based ALF with slice filter sets, only IntensitySA and GeometricBA | [JVET-K0235](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3745) |
| 2.4.2.2.c | CTB-based ALF with slice filter sets, only GeometricBA | [JVET-K0235](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3745) |
| 2.4.2.2.d | CTB-based ALF with slice filter sets, only IntensitySA | [JVET-K0235](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3745) |
| 2.4.2.2.e | CTB-based ALF with slice filter sets, single filter is selected at CTB  (including merge with neighbor signalling) | [JVET-K0235](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3745) |
| 2.4.2.3 | Unified Adaptive Loop Filter for Luma and Chroma | [JVET-K0132](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3638) |
| 2.4.2.4 | Modified ALF classification: horizontal and vertical gradients  are calculated using Sobel filter | [JVET-K0151](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3659) |

Properties of Methods:

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Test# | Block classification granularity |  | | # Samples for 1-D Laplacian value calculations  (worst case) | Line buffer size | Filter supports | FF | Filter | Max. Num. of filters per slice to be stored | Classification method selection | Filter decision require whole slice data? |
| BMS | 2×2 (L), N/A (C) |  | | # samples in the whole slice (M) | 8 (L), 4 (C) | 9×9/7×7/5×5 diamond (L);  5×5 diamond (C) | Y | × | 25 (L)  1 (C) | N/A | Y |
| 2.4.1.1.a | 4×4 (L), N/A (C) |  | | M/2  (one dimension  subsample by 2) | 8 (L), 4 (C) | Same as BMS | Y | × | 25 (L)  1 (C) | N/A | Y |
| 2.4.1.1.b | 2×2 (L), N/A (C) |  | | M/2  (one dimension  subsample by 2) | 8 (L), 4 (C) | Same as BMS | Y | × | 25 (L)  1 (C) | N/A | Y |
| 2.4.1.2 | 2×2 (L), N/A (C) |  | | M/4 for highest temporal layer; and M for others  (two dimensions subsampled by 2) | 8 (L), 4 (C) | Same as BMS | Y | × | 25 (L)  1 (C) | N/A | Y |
| 2.4.1.3 | 2×2 (L), N/A (C) |  | | M | 8 (L), 4 (C) | Same as BMS | N | × and << | 25 (L)  1 (C) | N/A | Y |
| 2.4.1.4.a | 2×2 (L), N/A (C) |  | | M | 6(L), 4(C) | 7×7 (L), 5×5 (C) | Y | × | 25 (L)  1 (C) | N/A | Y |
| 2.4.1.4.b | 2×2 (L), N/A (C) |  | | M | 6(L), 6(C) | 7×7 (L), 7×7 (C) | Y | × | 25 (L)  1 (C) | N/A | Y |
| 2.4.1.4.c | 4×4 (L), N/A (C) |  | | M | 6(L), 4(C) | 7×7 (L), 5×5 (C) | Y | × | 25 (L)  1 (C) | N/A | Y |
| 2.4.1.4.d | 4×4 (L), N/A (C) |  | | M | 6(L), 6(C) | 7×7 (L), 7×7 (C) | Y | × | 25 (L)  1 (C) | N/A | Y |
| 2.4.1.4.e | 2×2 (L), N/A (C) |  | | M | 8(L), 4(C) | 9×9 (L), 5×5 (C) | Y | × | 25 (L)  1 (C) | N/A | Y |
| 2.4.1.4.f | 2×2 (L), N/A (C) |  | | M | 8(L), 6(C) | 9×9 (L), 7×7 (C) | Y | × | 25 (L)  1 (C) | N/A | Y |
| 2.4.1.4.g | 4×4 (L), N/A (C) |  | | M | 8(L), 4(C) | 9×9 (L), 5×5 (C) | Y | × | 25 (L)  1 (C) | N/A | Y |
| 2.4.1.4.h | 4×4 (L), N/A (C) |  | | M | 8(L), 6(C) | 9×9 (L), 7×7 (C) | Y | × | 25 (L)  1 (C) | N/A | Y |
| 2.4.1.4.i |  | | Same as in BMS | | | | | | 25 (L)  1 (C) | N/A | Y |
| 2.4.2.1 | 2×2, 1×1 (L),  N/A (C) |  | | M | 8 (L), 4 (C) | Same as BMS | Y | × | 16 (L)  1 (C) | Slice-level | Y |
| 2.4.2.2.a | 2×2, 1×1 (L),  N/A (C) |  | | M | 8(L), 8(C) | 9×9 Cross +3×3Square with half or symmetry | N | × | 16 (L)  2 (C) | Slice-level | Y |
| 2.4.2.2.b | 2×2, 1×1 (L),  N/A (C) |  | | M | 8(L), 8(C) | 9×9 Cross +3×3Square with half or symmetry | N | × | 16 (L)  2 (C) | Slice-level | Y |
| 2.4.2.2.c | 2×2 (L)  N/A (C) |  | | M | 8(L), 8(C) | 9×9 Cross +3×3Square with half or symmetry | N | × | 16 (L)  2 (C) | Slice-level | Y |
| 2.4.2.2.d | 1×1 (L),  N/A (C) |  | | M | 8(L), 8(C) | 9×9 Cross +3×3Square with half or symmetry | N | × | 16 (L)  2(C) | Slice-level | Y |
| 2.4.2.2.e | N/A (L),  N/A (C) |  | | 0 | 8(L), 8(C) | 9×9 Cross +3×3Square with half or symmetry | N | × | 16L + 2C + (# of CTUs in one CTU row) \* 3 (L)  2 (C) | N/A | Y |
| 2.4.2.3 | 2×2 (L),  1×1 (C, inherited from luma) |  | | M | 8 (L), 4 (C) | 9×9/7×7/5×5 diamond (L);  5×5 diamond (C) | Y | × | 25 (L)  1 (C) | N/A | Y |
| 2.4.2.4 |  | | Same as BMS (except gradient calculation is replaced by Sobel filter) | | | | | | | | |

Results vs. VTM:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **AI** | | | | | **RA** | | | | | **LDB** | | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| AhG13 | -3.30% | -3.55% | -4.06% | 101% | 155% | -5.34% | -2.21% | -1.66% | 108% | 184% | - | - | - | - | - |
| 2.4.1.1.a | -3.11% | -3.55% | -4.06% | 110% | 205% | -5.18% | -2.26% | -1.65% | 113% | 268% | -4.41% | -1.80% | -1.85% | 119% | 216% |
| 2.4.1.1.b | -3.19% | -3.55% | -4.06% | 111% | 209% | -5.25% | -2.19% | -1.64% | 114% | 273% | -4.45% | -1.81% | -1.94% | 120% | 219% |
| 2.4.1.2 | -3.30% | -3.55% | -4.06% | 102% | 170% | -5.26% | -2.23% | -1.67% | 109% | 200% | -4.55% | -1.80% | -1.95% | 114% | 170% |
| 2.4.1.3 | -3.04% | -3.44% | -3.95% | 102% | 149% | -5.15% | -2.19% | -1.56% | 107% | 176% | -4.52% | -1.79% | -2.16% | 112% | 154% |
| 2.4.1.4.a\* | -3.10% | -3.56% | -4.07% | 101% | 124% | -5.10% | -2.29% | -1.69% | 105% | 135% | -4.32% | -1.79% | -1.77% | 106% | 119% |
| 2.4.1.4.b\* | -3.09% | -4.20% | -4.73% | 101% | 123% | -5.11% | -3.16% | -2.66% | 104% | 134% | -4.31% | -2.63% | -2.72% | 107% | 122% |
| 2.4.1.4.c\* | -2.99% | -3.56% | -4.07% | 101% | 114% | -4.99% | -2.24% | -1.66% | 104% | 122% | -4.19% | -1.64% | -1.85% | 107% | 114% |
| 2.4.1.4.d\* | -2.98% | -4.20% | -4.73% | 102% | 114% | -4.98% | -3.14% | -2.66% | 104% | 122% | -4.23% | -2.44% | -2.57% | 107% | 116% |
| 2.4.1.4.e\* | -3.28% | -3.54% | -4.05% | 103% | 130% | -5.33% | -2.21% | -1.63% | 108% | 143% | -4.56% | -1.81% | -2.01% | 110% | 125% |
| 2.4.1.4.f\* | -3.27% | -4.19% | -4.71% | 102% | 129% | -5.35% | -3.11% | -2.57% | 107% | 142% | -4.63% | -2.61% | -2.88% | 111% | 128% |
| 2.4.1.4.g\* | -3.17% | -3.54% | -4.05% | 102% | 119% | -5.21% | -2.20% | -1.66% | 107% | 129% | -4.46% | -1.83% | -2.01% | 110% | 117% |
| 2.4.1.4.h\* | -3.16% | -4.19% | -4.71% | 103% | 121% | -5.23% | -3.15% | -2.63% | 108% | 129% | -4.49% | -2.59% | -2.72% | 111% | 116% |
| 2.4.1.4.i | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 2.4.2.1.a | -3.32% | -3.55% | -4.06% | 103% | 162% | -5.61% | -2.22% | -1.66% | 111% | 197% | -5.19% | -1.74% | -2.04% | 117% | 167% |
| 2.4.2.1.b | -3.31% | -3.55% | -4.06% | 102% | 165% | -5.53% | -2.21% | -1.67% | 109% | 200% | -4.97% | -1.96% | -2.13% | 115% | 167% |
| 2.4.2.1.c | -3.31% | -3.55% | -4.06% | 102% | 164% | -5.60% | -2.21% | -1.64% | 109% | 197% | -5.14% | -1.99% | -1.94% | 113% | 165% |
| 2.4.2.1.d | -3.30% | -3.55% | -4.06% | 102% | 166% | -5.53% | -2.24% | -1.66% | 110% | 201% | -4.90% | -1.67% | -2.04% | 115% | 169% |
| 2.4.2.2.a | -2.70% | -4.94% | -5.69% | 102% | 133% | -4.85% | -5.25% | -5.08% | 101% | 145% | -4.60% | -7.42% | -7.19% | 102% | 137% |
| 2.4.2.2.b | -2.69% | -4.94% | -5.69% | 101% | 132% | -4.82% | -5.24% | -5.09% | 101% | 144% | -4.52% | -7.51% | -6.97% | 101% | 135% |
| 2.4.2.2.c | -2.66% | -4.94% | -5.69% | 101% | 134% | -4.66% | -5.26% | -5.10% | 100% | 148% | -4.19% | -7.48% | -6.97% | 101% | 136% |
| 2.4.2.2.d | -2.07% | -4.96% | -5.71% | 101% | 117% | -3.72% | -5.12% | -4.95% | 100% | 123% | -3.38% | -7.30% | -6.96% | 101% | 118% |
| 2.4.2.2.e | -1.89% | -4.97% | -5.71% | 101% | 114% | -3.41% | -5.16% | -4.96% | 100% | 118% | -2.94% | -7.33% | -6.95% | 101% | 113% |
| 2.4.2.3 | -3.24% | -4.79% | -5.47% | 105% | 144% | -5.31% | -6.69% | -5.38% | 111% | 173% | -4.62% | -8.68% | -7.52% | 116% | 154% |
| 2.4.2.4\*\* | -3.25% | -3.51% | -4.02% | 109% | 268% | -5.34% | -2.21% | -1.63% | 110%\*\* | 370% | -4.68% | -1.55% | -2.04% | 124% | 275% |

Results vs. BMS:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **AI** | | | | | **RA** | | | | | | **LDB** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | | **U** | **V** | **EncT** | **DecT** |
| AhG13 | -2.74% | -3.43% | -3.57% | 100% | 139% | -5.06% | -2.20% | -1.33% | 102% | 152% | - | | - | - | - | - |
| 2.4.1.1.a | -2.59% | -3.43% | -3.57% | 101% | 171% | -4.91% | -2.23% | -1.37% | 101% | 183% | -4.14% | | -1.65% | -1.57% | 102% | 171% |
| 2.4.1.1.b | -2.65% | -3.43% | -3.57% | 101% | 174% | -4.96% | -2.25% | -1.36% | 101% | 186% | -4.24% | | -1.72% | -1.62% | 105% | 173% |
| 2.4.1.2 | -2.74% | -3.43% | -3.57% | 100% | 154% | -4.98% | -2.24% | -1.34% | 102% | 169% | -4.24% | | -1.63% | -1.64% | 103% | 154% |
| 2.4.1.3 | -2.50% | -3.32% | -3.46% | 101% | 139% | -4.84% | -2.20% | -1.25% | 102% | 150% | -4.29% | | -1.74% | -1.92% | 103% | 143% |
| 2.4.1.4.a\* | Missing |  |  |  |  | Missing |  |  |  |  | Missing | |  |  |  |  |
| 2.4.1.4.b\* | -2.55% | -4.11% | -4.25% | 101% | 116% | -4.83% | -3.22% | -2.43% | 101% | 122% | -4.04% | | -2.39% | -2.31% | 101% | 118% |
| 2.4.1.4.c\* | Missing |  |  |  |  | Missing |  |  |  |  | Missing | |  |  |  |  |
| 2.4.1.4.d\* | -2.47% | -4.11% | -4.25% | 101% | 111% | -4.71% | -3.22% | -2.42% | 101% | 114% | -3.98% | | -2.31% | -2.55% | 101% | 111% |
| 2.4.1.4.e\* | Missing |  |  |  |  | Missing |  |  |  |  | Missing | |  |  |  |  |
| 2.4.1.4.f\* | -2.71% | -4.09% | -4.24% | 100% | 122% | -5.05% | -3.17% | -2.39% | 101% | 127% | -4.35% | | -2.48% | -2.53% | 103% | 124% |
| 2.4.1.4.g\* | Missing |  |  |  |  | Missing |  |  |  |  | Missing | |  |  |  |  |
| 2.4.1.4.h\* | -2.63% | -4.09% | -4.23% | 101% | 115% | -4.96% | -3.16% | -2.42% | 102% | 119% | -4.30% | | -2.48% | -2.62% | 103% | 118% |
| 2.4.1.4.i | - | - | - | - | - | - | - | - | - | - | - | | - | - | - | - |
| 2.4.2.1.a | -2.77% | -3.43% | -3.57% | 100% | 145% | -5.34% | -2.19% | -1.31% | 102% | 159% | -4.86% | | -1.76% | -1.83% | 104% | 148% |
| 2.4.2.1.b | -2.76% | -3.43% | -3.57% | 100% | 144% | -5.26% | -2.22% | -1.32% | 102% | 157% | -4.71% | | -1.75% | -1.73% | 104% | 149% |
| 2.4.2.1.c | -2.76% | -3.43% | -3.57% | 100% | 145% | -5.34% | -2.17% | -1.31% | 102% | 159% | -4.83% | | -1.76% | -1.60% | 103% | 149% |
| 2.4.2.1.d | -2.75% | -3.43% | -3.57% | 100% | 146% | -5.27% | -2.24% | -1.32% | 102% | 158% | -4.63% | | -1.89% | -1.88% | 104% | 151% |
| 2.4.2.2.a | -2.18% | -4.95% | -5.60% | 100% | 124% | -4.62% | -5.40% | -5.18% | 101% | 129% | -4.35% | | -7.24% | -6.96% | 100% | 128% |
| 2.4.2.2.b | -2.17% | -4.95% | -5.60% | 100% | 123% | -4.57% | -5.39% | -5.17% | 101% | 127% | -4.24% | | -7.40% | -6.72% | 101% | 126% |
| 2.4.2.2.c | -2.13% | -4.94% | -5.60% | 100% | 125% | -4.43% | -5.42% | -5.20% | 100% | 131% | -3.95% | | -7.42% | -6.79% | 100% | 128% |
| 2.4.2.2.d | -1.66% | -4.99% | -5.63% | 100% | 113% | -3.56% | -5.29% | -5.05% | 100% | 114% | -3.19% | | -7.37% | -7.11% | 101% | 115% |
| 2.4.2.2.e | -1.49% | -5.00% | -5.65% | 100% | 110% | -3.27% | -5.31% | -5.09% | 100% | 112% | -2.78% | | -7.41% | -6.80% | 101% | 111% |
| 2.4.2.3 | -2.70% | -4.46% | -4.68% |  |  | -5.04% | -6.14% | -4.24% |  |  | -4.36% | | -8.00% | -6.87% |  |  |
| 2.4.2.4\*\* | -2.69% | -3.39% | -3.53% | 103% | 232% | -5.04% | -2.21% | -1.30% | 103%\*\* | 263% | -4.39% | | -1.76% | -1.81% | 104%\*\* | 247% |

The following table shows additional optional tests (only for BMS):

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **AI** | | | | | **RA** | | | | | **LDB** | | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | | **V** | **EncT** | **DecT** |
| 2.4.1.1 | 0.15% | 0.01% | 0.01% | 101% | 97% | 0.17% | -0.03% | -0.04% | 100% | 97% | 0.10% | -0.03% | 0.08% | | 100% | 97% |
| 2.4.1.2 | - | - | - | - | - | 0.09% | -0.04% | -0.01% | 100% | 97% |  |  |  | |  |  |
| 2.4.1.3 | 0.25% | 0.11% | 0.11% | 101% | 100% | 0.24% | 0.00% | 0.07% | 100% | 100% | -0.06% | -0.12% | -0.28% | | 100% | 101% |
| 2.4.1.4.i | -0.87% | -0.57% | -0.10% | 100% | 100% | -0.19% | -0.55% | -0.50% | 100% | 100% | -0.44% | -0.33% | -0.09% | | 100% | 99% |
| 2.4.2.3 | 0.04% | -1.06% | -1.14% | 100% | 100% | 0.03% | -4.02% | -2.95% | 101% | 100% | -0.13% | -6.38% | -5.29% | | 102% | 101% |

2.4.1.4.i is modifying the parameters of deblocking (encoder only). Before such a change is done, it should be tested for visual quality.

From discussion in track B:

In general, we are seeking for complexity reduction compared to BMS-GALF (giving -5.3% compared to VTM) rather than increasing complexity.

Methods of increasing complexity (e.g. pixel based classification, multiple classifiers, classification or switching also for chroma) give at most another 0.25% over BMS-GALF

Options to reduce complexity:

- Simpler classifiers (no results from CE)

- Classification block sizes (4x4 loses 0.15% relative to BMS-GALF)

- Subsampling in classification

- Avoiding pre-defined filters (BMS-GALF has 400, and some CE results reduce them to less or zero)

- Filter size (7x7 loses 0.2%), which implies less operations and line buffer reduction.

- Multiplication simplifications (e.g. bit shifting loses <0.1%, see JVET-K0215)

- Omitting classification (e.g. switching CTU based, 2.4.2.2e loses 2% relative to BMS-GALF)

It was suggested to make a subjective comparison of VTM vs. a simplified classification based approach and an approach without classification, but several experts expressed they would not expect that differences would be visible.

Breakout activity (L. Zhang) to assess the implementation complexity (memory accesses, memory for storing predefined filters, operations per sample in terms of mul, add, comp, reload operations, etc.) of the CTU based approach and different aspects of the simplified classification based approaches. This should gives us data to assess the complexity aspect versus the performance.

Another aspect that requires clarification is the signalling of the filter coefficients. The current BMS-GALF solution of signalling between slice header and first CTU is not desirable. Also the case of multiple slices per picture should be supported.

The BoG report JVET-K0521 was presented in track B Sat. 14th 1900.

An analysis was done on the algorithmic and memory complexity of different algorithms.

The two solutions with lowest complexity are 2.4.1.4.c and 2.4.2.2.e.

Both are approximately identical in terms of number of multiplications, significantly reduced relative to BMS-GALF

2.4.1.4.c has more additions and shifts

The classification at 4x4 block level does not have high complexity as compared to the filtering itself.

It is agreed that the classification based approach provides the best performance (1.5% coding gain).

Some concern is however raised with regard to the representation/coding of filter coefficients, in particular concerning the prediction aspects. This should be further studied, to make the representation of ALF parameters more straightforward.

Decision(VTM): Adopt JVET-K0371 (based on subtest 2.4.1.4c, 4x4 classification based on Laplacian for luma only, 7x7 luma, 5x5 chroma filters); disable prediction of adaptive filters from fixed filter set; disable temporal prediction; put filter parameters into slice header; Enabling flag at CTU level.

Further investigation in ongoing CE: Prediction of filter parameters; enabling at sub-CTU level; other classification approaches from 2.4.1.1 and 2.4.1.2. Also study aspects of 2.4.1.3 that replace multiplication in filtering by shift operations.

Was revisited later. Draft text was provided (JVET-K0564), containing description of the method described above (syntax, semantics, decoding process).

[JVET-K0564](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4094) Specification draft for Adaptive Loop Filter [V. Seregin, N. Hu, M. Karczewicz (Qualcomm)] [late]

Some suggestions were made as follows:

* Specifiy ranges of variables
* Correctly specify bit depth of filtering operations
* Impose constraints that an encoder should not send coefficients which cause overflows
* Describe in a way that it is neutral about bit depth of the signal samples

Looks generally OK, but probably needs more detailed check by editors.

**Sub-CE5: Non-local filter**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| 2.5.1 | Non-local Structure-based Filter | [JVET-K0160](file:///C:\Users\admin\Desktop\proposal\current_document.php%3fid=3668) |
| 2.5.2 | Non-local mean in-loop filter | [JVET-K0236](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3746) |
| 2.5.3 | Noise Suppression Filter | [JVET-K0053](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3554) |

Notes: SVD: singular-value decomposition

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Test# | Line buffer size | Filter unit | Search window size | Overlapped units? | matching criterion | Division  Required  ? | How to filter | Matrix size/filter tap | Filter coeffs | On/off control | Stage |
| 2.5.1 | 8 (L), 4 (C) | 6×6 | 32×32 | Y | SSD | Y | SVD-based filtering | 21×6×6 | Derived on-the-fly | Frame and CTU level | After deblocking, before SAO |
| 2.5.2 | 16 (L) 8 (C) | 8×8 | 33 x 33 | N | SSD | Y | Linear filter with normalization | 16-tap | Based on estimated quantization noise (signalled index) and SSD between current and reference patch | Slice/CTB/  32×32 | After deblocking, before SAO |
| 2.5.3 | 0 (L) 0 (C) | 8×8 | 16×16 within CTB | N | SSD | Y | Noise Suppressor’s Collaborative Filter in Hadamard transform domain | Hadamard transform 8x1; 8x1-tap filter | Based on QP value | Slice/256×256/128×128/64×64/32×32 | After SAO, before ALF |

Results vs. VTM

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **AI** | | | | | **RA** | | | | | **LDB** | | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | | **V** | **EncT** | **DecT** |
| 2.5.1 | -2.55% | -4.26% | -5.03% | 108% | 4748% | -3.66% | -2.26% | -2.56% | 100% | 2431% | -2.20% | -0.84% | -1.27% | | 112% | 3086% |
| 2.5.2\* | -0.63% | -1.98% | -1.99% | 102% | 235% | -1.24% | -3.95% | -3.49% | 100% | 211% | -1.00% | -3.99% | -4.24% | | 101% | 179% |
| 2.5.3.a\* | -0.61% | -1.58% | -1.79% | 100% | 135% | -1.03% | -2.04% | -1.89% | 100% | 134% | -0.67% | -1.49% | -1.59% | | 100% | 128% |
| 2.5.3.b\* | -0.61% | 0.09% | 0.09% | 100% | 125% | -0.99% | -0.02% | 0.00% | 100% | 123% | -0.62% | 0.69% | 0.50% | | 100% | 119% |

Results vs. BMS

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **AI** | | | | | **RA** | | | | | **LDB** | | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | | **V** | **EncT** | **DecT** |
| 2.5.1 | -0.93% | -1.95% | -2.30% | 100% | 2581% | -1.47% | -1.88% | -2.16% | 97% | 1255% | -1.53% | -1.10% | -0.88% | | 103% | 1861% |
| 2.5.2\* | -0.31% | -1.45% | -1.48% | 100% | 167% | -0.57% | -3.57% | -3.25% | 100% | 139% | -0.76% | -3.94% | -3.99% | | 101% | 138% |
| 2.5.3.a\* | -0.34% | -0.81% | -0.90% | 100% | 119% | -0.62% | -2.41% | -2.42% | 100% | 114% | -0.73% | -2.58% | -2.35% | | 100% | 117% |
| 2.5.3.b\* | -0.34% | 0.11% | 0.11% | 100% | 113% | -0.60% | 0.18% | 0.18% | 100% | 109% | -0.68% | 0.40% | 0.43% | | 100% | 112% |

2.5.1: Current SVD as used in CE uses floating point implementation. There was a proposal on fixed point implementation by the last meeting, but this was not investigated in CE. Overall, decoder complexity of 2.5.1 is very high. The compression is reduced (but still around 1.5% rate reduction) when combined with GALF (in BMS). The method itself is probably more complex than GALF but provides less gain on top of VTM.

2.5.2/2.5.3 These two approaches provide 1.2% (non-local mean filter) and 1% (Hadamard based noise suppression). This reduces to roughly 0.6% when used in BMS (probably less gain when combined with ALF). Both of these approaches are more complex than e.g. the current design of the bilateral filter from CE2.1, which still gives similar gain (0.5%) for BMS.

No action at this moment, further study for significant complexity reduction. Would also be interesting to identify in which cases the non-local filters are able to provide gain that the other loop filters cannot provide. For the SVD based approach, the proponents report that the usage is between 10% and 50%, depending on sequence.

As a general statement, VVC should have an overall clean design, and it should be avoided to operate a large number of loop filters sequentially, unless they provide substantial individual gains, and not being overly complex.

[JVET-K0053](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3554) CE2: Noise Suppression Filter (Test 2.5.3) [R. Chernyak, V. Stepin, S. Ikonin, J. Chen (Huawei)]

[JVET-K0112](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3616) CE2: long-tap deblocking filter (Test 2.2.1.5) [W. Choi, C. Kim (Samsung)]

[JVET-K0129](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3634) CE2: Deblocking filter with asymmetric weighting (CE2-2.2.1) [T. Toma, K. Abe (Panasonic)]

[JVET-K0132](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3638) CE2.4.2.3 Unified Adaptive Loop Filter for Luma and Chroma [J. Zheng, Q. Yu, Y.Lin (HiSilicon)]

[JVET-K0151](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3659) CE2: Modified ALF classification (CE2-4.2.4) [M. Ikeda, T. Suzuki (Sony)]

[JVET-K0152](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3660) CE2: Long-tap deblocking filter for luma and chroma (CE2-2.1.6) [M. Ikeda, T. Suzuki (Sony)]

[JVET-K0153](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3661) CE2.3.5 CTU adaptive sample adaptive offset [T. Ikai (Sharp)]

[JVET-K0160](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3668) CE2: Non-local Structure-based Filter [X. Meng, C. Jia, Z. Wang, S. S. Wang, S. Ma (Peking Univ.), X. Zheng (DJI)]

[JVET-K0164](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3672) CE2: Subsampled sum-modified-Laplacian (Test 4.1.1) [S.-C. Lim, J. Kang, H. Lee, J. Lee, S. Cho, H. Y. Kim (ETRI)]

Note: This document was withdrawn by mistake. Proponents were asked to register it again under a new number. – Add this number

[JVET-K0176](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3684) CE2: SAO modification (CE2.3.4) [J. Chen, K. Choi, C. Kim (Samsung)]

[JVET-K0192](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3701) CE2-3.3 SAO Palette results and discussion [P. Bordes, F. Racape (Technicolor)]

[JVET-K0215](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3725) CE2: ALF with Multiplication Replaced by Bit-Shifting (Test 4.1.3) [S. Esenlik, Z. Zhao, J. Chen (Huawei)]

[JVET-K0231](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3741) CE2.1.2: Bilateral filter - spatial filter strength adjustment [Y.-C. Su, C.-Y. Chen, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-K0232](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3742) CE2.2.1.3: Long deblocking filters [C.-M. Tsai, T.-D. Chuang, C.-W. Hsu, C.-Y. Chen, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-K0233](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3743) CE2.3.1: SAO with EO sign constraints removal and more EO patterns [C.-Y. Chen, C.-Y. Lai, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-K0235](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3745) CE2.4.2.2: CTB-based ALF with slice filter sets [C.-Y. Chen, Y.-C. Su, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-K0236](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3746) CE2.5.2: Non-local mean in-loop filter [C.-Y. Lai, C.-Y. Chen, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-K0285](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3798) CE2.4.2.1: Multiple-feature based adaptive loop filter [W.-Q. Lim, J. Erfurt, M. Siekmann, H. Schwarz, D. Marpe, T. Wiegand (HHI)]

[JVET-K0307](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3821) CE2-2.1.1: Long deblocking filters and fixes [K. Andersson, Z. Zhang, R. Sjöberg (Ericsson)]

[JVET-K0315](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3829) CE2: Deblocking Improvements for Large CUs (Test 2.1.7) [W. Zhu, K. Misra, A. Segall, P. Cowan (Sharp)]

[JVET-K0324](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3838) CE2: Tests on SAO design in CE2.3.2 [A. Gadde, D. Rusanovskyy, M. Karczewicz (Qualcomm)]

[JVET-K0327](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3841) CE2: Adaptive Loop Filter Simplification (Test 2.4.1.2) [R. Vanam, Y. He, Y. Ye (InterDigital)]

[JVET-K0334](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3849) CE2: Tests on long deblocking (CE2.2.1.4) [D. Rusanovskyy, J. Dong, M. Karczewicz (Qualcomm)]

[JVET-K0371](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3889) CE2.4.1.4: Reduced filter shape size for ALF [M. Karczewicz, N. Shlyakhov, N. Hu, V. Seregin, W.-J. Chien (Qualcomm)]

[JVET-K0384](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3903) CE2.1.3: In-loop bilateral filter [A. Gadde, V. Seregin, W.-J. Chien, M. Karczewicz (Qualcomm)]

[JVET-K0386](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3905) CE2-2.2.2: Luma-adaptive deblocking filter [S. Nemoto, S. Iwamura, A. Ichigaya (NHK)] [late]

[JVET-K0393](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3913) CE2: Extended Deblocking Filter (CE2.2.1.2) [Kyohei Unno, Kei Kawamura, Yoshitaka Kidani, Sei Naito (KDDI)] [late]

[JVET-K0435](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3958) Crosscheck for CE2-5.1 [W. Zhang (Hulu)] [late]

## CE3: Intra prediction and mode coding (38)

Contributions in this category were discussed Wednesday 11 July 1140–1800 (chaired by GJS).

[JVET-K0023](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3869) CE3: Summary Report on Intra Prediction and Mode Coding [G. Van der Auwera, J. Heo, A. Filippov]

The goal of CE3 is to study intra prediction tools including mode coding for the VVC standard. There are 7 sub-tests defined. In CE3.1 tests are included targeting DC, planar, directional, and additional modes. CE3.2 tests reference sample filtering, interpolation, across boundary filtering and prediction sample filtering. CE3.3 targets intra mode coding, such as most probable mode list variations. In CE3.4 cross-component linear model and variations are tested. CE3.5 focuses on multi-reference line intra prediction tests and CE3.6 on non-linear intra prediction. In CE3.7 tests are performed regarding bidirectional intra prediction.

The following is the list of defined sub-tests in CE3:

* CE3.1: Intra modes
* CE3.2: Intra filtering and interpolation
* CE3.3: Intra mode coding
* CE3.4: Cross-component linear model (CCLM)
* CE3.5: Multi-reference line intra prediction
* CE3.6: Non-linear intra prediction
* CE3.7: Bidirectional intra prediction

The CE3 description originally defined 75 tests of which 7 were withdrawn and one redefined. This document summarizes the objective results (BD-rates, runtimes), cross-check reports and related input contributions.

Comments from cross-checkers were copied verbatim into the summary report.

**CE3.1 on ‘Intra modes’**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Short description** | **Doc. #** |
| 1.1.1 | Use 129 directional modes for all blocks | JVET-K0060  (Qualcomm) |
| 1.1.2 | Use variable number of directional modes (33, 65, or 129) depending on block size comparison with two SPS thresholds |
| 1.2.1 | DC mode with only shift operators | JVET-K0211  (Panasonic) |
| 1.3.1 | Wide-angle prediction | JVET-K0046  (Nokia) |
| 1.4.1 | Usage of line-based intra prediction mode | JVET-K0049  (HHI) |
| 1.4.2 | Fast Line-based intra prediction mode |
| 1.4.3 | Constrained Line-based intra prediction mode |
| 1.5.1 | Unequal weighted planar prediction (UWP) | JVET-K0055  (ARRIS, Sharp) |

CE3.1: ‘All Intra Main10’ results

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **All Intra Main10 – Over VTM1.0** | | | | | **All Intra Main10 – Over BMS1.0** | | | | |
| **Test#** | **Description** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 1.1.1 | Use 129 directional modes for all blocks | −1.30% | −1.43% | −1.42% | 119% | 101% | 0.04% | −0.10% | −0.13% | 102% | 100% |
| 1.1.2 | Use variable number of directional modes (33, 65, or 129) depending on block size | −1.44% | −1.45% | −1.48% | 116% | 101% | −0.14% | −0.13% | −0.18% | 101% | 101% |
| 1.2.1 | DC mode with only shift operators | 0.01% | 0.04% | 0.01% | 100% | 99% | 0.01% | 0.09% | 0.02% | 100% | 99% |
| 1.3.1 | Wide-angle prediction | −0.28% | −0.38% | −0.42% | 106% | 101% | −0.23% | −0.23% | −0.30% | 102% | 98% |
| 1.4.1 | Usage of line-based intra prediction mode | −2.34% | −2.15% | −2.47% | 293% | 120% | −0.70% | −0.99% | −1.15% | 125% | 105% |
| 1.4.2 | Fast line-based intra prediction mode | −2.00% | −1.83% | −2.09% | 164% | 112% | −0.35% | −0.55% | −0.57% | 106% | 102% |
| 1.4.3 | Constrained line-based intra prediction mode | −1.58% | −1.74% | −1.98% | 146% | 110% | −0.26% | −0.39% | −0.49% | 105% | 102% |
| 1.5.1 | Unequal weighted planar prediction (UWP) | −0.20% | −0.03% | −0.06% | 101% | 103% | −0.53% | −0.53% | −0.53% | 100% | 102% |

Focus on 1.2.1: There is a division by width + height, which is bad for non-square blocks. This proposal fixes that and has no coding loss, which seems ripe for action pending review of non-CE inputs K0122 and K0400. [Resolved per notes elsewhere]

Focusing on 1.4.1: This has the most coding gain in the table, but has an increase in encoder complexity (requiring a line-by-line operation similar to a 1-D transform). Test 1.4.2 has a reduced-complexity encoding with the same syntax and decoding process. It has an increase in decoding complexity and from the encoder perspective, there is an extra switch to evaluate. It was commented that 1.4.1 has substantial complexity.

Focusing on 1.3.1: It was commented that there is no significant complexity associated with that. It does have an extra switch bit and two variations to support in the decoder side. There are non-CE inputs on that. Those should be reviewed.

It was noted that these are AI results, whereas the impact of these techniques is roughly cut in half for RA conditions.

Focus on the BMS: It was suggested that we should consider adopting the 67 modes that are in the BMS. This question was deferred to after consideration of the mode coding.

**CE3.2 on ‘Intra filtering and interpolation’**

|  |  |  |
| --- | --- | --- |
| **Test #** | **Short Description** | **Doc. #** |
| 2.1.1 | Bilateral reference sample filter | JVET-K0061  (Qualcomm) |
| 2.2.1 | 4-tap cubic filter for extending reference samples | JVET-K0062  (Qualcomm) |
| 2.2.2 | Bilinear interpolation for extending reference samples | JVET-K0211  (Panasonic) |
| 2.3.1 | Interpolation filter selection between 6-tap cubic and 4-tap Gaussian filter based on block size | JVET-K0062  (Qualcomm) |
| 2.3.2 | Combine tests 2.2.1 + 2.3.1 |
| 2.3.3 | Interpolation filter selection between 4-tap cubic and 4-tap Gaussian filter based on intra prediction mode and block size | JVET-K0097  (LGE) |
| 2.3.4 | Interpolation filter selection between 6-tap cubic and 4-tap Gaussian filter based on intra prediction mode and block size |
| 2.4.1 | Simplified position dependent intra prediction combination (PDPC) | JVET-K0063  (Qualcomm) |
| 2.5.1 | Bilinear interpolation for projection and smoothing after projection | JVET-K0211  (Panasonic) |
| 2.6.1 | 6-tap combined filter without reference sample smoothing | JVET-K0165  (ETRI) |
| 2.7.1 | Bilateral reference sample filter | JVET-K0043  (HHI) |
| 2.8.2 | Mode dependent de-ringing filter based on bitstream flag | JVET-K0066  (Huawei) |
| 2.9.1 | Intra boundary filters | JVET-K0240  (MediaTek) |
| 2.10.1 | Multiple 4-tap filter | JVET-K0179  (Samsung) |
| 2.11.1 | Multi-combined intra prediction | JVET-K0180  (Samsung) |

CE3.2: ‘All Intra Main10’

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **All Intra Main10 – Over VTM1.0** | | | | | **All Intra Main10 – Over BMS1.0** | | | | |
| **Test #** | **Description** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 2.1.1 | Bilateral reference sample filter | 0.03% | −0.02% | −0.04% | 100% | 101% | −0.13% | −0.16% | −0.17% | 100% | 99% |
| 2.2.1 | 4-tap cubic filter for extending reference samples | −0.05% | −0.08% | −0.11% | 102% | 100% | −0.01% | −0.04% | −0.03% | 100% | 99% |
| 2.2.2 | Bilinear interpolation for extending reference samples | −0.05% | −0.08% | −0.08% | 101% | 100% | 0.00% | 0.01% | −0.03% | 100% | 99% |
| 2.3.1 | Interpolation filter selection between 6-tap cubic and 4-tap Gaussian filter based on block size | −0.57% | −0.39% | −0.43% | 107% | 103% | −0.44% | −0.48% | −0.53% | 101% | 101% |
| 2.3.2 | Combine tests 2.2.1 + 2.3.1 | −0.63% | −0.49% | −0.48% | 107% | 103% | −0.46% | −0.58% | −0.58% | 102% | 100% |
| 2.3.3 | Interpolation filter selection between 4-tap cubic and 4-tap Gaussian filter based on intra prediction mode and block size | −0.54% | −0.41% | −0.46% | 106% | 103% | −0.40% | −0.41% | −0.43% | 106% | 104% |
| 2.3.4 | Interpolation filter selection between 6-tap cubic and 4-tap Gaussian filter based on intra prediction mode and block size | −0.58% | −0.39% | −0.43% | 107% | 102% | −0.44% | −0.57% | −0.57% | 101% | 101% |
| 2.4.1 | Simplified position dependent intra prediction combination (PDPC) | −0.97% | −0.14% | −0.01% | 109% | 107% | −1.35% | −0.94% | −0.81% | 103% | 105% |
| 2.5.1 | Bilinear interpolation for projection and smoothing after projection | −0.01% | −0.11% | −0.13% | 102% | 100% | 0.02% | 0.01% | −0.04% | 100% | 99% |
| 2.6.1 | 6-tap combined filter without reference sample smoothing | −0.55% | −0.01% | 0.02% | 108% | 102% | −0.39% | −0.11% | −0.16% | 101% | 101% |
| 2.7.1 | Bilateral reference sample filter | −0.22% | −0.11% | −0.14% | 101% | 101% | −0.24% | −0.20% | −0.24% | 101% | 102% |
| 2.8.2 | Mode dependent de-ringing filter based on bitstream flag | −0.40% | −0.26% | −0.32% | 116% | 100% | −0.40% | −0.37% | −0.39% | 115% | 100% |
| 2.9.1 | Intra boundary filters | −0.75% | −0.69% | −0.75% | 104% | 102% | −0.87% | −0.55% | −0.59% | 103% | 106% |
| 2.10.1 | Multiple 4-tap filter | −0.41% | −0.19% | −0.16% | 112% | 103% | −0.13% | −0.11% | −0.16% | 103% | 102% |
| 2.11.1 | Multi-combined intra prediction | −0.12% | −0.04% | −0.06% | 196% | 100% | 0.01% | 0.04% | 0.00% | 212% | 101% |

Focus on 2.4.1: This has some complexity, but it was suggested that the gain (relative to VTM: 0.97% in AI, 0.48% in RA) is worth that. Some participants suggested instead using the intra boundary filter approach of 2.9.1, which is asserted to be less complex. In spirit, PDPC is a merging of boundary filtering with the intra prediction process into a single formula. Decision: Adopt PDPC (per K0063).

Focus on 2.3.3: It was remarked that 2.3.1 and 2.3.4 are conceptually similar and there may be some relationship with 2.6.1. Basically it was remarked that 2.3.3 appears to be a good starting point by having a shorter filter. A participant said they had experimented with 2.3.3 combined with PDPC and there was not a diminishment of the gain. It was agreed to further study this in a CE.

**CE3.3 on ‘Intra mode coding’**

VTM has 35 modes and an MPM list with 3 modes. BMS has 67 modes and a primary MPM list with 6 modes, and then a “selected list” of secondary modes, and then a set of remaining modes.

Defined tests

|  |  |  |
| --- | --- | --- |
| **Test #** | **Short Description** | **Doc. #** |
| 3.1.1 | Intra 67, 6 modes in primary MPM, secondary MPM, shape dependency | JVET-K0081  (Qualcomm) |
| 3.2.1 | Priority based list with primary MPM, secondary MPM and first few default modes prioritized (method 1 from Samsung): TM intra modes | JVET-K0181  (Samsung) |
| 3.2.2 | Priority based list with primary MPM, secondary MPM and first few default modes prioritized (method 2 from Huawei): 67 intra modes | JVET-K0365  (Huawei) |
| 3.2.3 | 6 MPM + Selected + Non-selected modes list (JEM macro JVET\_B0051\_NON\_MPM\_MODE), 67 intra modes | JVET-K0368  (Huawei) |
| 3.3.1 | MPM list construction based on dependency between neighboring intra modes | JVET-K0087  (LGE) |

CE3.3: Test results

CE3.3: ‘All Intra Main10’

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Test #** | **Description** | **All Intra Main10 - Over VTM1.0** | | | | | **All Intra Main10 - Over BMS1.0** | | | | |
| **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 3.1.1 | Intra 67, 6 modes in primary MPM, secondary MPM, shape dependency | -1.41% | -1.33% | -1.35% | 109% | 101% | -0.16% | -0.07% | -0.08% | 99% | 101% |
| 3.2.1 | Priority based list with primary MPM, secondary MPM and first few default modes prioritized (method 1 from Samsung): TM intra modes | -0.07%  Note 1 | -0.07% | -0.08% | 104% | 101% | -0.08% | 0.07% | 0.01% | 102% | 101% |
| 3.2.2 | Priority based list with primary MPM, secondary MPM and first few default modes prioritized (method 2 from Huawei): 67 intra modes | -1.48% | -1.29% | -1.28% | 112% | 104% | -0.25% | -0.03% | -0.03% | 102% | 102% |
| 3.2.3 | 6 MPM + 1 mode with 4 bits + 60 modes with 6 bits (disabled JEM macro JVET\_B0051\_NON\_MPM\_MODE, per JEM macro VCEG\_AZ07\_INTRA\_ANG\_MODES), 67 intra modes | -1.25% | -1.26% | -1.29% | 114% | 102% | 0.01% | 0.02% | -0.01% | 100% | 102% |
| 3.3.1 | MPM list construction based on dependency between neighbouring intra modes | -1.26% | -1.22% | -1.23% | 115% | 101% | -0.04% | 0.08% | 0.01% | 101% | 102% |

Note 1: for 3.2.1 the comparison to the VTM uses 35 prediction modes; the others have enabled 67 modes when comparing to the VTM.

Focusing on 3.2.3: It was noted that none of these have substantial gain over the BMS and remarked that 3.2.3 seems like a good solution since it is a straightforward approach. Another participant suggested using 67 intra modes with 3 MPMs and a 6-bit FLC for the remaining modes. Another participant remarked that K0175 reports that 6 MPMs rather than 3 for the VTM with 35 prediction modes has 0.2% AI gain with 13% encoding time increase. It was commented that the 0.2% gain might just be because of more searching (reflected in the encoding time increase) rather than the longer MPM list.

It was suggested that instead of the special treatment of one mode as in 3.2.3, to use truncated binarization of the remaining modes (i.e., four modes use 5 bits and 57 of them use 6 bits). It is expected that this would have the same performance as 3.2.3.

It was initially planned to adopt the truncated binarization approach, otherwise per 3.2.3, pending confirmation of some experiment results (LGE / Huawei / Qualcomm planned to test and provide text). Per section 12.2, a 3 MPM approach was adopted.

Further study whether r0educing the number of MPMs from 6 to 3 would have a significant effect.

**CE3.4 on ‘Cross-component linear model (CCLM)’**

Defined tests

|  |  |  |
| --- | --- | --- |
| **Test #** | **Short Description** | **Doc. #** |
| 4.1.1 | LM + MMLM + MFLM + LM-Angular | JVET-K0082  (Qualcomm) |
| 4.1.2 | MMLM + MFLM + LM-angular |
| 4.1.3 | MNLM: LM + MMLM/MFLM (B, C, E, F) + MMLM/MFLM (A, B, C, D) | JVET-K0073  (Foxconn) |
| 4.1.4 | MNLM: LM + MMLM/MFLM (B, C, E, F) + MMLM/MFLM (A, B, C, D) + MMLM/MFLM (C, D, F, H) |
| 4.1.5 | MNLM: LM + MMLM/MFLM (B, C, E, F) + MMLM/MFLM (A, B, C, D) + MMLM/MFLM (C, D, F, H) + MMLM/MFLM (A, B, E, G) |
| 4.1.6 | LM + MMLM + multi filter LM + extended LM-Angular | JVET-K0092  (LGE) |
| 4.1.7 | LM + MMLM + extended LM-Angular |
| 4.1.8 | LM only (or single model CCLM) | JVET-K0190  (Huawei) |
| 4.1.9 | LM only + CCLM Cb-to-Cr |
| 4.1.10 | LM+MMLM |
| 4.1.11 | LM+MMLM+MFLM |
| 4.2.1 | MDLM | JVET-K0191  (Huawei) |
| 4.2.2 | LM + MDLM |
| 4.2.3 | LM+MDLM+MMLM |
| 4.2.4 | LM+MDLM+MMLM+MFLM |
| 4.3.1 | Inter-color reference prediction | JVET-K0395 (KDDI) |
| 4.3.2 | Adaptive inter-residual prediction with fast RDO |
| 4.4.1 | LM + LM-left + LM-top | JVET-K0241 (MediaTek) |
| 4.4.2 | LM + LM-CbCr |
| 4.4.3 | LM + LM fusion |
| 4.4.4 | LM + LM-left + LM-top + LM-CbCr + LM fusion |

CE3.4: ‘All Intra Main10’

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Test #** | **Description** | **All Intra Main10 - Over VTM1.0** | | | | | **All Intra Main10 - Over BMS1.0** | | | | |
| **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 4.1.1 | LM + MMLM + MFLM + LM-Angular | -1.71% | -11.19% | -12.01% | 158% | 109% | -0.03% | -0.56% | -0.60% | 111% | 102% |
| 4.1.2 | MMLM + MFLM + LM-angular | -1.68% | -9.77% | -10.52% | 149% | 109% | -0.02% | 0.84% | 0.67% | 106% | 102% |
| 4.1.3 | MNLM: LM + MMLM/MFLM (B, C, E, F) + MMLM/MFLM (A, B, C, D) | -1.74% | -9.78% | -10.73% | 102% | 102% | -0.08% | 0.64% | 0.48% | 76% | 100% |
| 4.1.4 | MNLM: LM + MMLM/MFLM (B, C, E, F) + MMLM/MFLM (A, B, C, D) + MMLM/MFLM (C, D, F, H) | -1.74% | -10.07% | -11.03% | 106% | 102% | -0.07% | 0.29% | 0.11% | 79% | 100% |
| 4.1.5 | MNLM: LM + MMLM/MFLM (B, C, E, F) + MMLM/MFLM (A, B, C, D) + MMLM/MFLM (C, D, F, H) + MMLM/MFLM (A, B, E, G) | -1.74% | -10.20% | -11.17% | 110% | 102% | -0.06% | 0.24% | 0.05% | 81% | 100% |
| 4.1.6 | LM + MMLM + multi filter LM + extended LM-Angular | -1.71% | -11.01% | -11.82% | 141% | 106% | -0.03% | -0.30% | -0.23% | 101% | 101% |
| 4.1.7 | LM + MMLM + extended LM-Angular | -1.70% | -10.80% | -11.58% | 122% | 103% | -0.01% | -0.01% | 0.02% | 90% | 100% |
| 4.1.8 | LM only (or single model CCLM) | -1.19% | -9.01% | -8.00% | 110% | 102% | 0.47% | 1.88% | 4.39% | 82% | 99% |
| 4.1.9 | LM only + CCLM Cb-to-Cr | -1.53% | -9.80% | -10.48% | 113% | 104% | 0.14% | 1.09% | 1.42% | 83% | 101% |
| 4.1.10 | LM+MMLM | -1.66% | -10.36% | -11.38% | 122% | 104% | 0.01% | 0.32% | 0.26% | 89% | 100% |
| 4.1.11 | LM+MMLM+MFLM | -1.68% | -10.65% | -11.59% | 142% | 106% | 0.00% | 0.00% | 0.00% | 101% | 102% |
| 4.2.1 | MDLM | -1.20% | -9.50% | -8.54% | 121% | 103% | 0.48% | 1.26% | 3.63% | 88% | 100% |
| 4.2.2 | LM + MDLM | -1.61% | -11.11% | -12.07% | 135% | 106% | 0.07% | -0.33% | -0.23% | 97% | 102% |
| 4.2.3 | LM+MDLM+MMLM | -1.67% | -11.40% | -12.46% | 145% | 106% | 0.02% | -0.76% | -0.84% | 102% | 102% |
| 4.2.4 | LM+MDLM+MMLM+MFLM | -1.70% | -11.65% | -12.68% | 165% | 108% | -0.01% | -1.05% | -1.09% | 114% | 103% |
| 4.3.1 | Inter-color reference prediction | 0.07% | -1.62% | -1.62% | 110% | 100% | 0.08% | -0.23% | -0.27% | 110% | 100% |
| 4.3.2 | Adaptive inter-residual prediction with fast RDO | -0.42% | -2.70% | -3.35% | 128% | 102% | 0.02% | 0.02% | -0.07% | 117% | 99% |
| 4.4.1 | LM + LM-left + LM-top | -1.54% | -10.75% | -11.64% | 121% | 102% | 0.14% | -0.01% | 0.20% | 88% | 98% |
| 4.4.2 | LM + LM-CbCr | -1.50% | -10.16% | -11.23% | 119% | 102% | 0.18% | 0.69% | 0.73% | 87% | 98% |
| 4.4.3 | LM + LM fusion | -1.18% | -11.17% | -11.37% | 118% | 103% | 0.52% | -0.52% | 0.21% | 87% | 101% |
| 4.4.4 | LM + LM-left + LM-top + LM-CbCr + LM fusion | -1.19% | -12.09% | -12.62% | 136% | 103% | 0.53% | -1.59% | -1.09% | 98% | 101% |

BMS has single-model LM and Cb-to-Cr prediction and MMLM and MFLM.

It was remarked that K0336 has a single-model LM and MMLM without Cb-to-Cr and MFLM.

It was commented that some of these perform differently (better) if there is a separate tree for chroma. The CD report has test results for that. It was commented that K0074 reports on 4.1.1 and 4.1.5 with a separate tree.

Focus on 4.1.8 as the simplest, especially considering line buffer and computation requirements (1.19%/9.01%/8.00% for AI, 0.76%/10.39%/9.24% for RA). Decision: Adopt 4.1.8.

Further study is suggested for enhancement beyond that.

Supporting a separate tree is suggested.

**CE3.5 on ‘Multi-reference line intra prediction’**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Short description** | **Doc. #** |
| 5.1.1 | Mode dependent reference line selection | JVET-K0284  (Tencent) |
| 5.1.2 | Reference sample extension for multiline intra prediction |
| 5.2.1 | Multiple reference lines | JVET-K0162  (Technicolor) |
| 5.2.2 | Multiple reference lines + boundary filtering |
| 5.2.3 | Multiple reference lines and not used for top line of CTU | JVET-K0221  (Sony) |
| 5.2.4 | Multiple reference lines with 50:50 weighting |  |
| 5.2.5 | Multiple reference lines with 50:50 weighting (multiple reference lines not used for 4xN and Nx4) |
| 5.3.1 | r1: 6-tap combined filter without reference sample smoothing, r2: bi-linear | JVET-K0166 (ETRI) |
| 5.3.2 | r1: 4-tap filter, r2: bi-linear |
| 5.3.3 | r1: bi-linear, r2: bi-linear |
| 5.4.1 | MRL all block sizes | JVET-K0051 (HHI) |
| 5.4.2 | MRL restricted block sizes (encoder only) |
| 5.4.3 | MRL restricted block sizes (restr. signalling) |
| 5.4.4 | MRL + Mode dependent reference line selection |
| 5.5.1 | Use two extended reference lines | JVET-K0277 (ITRI) |
| 5.5.2 | Use three extended reference lines |

CE3.5: ‘All Intra Main10’ results

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **All Intra Main10 – Over VTM1.0** | | | | | **All Intra Main10 – Over BMS1.0** | | | | |
| **Test#** | **Description** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 5.1.1 | Mode dependent reference line selection | -0.86% | -0.15% | -0.09% | 203% | 99% | -0.88% | -0.42% | -0.54% | 179% | 101% |
| 5.1.2 | Reference sample extension for multiline intra prediction | -0.86% | -0.13% | -0.12% | 209% | 105% | -0.88% | -0.45% | -0.53% | 181% | 102% |
| 5.2.1 | Multiple reference lines | -0.31% | -0.51% | -0.54% | 123% | 105% | -0.21% | -0.21% | -0.24% | 105% | 101% |
| 5.2.2 | Multiple reference lines + boundary filtering | -0.91% | -1.40% | -1.49% | 134% | 109% | -0.97% | -1.29% | -1.43% | 108% | 104% |
| 5.2.3 | Multiple reference lines and not used for top line of CTU | -0.26% | -0.42% | -0.44% | 116% | 108% | -0.18% | -0.20% | -0.21% | 105% | 104% |
| 5.2.4 | Multiple reference lines with 50:50 weighting | -0.11% | -0.39% | -0.48% | 115% | 107% | -0.03% | -0.21% | -0.19% | 104% | 103% |
| 5.2.5 | Multiple reference lines with 50:50 weighting (multiple reference lines not used for 4xN and Nx4) | -0.42% | -0.34% | -0.34% | 113% | 107% | -0.28% | -0.05% | -0.02% | 104% | 103% |
| 5.3.1 | r1: 6-tap combined filter without reference sample smoothing, r2: bi-linear | -0.75% | -0.63% | -0.72% | 119% | 107% | -0.57% | -0.60% | -0.60% | 105% | 102% |
| 5.3.2 | r1: 4-tap filter, r2: bi-linear | -0.76% | -0.64% | -0.68% | 116% | 106% | -0.56% | -0.51% | -0.54% | 104% | 102% |
| 5.3.3 | r1: bi-linear, r2: bi-linear | -0.29% | -0.15% | -0.13% | 110% | 104% | -0.19% | 0.00% | -0.03% | 103% | 101% |
| 5.4.1 | MRL all block sizes | -0.9% | -0.4% | -0.4% | 216% | 100% | -0.5% | -0.2% | -0.3% | 136% | 100% |
| 5.4.2 | MRL restricted block sizes (encoder only) | -0.4% | -0.2% | -0.3% | 145% | 98% | -0.3% | -0.2% | -0.2% | 116% | 99% |
| 5.4.3 | MRL restricted block sizes (restr. signalling) | -0.7% | -0.2% | -0.2% | 147% | 99% | -0.4% | -0.1% | -0.1% | 116% | 100% |
| 5.4.4 | MRL + Mode dependent reference line selection | -0.7% | -0.3% | -0.3% | 147% | 101% | -0.5% | -0.1% | -0.1% | 117% | 100% |
| 5.5.1 | Use two extended reference lines | -0.89% | -0.18% | -0.15% | 176% | 100% | -0.79% | -0.42% | -0.49% | 146% | 99% |
| 5.5.2 | Use three extended reference lines | -1.05% | -0.19% | -0.16% | 208% | 99% | -0.93% | -0.52% | -0.59% | 161% | 100% |

It was commented that multiple reference line usage has a significant complexity impact, both for the encoder and in regard to line buffering for the decoder.

It was commented that some of these have combinations of techniques in them, e.g., 5.2.2 has a boundary filtering aspect.

It was commented that some of these have a problem with screen content.

It was commented that 5.4.4 and 5.2.5 combine well with PDPC.

Non-CE contributions K0482 and K0175 were said to contain reduced-complexity schemes.

No action; further study for minimizing line buffering (e.g., disabling at top of CTUs), minimizing search, checking effect on screen content.

**CE3.6 on ‘Non-linear intra prediction’**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Short description** | **Doc. #** |
| 6.1.1 | Intra prediction using neural networks | JVET-K0266  (HHI) |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **All Intra Main10 – Over VTM1.0** | | | | | **All Intra Main10 – Over BMS1.0** | | | | |
| **Test#** | **Description** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 6.1.1 | Nonlinear weighted intra prediction | −2.81% | −1.42% | −1.46% | 265% | 145% | −2.02% | −1.91% | −1.95% | 145% | 121% |

It was commented that K0196 has a method that is asserted to be simpler and have about 0.5% additional gain.

The complexity impact is said to be a matrix multiply (200×65 for a 32x32 block) and some ROM (~7 Mbytes).

Basically the order of the prediction and inverse transform is swapped. A variation would be to do two inverse transforms.

This is a separate mode; the encoder would not need to use it.

It was commented that this has a significant complexity impact on the latency of the processing.

Further study in CE is encouraged to reduce memory and consider latency impact.

**CE3.7 on ‘Bidirectional intra prediction’**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Short description** | **Doc. #** |
| 7.1.1 | Bi-directional Intra prediction (BDIP) | JVET-K0163  (Technicolor) |
| 7.2.1 | Distance-Weighted Directional Intra Prediction (DWDIP) | JVET-K0045  (Huawei) |
| 7.3.1 | Linear interpolation intra prediction (LIP) based on MPM flag following LIP flag signalling | JVET-K0090  (LGE) |
| 7.3.2 | LIP based on LIP flag following MPM flag signalling | JVET-K0090  (LGE) |

CE3.7: ‘All Intra Main10’ results

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **All Intra Main10 – Over VTM1.0** | | | | | **All Intra Main10 – Over BMS1.0** | | | | |
| **Test#** | **Description** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 7.1.1 | BDIP | -0.26% | -0.29% | -0.30% | 111% | 97% | -0.38% | -0.62% | -0.65% | 102% | 96% |
| 7.2.1 | DWDIP | -0.28% | -0.33% | -0.29% | 106% | 100% | -0.35% | -0.42% | -0.44% | 105% | 99% |
| 7.3.1 | LIP (LIP🡪MPM) | -0.55% | -0.52% | -0.53% | 106% | 102% | -0.67% | -1.01% | -1.04% | 105% | 103% |
| 7.3.2 | LIP (MPM🡪LIP) | -0.57% | -0.53% | -0.51% | 105% | 102% | -0.68% | -1.02% | -1.03% | 105% | 104% |

This adds additional modes.

This has some relationship with the interpolation filtering topic (subtest 2).

There could be some effect from PDPC.

Further study in a CE was suggested.

[JVET-K0043](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3544) CE3: Bilateral reference sample filter (Test 2.7.1) [P. Merkle, H. Schwarz, D. Marpe, T. Wiegand (HHI)]

[JVET-K0045](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3546) CE3: Distance-weighted directional intra-prediction (Test 7.2.1) [A. Filippov, V. Rufitskiy, J. Chen (Huawei)]

[JVET-K0046](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3547) CE3: Wide-angle intra prediction (Test 1.3.1) [J. Lainema (Nokia)]

[JVET-K0049](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3550) CE3: Line-based intra coding mode (Tests 1.4.1, 1.4.2 and 1.4.3) [S. De Luxán Hernández, H. Schwarz, D. Marpe, T. Wiegand (HHI)]

[JVET-K0051](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3552) CE3: Multiple reference line intra prediction (Test 5.4.1, 5.4.2, 5.4.3 and 5.4.4) [B. Bross, H. Schwarz, D. Marpe, T. Wiegand (HHI)]

[JVET-K0055](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3556) CE3: Unequal Weight Planar Prediction (Test 1.5.1) [K. Panusopone, S. Hong, Y. Yu, L. Wang (Arris), A. Segall (Sharp)]

[JVET-K0060](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3561) CE3: Variable number of directional intra modes (Tests 1.1.1 and 1.1.2) [G. Van der Auwera, V. Seregin, A.K. Ramasubramonian, M. Karczewicz (Qualcomm)]

[JVET-K0061](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3562) CE3: Bilateral intra reference sample filter (Test 2.1.1) [G. Van der Auwera, V. Seregin, A.K. Ramasubramonian, M. Karczewicz (Qualcomm)]

[JVET-K0062](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3563) CE3: Intra reference sample interpolation (Tests 2.2.1, 2.3.1, 2.3.2) [G. Van der Auwera, V. Seregin, A.K. Ramasubramonian, M. Karczewicz (Qualcomm)]

[JVET-K0063](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3564) CE3: Simplified PDPC (Test 2.4.1) [G. Van der Auwera, V. Seregin, A. Said, A.K. Ramasubramonian, M. Karczewicz (Qualcomm)]

[JVET-K0066](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3567) CE3: Mode dependent de-ringing filter (Test 2.8.2) [S. Ikonin, J.Chen (Huawei)]

[JVET-K0073](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3575) CE3: Multiple neighbor-based linear model (Tests 4.1.3, 4.1.4, and 4.1.5) [Y.-J. Chang, H.-Y. Jiang (Foxconn)]

[JVET-K0081](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3584) CE3: Two MPM modes and shape dependency (Test 3.1.1) [A.K. Ramasubramonian, G. Van der Auwera, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-K0082](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3585) CE3: LM-Angular prediction (Tests 4.1.1 and 4.1.2) [A.K. Ramasubramonian, G. Van der Auwera, V. Seregin, M. Karczewicz (Qualcomm)]

[JVET-K0087](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3590) CE3-3.3.1: MPM list construction based on dependency between neighboring intra modes [L. Li, J. Heo, J. Choi, S. Yoo, J. Lim (LGE)]

[JVET-K0090](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3593) CE3: Linear interpolation intra prediction (Tests 7.3.1, 7.3.2) [J. Heo, J. Choi, S. Yoo, L. Li, J. Lim (LGE)]

[JVET-K0092](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3595) CE3: Extended LM angular prediction (Test 4.1.6 and 4.1.7) [J. Choi, J. Heo, S. Yoo, L. Li, J. Lim (LGE)]

[JVET-K0097](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3600) CE3-2.3.3 and CE3-2.3.4: Interpolation filter selection regarding intra mode and block size [S. Yoo, J. Heo, J. Choi, L. Li, J. Lim (LGE)]

[JVET-K0162](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3670) CE3.5: Multiple Reference Intra Prediction (tests 5.2.1 and 5.2.2) [G. Rath, F. Urban, F. Racapé (Technicolor)]

[JVET-K0163](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3671) CE3.7: Bi-Directional Intra Prediction (test 7.1.1) [G. Rath, F. Urban, F. Racapé (Technicolor)]

[JVET-K0165](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3673) CE3: Combined filter (Test 2.6.1) [J. Lee, H. Lee, S.-C. Lim, J. Kang, H. Y. Kim (ETRI)]

[JVET-K0166](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3674) CE3: Multi-line based intra prediction (Test 5.3.1, 5.3.2, 5.3.3) [J. Lee, H. Lee, S.-C. Lim, J. Kang, H. Y. Kim (ETRI)]

[JVET-K0179](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3687) CE3: Multiple 4-tap interpolation filter (CE3 Test 2.10.1) [N. Choi, M. W. Park, C. Kim (Samsung)]

[JVET-K0180](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3688) CE3: multi-combined intra prediction (MIP, CE3 Test 2.11.1) [N. Choi, M. W. Park, C. Kim (Samsung)]

[JVET-K0181](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3690) CE3: Priority based list with primary MPM, secondary MPM and first few default modes prioritized (CE3 Test 3.2.1) [N. Choi, Y. Piao, C. Kim (Samsung)]

[JVET-K0190](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3699) CE3: Tests of cross-component linear model in BMS1.0 (Test 4.1.8, 4.1.9, 4.1.10, 4.1.11) [X. Ma, H. Yang, J. Chen (Huawei)]

[JVET-K0191](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3700) CE3: Multi-directional LM (Test 4.2.1, 4.2.2, 4.2.3, 4.2.4) [X. Ma, H. Yang, J. Chen (Huawei)]

[JVET-K0211](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3720) CE3: DC mode without divisions and modifications to intra filtering (Tests 1.2.1, 2.2.2 and 2.5.1) [V. Drugeon (Panasonic)]

[JVET-K0240](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3750) CE3.2.9.1: Intra boundary filters [Z.-Y. Lin, T.-D. Chuang, C.-Y. Chen, C.-W. Hsu, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-K0241](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3751) CE3.4.4: Additional LM-based modes for intra chroma prediction [C.-M. Tsai, C.-W. Hsu, C.-Y. Chen, T.-D. Chuang, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-K0266](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3777) CE3: Non-linear weighted intra prediction (Test 6.1.1) [P. Merkle, J. Pfaff, P. Helle, R. Rischke, M. Schäfer, B. Stallenberger, H. Schwarz, D. Marpe, T. Wiegand (Fraunhofer HHI)]

[JVET-K0277](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3790) CE3: Number of extended reference line for intra prediction (Test 5.5.1 and 5.5.2) [P.-H. Lin, C.-H. Yao, C.-C. Lin, S.-P. Wang, P. Sung, C.-L. Lin (ITRI)] [late]

[JVET-K0284](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3797) CE3: Mode dependent reference line selection (Test 5.1.1 and 5.1.2) [L. Zhao, X. Zhao, X. Li, S. Liu (Tencent)]

[JVET-K0365](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3883) CE3.2.2: Intra mode signalling with priority based MPM and non-MPM list construction [A.M. Kotra, Z. Zhao, J. Chen (Huawei)]

[JVET-K0368](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3886) CE 3.2.3: Intra mode signalling without non-MPM list [A.M. Kotra, B. Wang, Z. Zhao, J. Chen (Huawei)]

[JVET-K0395](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3915) CE3: Inter-color reference prediction (CE3-4.3.1) [Kei Kawamura, Yoshitaka Kidani, Sei Naito(KDDI)] [late]

[JVET-K0396](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3916) CE3: Adaptive inter-residual prediction (CE3-4.3.2) [Kei Kawamura, Yoshitaka Kidani, Sei Naito (KDDI)] [late]

## CE4: Inter prediction and motion vector coding (35)

Contributions in this category were discussed Wednesday 11 July in Track B 1900–2100 (chaired by JRO), continued Thursday 12th 1300-2100 and Friday 13th morning.

[JVET-K0024](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3852) CE4: Summary report on inter prediction and motion vector coding [H. Yang, S. Liu, K. Zhang]

This contribution provides a summary report of Core Experiment 4 on inter prediction and motion vector coding. CE4 comprises seven categories, 1) affine motion compensation, 2) merge mode enhancement, 3) motion vector coding, 4) generalized bi-prediction, 5) reference picture boundary padding, 6) local illumination compensation, and 7) interpolation filter improvement. Test results against VTM anchor are provided to show the coding efficiency and complexity trade-off of each tool. Test results against BMS anchor are also provided to show the interaction with BMS coding tools. Crosschecking results are integrated in this contribution.

**CE4.1: Affine MC**

**Aspects of affine motion compensation stage** (discussed Wed 11th 1900-2100, chaired by JRO)

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| AFFINE | BMS AFFINE as benchmark |  |
| 4.1.2.a | Affine flexing using sub-block size of 4x4 | JVET-K0047 |
| 4.1.2.b | Affine flexing using sub-block size of 8x8 | JVET-K0047 |
| 4.1.5.d | EIF on top of BMS Affine | JVET-K0185 |

Random access results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| AFFINE | -2.99% | -2.19% | -2.21% | 137% | 112% | -1.92% | -1.34% | -1.36% | 108% | 102% |
| 4.1.2.a | -3.31% | -2.38% | -2.37% | 131% | 113% | -2.13% | -1.46% | -1.45% | 105% | 100% |
| 4.1.2.b | -2.90% | -1.97% | -1.93% | 126% | 107% | -1.81% | -1.10% | -1.09% | 105% | 98% |
| 4.1.5.d | -3.23% | -2.41% | -2.52% | 132% | 102% | -1.85% | -1.42% | -1.44% | 109% | 99% |

Low delay B results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| AFFINE | -2.06% | -1.33% | -1.52% | 168% | 108% | -1.98% | -1.52% | -1.70% | 125% | 104% |
| 4.1.2.a | -2.27% | -1.48% | -1.48% | 161% | 111% | -2.18% | -1.79% | -1.61% | 122% | 102% |
| 4.1.2.b | -1.94% | -1.08% | -1.17% | 152% | 111% | -1.94% | -1.40% | -1.65% | 119% | 102% |
| 4.1.5.d | -2.47% | -1.61% | -1.94% | 158% | 102% | -1.89% | -1.58% | -1.64% | 122% | 100% |

Current affine inter prediction in BMS firstly divides a coding unit to sub-blocks, and then calculates motion vector for each sub-block based on affine motion model. When doing MC for each sub-block, translational motion is assumed. This was a reasonable trade-off when affine MC was proposed. Because “true” affine MC requires per-pixel operation and existing MC operation is too expensive.

The two techniques here try to perform “true” affine MC by low-complexity MC operation.

Flexing in test 4.1.2 performs additional motion compensation horizontally for each line of a sub-block and vertically for each columns of a sub-block, separately, trying to approximate per-pixel affine MC. Whether horizontal or vertical MC is performed first depends on the whether the motion of a CU is more rotational featured and more zooming featured. Existing 8-tap DCTIF is used for interpolation. A look-up table implementation is provided in the implementation.

EIF in test 4.1.5.b performs per-pixel affine MC using a two-stage interpolation, bi-linear + 3-tap filter. It is asserted the new interpolation mechanism is much simpler than 8-tap DCTIF and thus could do per-pixel MC with low cost.

Aspects of affine motion compensation stage (discussed Wed 1900-2100)

4.1.1, 4.1.2 and 4.1.5 are touching the motion compensation stage

- 4.1.1b is a simplification of BMS affine, but it has 0.5% worse performance than BMS affine;

- 4.1.2 (affine flexing) am 4.1.5 (EIF) are doing more precise motion comp.

Affine flexing adds another step after regular block-based motion comp. This uses the regular 8-tap filters, but applies padding outside of the prediction block, such that no additional memory accesses are required. Entire lines are shifted horizontally, and then the result is shifted column-wise. This requires in worst case 8+8 additional multiplications per sample. The gain is -0.3% for 4x4 subblocks, and some small loss is observed when affine is performed on 8x8 subblocks. This additional complexity is not a good tradeoff versus the relative small gain.

EIF is pixel-based motion comp, which used a different motion compensation for affine blocks, by applying bilinear interpolation followed by a sharpening filter. The gain is approx. -0.2%. It is claimed that the method is less complex than the conventional motion comp (in terms of number of operations), however it probably is also less regular, as it consists of three subsequent stages: a) determining individual shift positions of samples b) performing bilinear interpolation c) sharpening filter. Likewise, the specification is likely more complicated.

From the current results, a solution using identical MC operation for affine cases as in other modes appears to be the preferred solution. However, compared to the BMS affine, it would be simpler to use fixed subblock size of 4x4 (as per CE 4.1.1a), i.e. remove the adaptive subblock size from BMS.

Decision: Adopt JVET-K0184 to BMS; (CE4.1.1a 4x4 fixed subblock size). It was later decided in the JVET Sunday plenary that the affine tools (with modifications as adopted for BMS initially) will be moved to VTM.

It is also mentioned that pixel-based MC may have subjective advantage, but it currently not evident whether this applies to relevant application cases.

Was revisited later. Draft text was provided and reviewed Tue 17 1545.

[JVET-K0565](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4095) Draft text for affine motion compensation [H. Yang, H. Chen, Y. Zhao, J. Chen (Huawei)] [late]

Suggestions made during discussion:

* High level enabling flag at SPS as already in software is missing (PPS or slice needs further consideration)
* Put into SPS 2 flags (1 for affine enabling, 1 for 6-parameter enabling)
* All affine operations and syntax should be disabled by the first flag
* All affine operations and syntax related to 6-parameter should be disabled by the second flag

It is generally noted that the text is appropriate and complete, but needs more detailed investigation in context of integration by editors.

For editors of draft text:

Some alignment of subblock syntax with ATMVP would be necessary

An implication of adopting affine and ATMVP requires also transferring some additional elements from HEVC, such as merge, MV prediction, MV difference coding

What is further needed for affine, is the 1/16 sample precision interpolation filters. Otherwise, the MC process of HEVC can be retained for subblocks. Rounding to 1/4 pel also needs to be described

8x8 MV compression requires action. Storage to be done with high (1/16 pel) precision. Note: It should be further studied if high precision is necessary.

CE4.1 continuation Track B Thursday 12th 1300-1500 (chaired by JRO)

**Aspects of affine motion vector prediction**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| AFFINE | BMS AFFINE as bench mark |  |
| 4.1.3.a | Affine MVP construction | JVET-K0337 |
| 4.1.4.b | Affine MVP construction with added spatial candidates | JVET-K0218 |
| 4.1.5.a | Affine MVP construction | JVET-K0185 |
| 4.1.6 | Affine MVP construction | JVET-K0244 |
| 4.1.7.a | Up to two affine candidate for affine inter | JVET-K0094 |

Random access results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| AFFINE | -2.99% | -2.19% | -2.21% | 137% | 112% | -1.92% | -1.34% | -1.36% | 108% | 102% |
| 4.1.3.a | -3.22% | -2.40% | -2.43% | 134% | 110% | -2.12% | -1.54% | -1.56% | 112% | 107% |
| 4.1.4.a | -3.90% | -2.95% | -2.98% | 148% | 115% | -2.72% | -2.01% | -2.02% | 113% | 107% |
| 4.1.4.b | -3.89% | -2.93% | -2.95% | 148% | 115% | -2.71% | -2.02% | -2.05% | 113% | 106% |
| 4.1.5.a | -3.24% | -2.41% | -2.47% | 138% | 110% | -2.10% | -1.54% | -1.53% | 110% | 101% |
| 4.1.6 | -3.21% | -2.38% | -2.40% | 132% | 107% | -2.10% | -1.50% | -1.56% | 111% | 88% |
| 4.1.7.a | -3.23% | -2.37% | -2.43% | 142% | 114% | -2.10% | -1.52% | -1.56% | 111% | 102% |

Low delay B results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| AFFINE | -2.06% | -1.33% | -1.52% | 168% | 108% | -1.98% | -1.52% | -1.70% | 125% | 104% |
| 4.1.3.a | -2.08% | -1.46% | -1.62% | 162% | 92% | -2.08% | -1.62% | -1.92% | 127% | 108% |
| 4.1.4.a | -2.38% | -1.61% | -1.59% | 213% | 107% | -2.39% | -1.89% | -2.07% | 142% | 109% |
| 4.1.4.b | -2.37% | -1.56% | -1.65% | 213% | 107% | -2.39% | -1.98% | -2.03% | 143% | 110% |
| 4.1.5.a | -2.13% | -1.26% | -1.59% | 169% | 107% | -2.07% | -1.69% | -1.57% | 126% | 103% |
| 4.1.6 | -2.10% | -1.43% | -1.48% | 159% | 104% | -2.04% | -1.57% | -1.74% | 128% | 108% |
| 4.1.7.a | -2.12% | -1.46% | -1.47% | 181% | 115% | -2.09% | -1.87% | -1.89% | 129% | 104% |

In BMS affine inter mode, two motion vectors at the top-left and top-right corners of a coding unit is encoded. The predictor for the motion vector pair is constructed using motion vectors of neighboring blocks.

In the five tests here, a new types of candidate predictor is proposed. This candidate inherits the affine model from neighboring blocks and use that model to derive the motion vector at the control points of the coding block. The derived motion vector pair is then used as a predictor for CPMVs of the coding block.

Tests here seems converged to constructing the MVP list with two types of predictors, inherited predictor and constructed predictor. And the number of predictors is 2.

The differences among the tests are,

* The position from where a candidate is derived
* The order in which the MVP list is constructed
* The number of candidates of a particular type
* Whether MV scaling is allow when deriving a candidate
* Whether a top-right MV derived from top-left and bottom-left MV could be used as the predictor for the top-right CPMV

Proponents are requested to provide an analysis about the number of operations, MV comparisons, memory usage, additional storage, etc. for the list construction and the inheritance, also in comparison with BMS affine.

Follow-up Sat. 14th morning:

Analysis was done (will be registered as new doc – add number

4.1.6 and 4.1.7 have higher complexity than BMS

4.1.8 has same complexity as BMS affine

4.1.3 has least complexity (its worst case is for inherited candidates, but still significantly lower than BMS – no scaling, no mult., no div.)

4.1.4 is same as 4.1.3 with more candidates (here analysed for 4-parameter model, for which case we don’t have CE results), complexity vs. 4.1.3 is higher by a factor of 1.4

4.1.5 is still roughly by a factor 3 less complex than BMS affine, but has some scaling/mul./div.

From this, the best choice in the domain of 4-parameter model prediction is the solution of 4.1.3a. It has the same compression performance as 4.1.5, but is less complex.

Decision (BMS): Adopt JVET-K0337 (4.1.3a, affine MVP list construction)

It is noted that the decoder run time increases. This is asserted to be due to the fact that affine prediction mode is now selected more frequently.

It is noted that inheritance requires additional buffering of affine model, including PU size. This is also the case for merge in BMS affine, whereas prediction in BMS affine currently does not use inheritance.

Performance of 4.1.4 is better as it also uses 6 parameter model (see other proposals below).

**Aspects of MVD coding**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| AFFINE | BMS AFFINE as bench mark |  |
| 4.1.3.b | 4.1.3.a + MVD prediction | JVET-K0337 |
| 4.1.5.b | 4.1.5.a + MVD coding | JVET-K0185 |

Random access results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| AFFINE | -2.99% | -2.19% | -2.21% | 137% | 112% | -1.92% | -1.34% | -1.36% | 108% | 102% |
| 4.1.3.a | -3.22% | -2.40% | -2.43% | 134% | 110% | -2.12% | -1.54% | -1.56% | 112% | 107% |
| 4.1.3.b | -3.42% | -2.55% | -2.61% | 136% | 112% | -2.24% | -1.61% | -1.67% | 111% | 108% |
| 4.1.5.a | -3.24% | -2.41% | -2.47% | 138% | 110% | -2.10% | -1.54% | -1.53% | 110% | 101% |
| 4.1.5.b | -3.23% | -2.41% | -2.43% | 139% | 110% | -2.10% | -1.52% | -1.53% | 110% | 101% |

Low delay B results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| AFFINE | -2.06% | -1.33% | -1.52% | 168% | 108% | -1.98% | -1.52% | -1.70% | 125% | 104% |
| 4.1.3.a | -2.08% | -1.46% | -1.62% | 162% | 92% | -2.08% | -1.62% | -1.92% | 127% | 108% |
| 4.1.3.b | -2.23% | -1.50% | -1.60% | 164% | 99% | -2.21% | -1.80% | -1.84% | 127% | 109% |
| 4.1.5.a | -2.13% | -1.26% | -1.59% | 169% | 107% | -2.07% | -1.69% | -1.57% | 126% | 103% |
| 4.1.5.b | -2.11% | -1.40% | -1.51% | 170% | 107% | -2.07% | -1.72% | -1.86% | 126% | 104% |

Test 4.1.3.b tries to exploit the similarity of CPMVs of a coding block, refining the MV predictor of a CPMV by the MVD of the other one.

Test 4.1.5.b tries to save the bits for coding two MVDs by explicitly signaling whether the pair of MVDs in a prediction direction is skipped or not.

Decision: Adopt Test 4.1.3b to BMS affine. Scheme as shown in equations below.

**Aspects of Alternative motion models**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| AFFINE | BMS AFFINE as bench mark |  |
| 4.1.3.c | 4.1.3.b + CU level 4-para/6-para switching | JVET-K0337 |
| 4.1.3.d | 4.1.3.c + slice level 4-para/6-para switching | JVET-K0337 |
| 4.1.5.c | 4.1.5.b + 4 para/6-para switching | JVET-K0185 |
| 4.1.7.b | Four and six parameter model | JVET-K0094 |
| 4.1.7.c | Four and six parameter model with the conditional affine parameter signaling | JVET-K0094 |
| 4.1.8.a | Alternative 3-para (scaling) motion model | JVET-K0124 |
| 4.1.8.b | Alternative 3-para (rotation) motion model | JVET-K0124 |
| 4.1.8.c | 4.1.8.a + 4.1.8.b | JVET-K0124 |

Random access results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| AFFINE | -2.99% | -2.19% | -2.21% | 137% | 112% | -1.92% | -1.34% | -1.36% | 108% | 102% |
| 4.1.3.b | -3.42% | -2.55% | -2.61% | 136% | 112% | -2.24% | -1.61% | -1.67% | 111% | 108% |
| 4.1.3.c | -3.90% | -2.94% | -2.92% | 160% | 113% | -2.72% | -1.96% | -1.97% | 118% | 107% |
| 4.1.3.d | -3.88% | -2.95% | -2.92% | 146% | 113% | -2.69% | -1.98% | -2.00% | 112% | 106% |
| 4.1.5.b | -3.23% | -2.41% | -2.43% | 139% | 110% | -2.10% | -1.52% | -1.53% | 110% | 101% |
| 4.1.5.c | -3.83% | -2.86% | -2.89% | 144% | 111% | -2.65% | -1.91% | -1.94% | 111% | 102% |
| 4.1.7.b | -3.53% | -2.56% | -2.62% | 200% | 116% | -2.40% | -1.69% | -1.70% | 128% | 102% |
| 4.1.7.c | -3.48% | -2.54% | -2.57% | 159% | 116% | -2.31% | -1.67% | -1.65% | 114% | 102% |
| 4.1.8.a | -3.10% | -2.27% | -2.29% | 181% | 114% | -1.99% | -1.40% | -1.42% | 118% | 101% |
| 4.1.8.b | -3.07% | -2.26% | -2.28% | 181% | 112% | -1.97% | -1.41% | -1.44% | 118% | 101% |
| 4.1.8.c | -3.19% | -2.32% | -2.31% | 221% | 114% | -2.06% | -1.44% | -1.50% | 127% | 102% |

Low delay B results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| AFFINE | -2.06% | -1.33% | -1.52% | 168% | 108% | -1.98% | -1.52% | -1.70% | 125% | 104% |
| 4.1.3.b | -2.23% | -1.50% | -1.60% | 164% | 99% | -2.21% | -1.80% | -1.84% | 127% | 109% |
| 4.1.3.c | -2.35% | -1.53% | -1.82% | 210% | 103% | -2.37% | -1.91% | -1.89% | 145% | 107% |
| 4.1.3.d | -2.37% | -1.53% | -1.53% | 199% | 105% | -2.36% | -1.92% | -1.74% | 139% | 109% |
| 4.1.5.b | -2.11% | -1.40% | -1.51% | 170% | 107% | -2.07% | -1.72% | -1.86% | 126% | 104% |
| 4.1.5.c | -2.36% | -1.54% | -1.47% | 174% | 108% | -2.31% | -1.69% | -2.13% | 128% | 104% |
| 4.1.7.b | -2.33% | -1.49% | -1.60% | 297% | 116% | -2.35% | -1.85% | -1.90% | 170% | 105% |
| 4.1.7.c | -2.28% | -1.43% | -1.42% | 218% | 117% | -2.27% | -1.61% | -1.75% | 141% | 106% |
| 4.1.8.a | -2.12% | -1.31% | -1.39% | 252% | 113% | -2.10% | -1.56% | -1.68% | 152% | 104% |
| 4.1.8.b | -2.12% | -1.28% | -1.34% | 252% | 113% | -2.08% | -1.72% | -1.64% | 152% | 104% |
| 4.1.8.c | -2.15% | -1.35% | -1.55% | 324% | 114% | -2.14% | -1.71% | -1.77% | 176% | 104% |

Most of tests here add an additional 6-param affine model. 6-param affine model requires coding 3 CPMVs, compared with coding 2CPMVs for 4-param affine model in BMS affine. Adaptive selection at CU level or slice level are proposed.

Test 4.1.8 propose two types of 3-param model models, scaling model and rotation model. The 3-param model is represented by 1.5 motion vector, top-left motion vector and x or y component of the top-right motion vector.

Generally, additional motion model introduces operations handling the additional motion parameters, e.g. motion estimation, motion vector prediction, motion vector difference coding, motion parameter storage, etc.

Switchable 4/6 parameter model provides approx. 0.5-0.6% gain. The switchable 3/3/4 approach does not provide comparable gain (up to 0.2%).

Proponents are requested to provide an analysis about the number of operations, memory usage, etc. for the list construction and the inheritance, also in comparison with BMS affine.

Analysis was shown Sat. 14th 9-10.

The solution of 4.1.3c has least complexity (still significantly less complex than the current BMS with 4 parameters, no scal. No mul. No div.). 4.1.4 is based on it with more candidates, by factor 1.4 more complex. 4.1.5 is still less complex than BMS, but higher complex than 4.1.3/4.1.4. Though it has less shift and additions, it requires scaling/mul/div.

After all, 4.1.3. is asserted to be the simplest solution. For 6 parameters it is 1.5x complex as for the case of 4 parameters, and still significantly less than current BMS with 4 parameters

Decision (BMS): Adopt JVET-K0337 (4.1.3c, 4/6 parameter model, no slice level switch).

Decision (BMS/enc): Adopt JVET-K0185 fast encoder from 4.1.5c in combination with 4.1.3c

Further study of slice level switch (as per 4.1.3d, 4.1.4.x) in CE

The harmonization with affine merge should further be studied

Line buffer reduction should further be studied

The signalling and coding mechanisms of 4.1.7c should also further be studied in combination with the new BMS.

The possibility of further switching between 3/3/4/6 models (as from the results of 3.18c might give additional gain) should also further be studied. However, such a modification should only be made when it is not penalized by increased encoder run time, and the methods of MV prediction and merge should be harmonized.

**Aspects of Affine model seed storage**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| AFFINE | BMS AFFINE as bench mark |  |
| 4.1.4.a | Affine model seed storage (based on 4.1.3.d) | JVET-K0337 |

Random access results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| AFFINE | -2.99% | -2.19% | -2.21% | 137% | 112% | -1.92% | -1.34% | -1.36% | 108% | 102% |
| 4.1.3.d | -3.88% | -2.95% | -2.92% | 146% | 113% | -2.69% | -1.98% | -2.00% | 112% | 106% |
| 4.1.4.a | -3.90% | -2.95% | -2.98% | 148% | 115% | -2.72% | -2.01% | -2.02% | 113% | 107% |

Low delay B results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| AFFINE | -2.06% | -1.33% | -1.52% | 168% | 108% | -1.98% | -1.52% | -1.70% | 125% | 104% |
| 4.1.3.d | -2.37% | -1.53% | -1.53% | 199% | 105% | -2.36% | -1.92% | -1.74% | 139% | 109% |
| 4.1.4.a | -2.38% | -1.61% | -1.59% | 213% | 107% | -2.39% | -1.89% | -2.07% | 142% | 109% |

In BMS, the affine model seeds are stored in the top-left, top-right and bottom-left 4x4 sub-blocks in the considered CU. In the proposed solution, the affine model seeds are stored separately as a motion information associated to the whole CU.

No action. No impact on performance, and the optimization in terms of storage saving might better be done after stabilization of affine approaches.

**CE4.2: Merge mode enhancements** (discussed in Track B Thu. 1500-1900, chaired by JRO)

**Proposals on CU based candidate – Long distance spatial candidates**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| 4.2.1 | Non-adjacent spatial merge candidates (number of candidates is 5) | JVET-K0228 |
| 4.2.3.a | Non-adjacent spatial merge candidate candidates (number of candidates is 5) | JVET-K0339 |
| 4.2.3.b | 4.2.3.a + Enlarge merge candidate list size to 10 | JVET-K0339 |
| 4.2.3.c | 4.2.3.a + Enlarge merge candidate list size to 15 | JVET-K0339 |
| 4.2.3.d | 4.2.3.c + Merge candidate list reordering | JVET-K0339 |
| 4.2.8.b | Adding middle spatial and multiple temporal candidates (number of candidates is 10) | JVET-K0245 |
| 4.2.13.a | Additional merge candidates (number of candidates is 10) | JVET-K0286 |
| 4.2.13.b | Further reduce candidate number (number of candidates is 8) | JVET-K0286 |
| 4.2.15.a | Extended spatial positions from 6 to 27 (number of candidates is 11) | JVET-K0198 |

Random access results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 4.2.1 | -0.45% | -0.45% | -0.44% | 100% | 100% | -0.22% | -0.21% | -0.20% | 98% | 97% |
| 4.2.3.a | -0.66% | -0.55% | -0.49% | 100% | 101% | -0.35% | -0.27% | -0.31% | 100% | 101% |
| 4.2.3.b | -0.88% | -0.81% | -0.77% | 105% | 101% | -0.62% | -0.54% | -0.57% | 102% | 101% |
| 4.2.3.c | -0.89% | -0.86% | -0.85% | 108% | 102% | -0.73% | -0.66% | -0.68% | 105% | 102% |
| 4.2.3.d | -0.88% | -0.85% | -0.85% | 109% | 103% | -0.76% | -0.71% | -0.78% | 105% | 103% |
| 4.2.8.b | -0.29% | -0.29% | -0.23% | 104% | 99% | -0.09% | -0.10% | -0.10% | 103% | 100% |
| 4.2.13.a | -0.88% | -0.89% | -0.88% | 108% | 104% | -0.68% | -0.73% | -0.71% | 100% | 100% |
| 4.2.13.b | -0.88% | -0.86% | -0.85% | 107% | 104% | -0.61% | -0.59% | -0.61% | 100% | 101% |
| 4.2.15.a | -1.14% | -1.12% | -1.17% | 109% | 101% | -1.22% | -1.21% | -1.30% | 106% | 102% |

Low delay B results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 4.2.1 | -0.12% | 0.07% | 0.11% | 100% | 99% | -0.09% | -0.06% | 0.07% | 99% | 96% |
| 4.2.3.a | -0.36% | 0.06% | 0.18% | 102% | 105% | -0.25% | 0.08% | 0.25% | 99% | 99% |
| 4.2.3.b | -0.49% | -0.18% | -0.25% | 107% | 103% | -0.37% | -0.48% | -0.28% | 102% | 102% |
| 4.2.3.c | -0.49% | -0.24% | -0.14% | 112% | 103% | -0.44% | -0.24% | -0.08% | 103% | 102% |
| 4.2.3.d | -0.49% | -0.24% | -0.14% | 111% | 104% | -0.44% | -0.24% | -0.08% | 103% | 100% |
| 4.2.8.b | -0.39% | -0.35% | -0.30% | 104% | 99% | -0.10% | -0.02% | 0.29% | 103% | 99% |
| 4.2.13.a | -0.51% | -0.27% | -0.22% | 108% | 103% | -0.41% | -0.22% | -0.09% | 100% | 97% |
| 4.2.13.b | -0.52% | -0.23% | -0.37% | 107% | 105% | -0.39% | -0.32% | -0.13% | 100% | 99% |
| 4.2.15.a | -0.84% | -0.65% | -0.70% | 111% | 104% | -0.89% | -0.83% | -0.67% | 103% | 100% |

Tests here add additional spatial merge candidates. Differences are, 1) the position from where candidates are derived, 2) the order of searching more candidates, 3) the number of candidates.

Possible explanations for decoding time increase are, 1) constructing longer list, 2) searching for valid motion info within a search range.

From test 4.2.3 and 4.2.13, it seems the coding gain saturate as the number of candidates increase. 5 looks like the magic number.

Test 4.2.15a shows -1.14% coding gain which is the highest coding gain in this category. On the other hand, the number of RD checks should be taken into account when making a comparison since it may have an impact on the coding gain.

Generally speaking, techniques in this category requires accessing more motion data in the coded area, additional buffer including the line number may be required.

From the data given, it is obvious that the merge performance can be significantly increased when the number of candidates checked is highly increased, the more the better. A reasonable assessment of complexity impact is missing.

Proponents are requested to provide an analysis about the number of operations, comparisons, memory accesses and additional storage needs, etc. for the list construction, also in comparison with current method from VTM. Not finished yet Sat. morning, as there is no consensus yet among proponents. Should be further continued in BoG (X. Li). See further notes under BoG XXXX.

**Proposals on CU based candidate – Extended spatial candidates**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| 4.2.9.a | UMVE candidates as additional merge candidates with additional information signalled | JVET-K0115 |
| 4.2.9.b1 | UMVE candidates as independent merge candidates with additional information signalled, number of candidates is 120 | JVET-K0115 |
| 4.2.9.b2 | UMVE candidates as independent merge candidates with additional information signalled, number of candidates is 384 | JVET-K0115 |
| 4.2.9.b3 | UMVE candidates as independent merge candidates with additional information signalled, number of candidates is 32 | JVET-K0115 |
| 4.2.15.b | Merge index 0 with MV offsets (number of candidates is 13) | JVET-K0198 |

Random access results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 4.2.9.a | -1.06% | -1.13% | -1.13% | 142% | 118% | -0.36% | -0.34% | -0.31% | 124% | 108% |
| 4.2.9.b1 | -1.93% | -2.05% | -2.10% | 127% | 110% | -0.56% | -0.59% | -0.66% | 115% | 105% |
| 4.2.9.b2 | -1.53% | -1.61% | -1.63% | 144% | 113% | -0.63% | -0.60% | -0.63% | 126% | 107% |
| 4.2.9.b3 | -1.67% | -1.78% | -1.83% | 121% | 102% | -0.48% | -0.50% | -0.57% | 114% | 102% |
| 4.2.15.b | -1.04% | -1.09% | -1.15% | 126% | 101% | -0.58% | -0.68% | -0.80% | 116% | 101% |

Low delay B results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 4.2.9.a | -0.21% | -0.08% | -0.26% | 155% | 118% | 0.02% | 0.06% | 0.01% | 125% | 109% |
| 4.2.9.b1 | -0.36% | -0.52% | -0.49% | 129% | 108% | -0.09% | 0.03% | 0.17% | 112% | 106% |
| 4.2.9.b2 | -0.22% | -0.36% | -0.38% | 153% | 117% | -0.02% | -0.13% | 0.06% | 124% | 108% |
| 4.2.9.b3 | -0.24% | -0.31% | -0.44% | 122% | 103% | 0.01% | 0.06% | 0.06% | 111% | 101% |
| 4.2.15.b | -0.42% | -0.48% | -0.70% | 125% | 102% | -0.38% | -0.60% | -0.48% | 110% | 101% |

Test 4.2.9 constructs a list of four candidates by reusing existing merge candidates, and then extend the merge candidates in two aspects, 1) prediction direction specification, including constructing mirrored motion vector, 2) MV offsets around each starting point, in cross pattern and non-linearly spaced. Additional syntax for indicating the prediction direction, starting point selection, and MV offset are signalled. MV offset is represented by distance IDX and Direction IDX, instead of the regular x/y component representation. The number of candidates for b1, b2 and b3 is 120, 384 and 32, separately.

All the new candidates in test 4.2.9 is encapsulated as UMVE. It is signalled by using an index of a regular merge list (4.2.9.a) separately by a flag after skip/merge flag. Note that the coding gain gap of different signaling method is big.

Test 4.2.15.b derives 8 neighboring motion vectors around the first merge candidates. All the 8 candidates are horizontally and vertically 1-pel away from the position the first merge candidate points to. The 8 candidates are appended to the end of the current merge list.

Generally speaking, these proposals are kind of “hybrid” new mode between merge and MV prediction, could be interpreted as either adding more merge candidates around existing ones, or adding a difference on the selected merge candidate. Unlike MV prediction, x and y differences are coded jointly, e.g. by only allowing additional horizontal/vertical displacements (in 4.2.9).

Furthermore, all proposals are benefitting from using symmetric coding (using inverted motion vector and opposite POC for the L1 picture). This is explaining why the gain in RA is significantly higher than in LDB. There are other proposals (e.g. K0188) which apply such an approach also for MVP.

Gain in BMS is significantly lower than in VTM. This indicates interdependency with other tools. Further study necessary, how much gain come from symmetric coding, and combination with other MV coding tools.

**Proposals on CU based candidate – Combined/Split merge candidates**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| 4.2.2 | Restricted merge | JVET-K0279 |
| 4.2.8.c | Adding pairwise-average candidates and removing HEVC combined candidates (number of candidates is 10) | JVET-K0245 |
| 4.2.15.c | Combined average merge candidates | JVET-K0198 |

Random access results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 4.2.2 | -0.21% | -0.32% | -0.32% | 139% | 102% | -0.20% | -0.28% | -0.29% | 120% | 100% |
| 4.2.8.c | -0.41% | -0.37% | -0.38% | 103% | 100% | -0.51% | -0.37% | -0.36% | 103% | 101% |
| 4.2.15.c | -0.31% | -0.22% | -0.24% | 102% | 100% | -0.25% | -0.12% | -0.14% | 102% | 100% |

Low delay B results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 4.2.2 | 0.01% | 0.02% | -0.03% | 145% | 104% | -0.10% | -0.19% | -0.03% | 131% | 99% |
| 4.2.8.c | -0.13% | 0.03% | 0.11% | 102% | 100% | -0.27% | -0.21% | -0.19% | 102% | 101% |
| 4.2.15.c | 0.00% | 0.00% | 0.00% | 102% | 100% | 0.00% | 0.00% | 0.00% | 102% | 96% |

Test 4.2.2 construct two additional merge list, one with uni-prediction candidate in L0 direction, the other with uni-prediction candidate in L1 direction. The new uni-directional candidates are constructed by splitting regular merge candidates.

Test 4.2.8.a and test 4.2.15.c are quite similar. Combination of the first 4 regular merge candidates is performed to derive 6 pair candidates, and then each pair of candidates are averaged to get one combined candidate. The difference is in the reference picture selection when the two MV to be averaged points to different reference pictures.

Proponents are requested to provide an analysis about the number of operations, comparisons, memory accesses and additional storage needs, etc. for the list construction, also in comparison with current method from VTM. To be done in BoG (X. Li). See further notes under BoG XXXX.

**Proposals on Sub-block based candidate – Affine candidates**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| AFFINE | BMS AFFINE as benchmark |  |
| 4.2.3.e | 4.2.3.d + Unified merge candidate list (number of all candidates is 15) | JVET-K0339 |
| 4.2.4.a | Combined test 4.2.3.e with model storage at CU level | JVET-K0219 |
| 4.2.4.b | Combined test 4.2.3.e with added affine candidates (17 candidates with up to 8 Affine ones) | JVET-K0219 |
| 4.2.4.c | 4.2.4.a + 4.2.4.b | JVET-K0219 |
| 4.2.7 | Affine merge with up to two affine merge candidate | JVET-K0094 |
| 4.2.8.a | Adding affine merge candidates (number of all candidates is 6) | JVET-K0245 |
| 4.2.10.a | Model based affine candidates (number of affine candidates is 2) | JVET-K0186 |
| 4.2.10.b | Additional control point based affine candidates (number of affine candidates is 5) | JVET-K0186 |
| 4.2.12.a | Use affine model of the neighbor coding unit with largest size as the merge candidate (uses 8-parameter model) | JVET-K0355 |
| 4.2.12.b | Add a separate merge list with 8 candidates using different motion models | JVET-K0355 |
| 4.2.12.c | 4.2.12.a + 4.2.12.b | JVET-K0355 |

Random access

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| AFFINE | -2.99% | -2.19% | -2.21% | 137% | 112% | -1.92% | -1.34% | -1.36% | 108% | 102% |
| 4.2.3.d | -0.88% | -0.85% | -0.85% | 109% | 103% | -0.76% | -0.71% | -0.78% | 105% | 103% |
| 4.2.3.e | **/** | **/** | **/** | **/** | **/** | -1.05% | -0.97% | -1.04% | 104% | 104% |
| 4.2.4.a | -4.03% | -3.28% | -3.26% | 146% | 117% | -2.85% | -2.22% | -2.28% | 119% | 99% |
| 4.2.4.b | -4.03% | -3.24% | -3.27% | 148% | 119% | -2.80% | -2.15% | -2.24% | 117% | 101% |
| 4.2.4.c | -4.08% | -3.35% | -3.35% | 150% | 129% | -2.86% | -2.22% | -2.29% | 119% | 107% |
| 4.2.7 | -3.10% | -2.26% | -2.32% | 143% | 115% | -1.97% | -1.36% | -1.39% | 111% | 102% |
| 4.2.8.a | -3.36% | -2.58% | -2.62% | 132% | 122% | -2.12% | -1.53% | -1.53% | 110% | 89% |
| 4.2.10.a | -3.09% | -2.27% | -2.28% | 137% | 110% | -1.97% | -1.39% | -1.42% | 110% | 101% |
| 4.2.10.b | -4.08% | -3.22% | -3.29% | 155% | 117% | -2.75% | -2.02% | -2.13% | 120% | 103% |
| 4.2.12.a | -3.05% | -2.21% | -2.22% | 138% | 112% | -1.94% | -1.40% | -1.41% | 111% | 102% |
| 4.2.12.b | -3.97% | -2.75% | -2.72% | 167% | 129% | -2.69% | -1.52% | -1.55% | 120% | 108% |
| 4.2.12.c | -3.99% | -2.74% | -2.72% | 167% | 129% | -2.69% | -1.56% | -1.59% | 121% | 108% |

Low delay B

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| AFFINE | -2.06% | -1.33% | -1.52% | 168% | 108% | -1.98% | -1.52% | -1.70% | 125% | 104% |
| 4.2.3.d | -0.49% | -0.24% | -0.14% | 111% | 104% | -0.44% | -0.24% | -0.08% | 103% | 100% |
| 4.2.3.e | **/** | **/** | **/** | **/** | **/** | -0.52% | -0.35% | -0.31% | 102% | 104% |
| 4.2.4.a | -2.50% | -1.53% | -1.57% | 181% | 100% | -2.47% | -1.83% | -1.83% | 134% | 104% |
| 4.2.4.b | -2.60% | -1.72% | -1.64% | 182% | 105% | -2.52% | -2.05% | -2.10% | 136% | 105% |
| 4.2.4.c | -2.67% | -1.86% | -1.82% | 187% | 117% | -2.57% | -2.11% | -1.94% | 136% | 113% |
| 4.2.7 | -2.09% | -1.37% | -1.49% | 181% | 115% | -2.05% | -1.81% | -1.97% | 130% | 104% |
| 4.2.8.a | -2.23% | -1.53% | -1.67% | 159% | 110% | -2.21% | -1.81% | -1.75% | 126% | 113% |
| 4.2.10.a | -2.11% | -1.37% | -1.39% | 168% | 107% | -2.05% | -1.73% | -1.73% | 125% | 102% |
| 4.2.10.b | -2.67% | -1.99% | -1.96% | 194% | 113% | -2.51% | -1.86% | -2.25% | 146% | 105% |
| 4.2.12.a | -2.11% | -1.36% | -1.48% | 169% | 109% | -2.11% | -1.58% | -1.63% | 127% | 105% |
| 4.2.12.b | -2.63% | -1.48% | -1.46% | 203% | 122% | -2.61% | -1.58% | -1.69% | 141% | 112% |
| 4.2.12.c | -2.66% | -1.42% | -1.25% | 203% | 122% | -2.63% | -1.63% | -1.63% | 142% | 111% |

Since tests here inherently requires affine motion compensation, BMS affine is suggested to be turned on for both anchor the test. And therefore, the RD performance of BMS affine is listed for comparison.

In BMS affine merge mode, only one affine merge candidate is allowed. It is inherited from the first available neighboring block with affine mode. Using the inherited affine model, two motion vectors at the top-left and top-right corners of the current CU is derived, and then the motion vector for each sub-block is computed using the two derived CPMVs.

The commonality among all tests are

* BMS affine merge mode is extended by adding more inherited affine merge candidates.
* Various constructed affine merge candidates are added, including 4-param / 6-param affine model based candidates.

Differences are

* The position from where a candidate is derived, including added neighboring and long distance spatial positions, and temporal positions.
* The number of inherited affine candidates and constructed affine candidates.
* Reference index derivation when constructing an affine merge candidates.
* Whether affine candidates are put in a separate list or in the existing merge list.
* In case of separate list, the order in which affine candidates are organized.
* In case of no separate list, the way of inserting affine candidates in the existing candidate list.

Note that the motion of four neighboring blocks are used to construct a bilinear motion model in test 4.2.12.

Highest coding gain from adding more affine merge candidates shown here is about 1% in VTM tool test.

One aspect that differentiates proposals is whether a separate merge list is used, and whether the affine-specific candidates are only inherited or also constructed (see under CE4.1, affine MV pred.)

Results are difficult to interpret, as the gains come from mixture of using modified merge methods, other motion models, etc. For subsequent experiments, this should be put on a more common basis. Proponents are asked to form a side activity (coordinated by H. Chen) with the goal of establishing a common basis for the subsequent round of CE. Different methods tested should also be made comparable in terms of the amount of encoder and RD optimization.

Proponents are requested to provide an analysis about the number of operations, MV comparisons, memory usage, additional storage, etc. for the list construction and the inheritance, also in comparison with BMS affine. No modification of the affine merge mode should be made that would end up with higher worst case complexity as compared to the adopted MV prediction (as per CE4.1) Was reviewed Monday afternoon in track B (and later uploaded as JVET-K0558).

From the results, it seems that the concept of separate list seems to be useful in terms of worst case decoder complexity, as the list construction for affine needs to be quite different from the normal merge mode, even though it requires signalling an additional flag.

**Proposals on Sub-block based candidate – Planar candidates**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| 4.2.11 | MV Planar | JVET-K0349 |
| 4.2.14 | MV Planar | JVET-K0135 |

Random access results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 4.2.11 | -0.32% | -0.26% | -0.23% | 107% | 110% | -0.33% | -0.19% | -0.20% | 103% | 104% |
| 4.2.14 | -0.91% | -0.94% | -0.94% | 108% | 119% | -0.49 % | -0.46% | -0.43 % | 104% | 107% |

Low delay B results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 4.2.11 | -0.09% | 0.00% | -0.02% | 111% | 108% | -0.22% | -0.10% | 0.05% | 106% | 103% |
| 4.2.14 | -0.70% | -0.64% | -0.75% | 109% | 112% | -0.51% | -0.50% | -0.42% | 105% | 104% |

Basic idea of the two MV planar mechanisms here is extrapolating motion vectors of each sub-block inside a coding block similarly as intra prediction planar mode.

In case of no motion info in the L-shaped neighboring area, the motion vector is derived in the same manner as intra reference pixel padding, separately for L0 and L1.

In test 4.2.11, inter prediction direction “inter\_pred\_idc” and reference index “ref\_idx” for both L0 and L1 lists are signalled. In contrast, prediction direction is derived at the decoder side, and the ref\_idx is set to 0 for both L0 and L1.

Another major difference from test 4.2.11 is using the bottom-right corner temporal motion vectors for planar prediction.

Targets similar aspects as affine, but with less coding gain. However, BMS results unveil that there is still some additive gain when combined. Should be further studied, specifically in terms of interrelation with new methods in affine parameter coding and affine sub-block merge, which seem to promise more gain.

**Proposals on Sub-block based candidate – ATMVP modification**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| ATMVP | BMS ATMVP as benchmark |  |
| ATMVP +HPMV | BMS ATMVP + HPMV as benchmark |  |
| 4.2.5.a | ATMVP with one fixed collocated picture | JVET-K0341 |
| 4.2.5.b | Picture/slice-level adaptation of ATMVP sub-block-size | JVET-K0341 |
| 4.2.5.c | 4.2.5.a + 4.2.5.b | JVET-K0341 |
| 4.2.6.a | Simplified ATMVP with one fixed collocated picture | JVET-K0079 |
| 4.2.6.b | 4.2.6.a + 4.2.5.b | JVET-K0079 |
| 4.2.8.d | Double ATMVP candidates with uniformity check and decoder-side speedup | JVET-K0245 |
| 4.2.16.a | ATMVP in BMS | JVET-K0338 |

Random access results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| ATMVP | -0.87% | -0.81% | -0.83% | 104% | 120% | -0.74% | -0.62% | -0.61% | 100% | 105% |
| ATMVP+  HPMV | -1.10% | -0.88% | -0.85% | 102% | 122% | / | / | / | / | / |
| 4.2.5.a | -1.09% | -0.87% | -0.87% | 105% | 122% | -0.73% | -0.63% | -0.63% | 100% | 105% |
| 4.2.5.b | -1.06% | -0.82% | -0.80% | 103% | 106% | -0.69% | -0.57% | -0.57% | 100% | 101% |
| 4.2.5.c | -1.04% | -0.81% | -0.80% | 102% | 105% | -0.68% | -0.56% | -0.59% | 99% | 101% |
| 4.2.6.a | -1.09% | -0.86% | -0.84% | 106% | 114% | -0.73% | -0.66% | -0.65% | 101% | 106% |
| 4.2.6.b | -1.03% | -0.79% | -0.77% | 102% | 105% | -0.69% | -0.62% | -0.59% | 110% | 100% |
| 4.2.8.d | -1.00% | -0.92% | -0.90% | 105% | 118% | -0.77% | -0.67% | -0.62% | 102% | 106% |
| 4.2.16.a | -0.87% | -0.81% | -0.83% | 103% | 115% | 0.75% | 0.63% | 0.62% | 100% | 95% |

Low delay B results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| ATMVP | -1.08% | -1.27% | -1.23% | 105% | 114% | -1.02% | -1.08% | -1.25% | 99% | 106% |
| ATMVP+  HPMV | -1.31% | -1.29% | -1.25% | 104% | 116% | / | / | / | / | / |
| 4.2.5.a | -1.31% | -1.28% | -1.14% | 106% | 115% | -1.10% | -1.03% | -1.43% | 99% | 108% |
| 4.2.5.b | -1.29% | -1.29% | -1.17% | 103% | 109% | -1.02% | -1.11% | -1.28% | 99% | 104% |
| 4.2.5.c | -1.30% | -1.26% | -1.28% | 103% | 108% | -1.07% | -1.13% | -1.38% | 99% | 104% |
| 4.2.6.a | -1.30% | -1.10% | -1.10% | 105% | 113% | -1.05% | -1.14% | -1.35% | 99% | 106% |
| 4.2.6.b | -1.27% | -1.12% | -1.10% | 102% | 106% | -1.08% | -0.98% | -1.29% | 100% | 104% |
| 4.2.8.d | -1.19% | -1.38% | -1.32% | 104% | 109% | -1.15% | -1.35% | -1.48% | 101% | 105% |
| 4.2.16.a | -1.08% | -1.27% | -1.23% | 104% | 112% | 1.03% | 1.11% | 1.28% | 101% | 94% |

Test 4.2.5.a and test 4.2.6.a propose using one fixed collocated picture for fetching motion for ATMVP mode. The difference is in handling the situation when the motion vector of a neighboring block points to a reference picture other than the fixed collocated picture. Test 4.2.5.a performs MV scaling while test 4.2.6.a simply use the collocated blocks.

The same slice level adaptation of sub-block size is applied in 4.2.5.b and 4.2.6.b.

In test 4.2.8.d, two MVs from merge candidate list instead of one are used as temporal MVs to generate two ATMVP candidates. Two types of uniformity check are then performed to the two candidates

* If the motion of the four corners are the same for the two ATMVP candidates, the second candidate is removed from the list.
* If the motion of the four corners the same as the center motion for an ATMVP candidate, it is treated as TMVP.

These proposals are targeting more of implementation aspects, but affine technology is still subject to changes. Some of the aspects of 4.2.8d would not even require normative specification. No need for action at this moment.

**Proposals on Sub-block based candidate – STMVP**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| 4.2.16.b | STMVP in JEM | JVET-K0338 |
| 4.2.16.c | 4.2.16.a + 4.2.16.b | JVET-K0338 |

Random access results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 4.2.16.b | -0.76% | -1.01% | -0.95% | 105% | 121% | 0.06% | -0.19% | -0.20% | 100% | 103% |
| 4.2.16.c | -1.30% | -1.36% | -1.38% | 109% | 128% | -0.37% | -0.39% | -0.41% | 101% | 105% |

Low delay B results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 4.2.16.b | -0.68% | -0.54% | -0.58% | 106% | 115% | 0.11% | 0.13% | 0.23% | 101% | 104% |
| 4.2.16.c | -1.31% | -1.37% | -1.39% | 111% | 120% | -0.37% | -0.14% | -0.09% | 101% | 101% |

The tests provide results of JEM STMVP and the combination with ATMVP. Comparison with 4.2.16a shows that STMVP has worse performance/complexity tradeoff than ATMVP, and also the combination does not provide that much benefit in BMS particularly. No action.

**CE4.3: Motion Vector Coding** (discussed track B Thursday 12th 1900-2100, chaired by JRO)

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| 4.3.1 | MV prediction between two directions for AMVP mode | JVET-K0234 (Ericsson) |
| 4.3.2 | Symmetrical mode for bi-prediction | JVET-K0188 (Huawei) |
| 4.3.3.a | Adaptive MV resolution (1/4 or 1 or 4) | JVET-K0357 (Qualcomm) |
| 4.3.3.b | 4.3.3.a + Efficient MVD coding |
| 4.3.4 | Adaptive motion vector resolution in J0018 (1/4 or 1 or 4) | JVET-K0247 (MediaTek) |
| 4.3.5.a | AMVR in J0024 with three MV resolutions ( ¼ or 1 or 4) | JVET-K0116 (Samsung) |
| 4.3.5.b | AMVR in J0024 with five MV resolutions (1/4 or ½ or 1 or 2 or 4) |
| 4.3.5.c | 4.3.5.a + combined signaling of MVR index and MVP index |
| 4.3.5.d | 4.3.5.b + combined signaling of MVR index and MVP index |
| 4.3.6 | Unicity Check for AMVP candidate list | JVET-K0208 (Technicolor) |

Random access results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| **BMS AMVR** | **-1.75%** | **-2.36%** | **-2.39%** | **116%** | **100%** | **-1.10%** | **-1.51%** | **-1.49%** | **103%** | **98%** |
| 4.3.1 | -0.29% | -0.27% | -0.24% | 102% | 100% | -0.13% | -0.11% | -0.14% | 101% | 99% |
| 4.3.2 | -0.84% | -0.62% | -0.56% | 105% | 100% | -0.15% | -0.14% | -0.18% | 102% | 100% |
| 4.3.3.a | -1.75% | -2.36% | -2.39% | 118% | 101% | -1.10% | -1.51% | -1.49% | 103% | 98% |
| 4.3.3.b | -1.79% | -2.40% | -2.43% | 117% | 101% | -1.10% | -1.56% | -1.49% | 103% | 98% |
| 4.3.4 | -1.79% | -2.41% | -2.52% | 121% | 100% | -1.12% | -1.57% | -1.58% | 105% | 98% |
| 4.3.5.a | -1.46% | -1.91% | -1.98% | 91% | 98% | -0.74% | -1.35% | -1.38% | 91% | 89% |
| 4.3.5.b | -1.60% | -2.17% | -2.35% | 111% | 99% | -0.81% | -1.49% | -1.56% | 97% | 89% |
| 4.3.5.c | -1.04% | -1.48% | -1.57% | 93% | 99% | -0.52% | -1.08% | -1.11% | 92% | 89% |
| 4.3.5.d | -1.16% | -1.75% | -1.87% | 114% | 99% | -0.59% | -1.24% | -1.33% | 98% | 89% |
| 4.3.6 | -0.07% | -0.08% | -0.03% | 101% | 102% | -0.05% | -0.12% | -0.06% | 101% | 103% |

Low delay B results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| **BMS AMVR** | **-0.68%** | **-1.27%** | **-1.39%** | **131%** | **102%** | **-0.57%** | **-0.99%** | **-1.10%** | **107%** | **97%** |
| 4.3.1 | / | / | / | / | / | / | / | / | / | / |
| 4.3.2 | / | / | / | / | / | / | / | / | / | / |
| 4.3.3.a | -0.68% | -1.27% | -1.39% | 129% | 100% | -0.57% | -0.99% | -1.10% | 108% | 98% |
| 4.3.3.b | -0.76% | -1.20% | -1.35% | 129% | 100% | -0.67% | -1.21% | -1.13% | 111% | 98% |
| 4.3.4 | -0.72% | -1.39% | -1.56% | 142% | 99% | -0.70% | -1.16% | -1.31% | 115% | 99% |
| 4.3.5.a | -0.65% | -1.25% | -1.33% | 130% | 98% | -0.65% | -1.25% | -1.33% | 130% | 98% |
| 4.3.5.b | -0.67% | -1.55% | -1.66% | 167% | 99% | -0.67% | -1.55% | -1.66% | 167% | 99% |
| 4.3.5.c | -0.53% | -1.21% | -1.34% | 132% | 98% | -0.53% | -1.21% | -1.34% | 132% | 98% |
| 4.3.5.d | -0.56% | -1.34% | -1.59% | 166% | 78% | -0.47% | -1.02% | -1.21% | 116% | 91% |
| 4.3.6 | -0.01% | -0.09% | 0.04% | 101% | 96% | -0.19% | -0.20% | 0.00% | 100% | 104% |

4.3.1: Does not provide significant gain, in particular in BMS -> interrelation with other tools significant

4.3.2: Symmetrical mode is simple and has high benefit in VTM. However, the gain is only low in BMS. Further study: Interrelation with other tools, elaborate how to define syntax and semantics -> seems to be necessary to impose certain conditions on reference picture list. Seems also beneficial to introduce an enabling flag at slice level.

4.3.5: The results indicate that increasing the number of switchable MV precisions from 3 to 5 does not give benefit. The decrease in encoding time of 4.3.5c is due to a fast ME algorithm which obviously loses performance but could likewise be applied to the anchor as well.

The other modifications of AMVR only show marginal modifications of performance and/or encoder/decoder run time. In particular, moving the AMVR flag to another position in the syntax does not seem beneficial. No action on 4.3.3b, 4.3.4 and 4.3.6

Decision: Include BMS-AMVR (4.3.3a) into VTM. This tool is well understood, simple and provides significant gain (-1.75% in RA VTM, 1.1% in BMS)

[JVET-K0566](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4096) Draft text for advanced temporal motion vector prediction (ATMVP) [X.Xiu, Y. Ye (InterDigital), H. Huang, W.-J. Chien (Qualcomm), H. Jang (LGE)] [late]

Revisit: ATMVP/AMVP review (JVET-K0566) still to be done

Remaining discussions CE4 track B, Friday 13th, 900- (chaired by JRO)

**CE4.4: Generalized Bi-prediction**

JVET- K0248 (MediaTek)

In this contribution, GBi is presented to allow applying different weights to predictors from L0 and L1. The predictor generation is shown in Equ. (2).

|  |  |
| --- | --- |
| *PGBi = ( (1-w1)\* PL0 + w1 \* PL1 + RoundingOffsetGBi) >> shiftNumGBi*, | (2) |

Weights for true bi-prediction cases in random access (RA) condition.

|  |  |  |
| --- | --- | --- |
| **GBi Index** | **Weight value of *w1*** | **Binarization of GBi Index** |
| 0 | 3/8 | 00 |
| 1 | 1/2 | 1 |
| 2 | 5/8 | 01 |

Weights in generalized bi-prediction

|  |  |  |
| --- | --- | --- |
| **GBi Index** | **Weight value of *w1*** | **Binarization of GBi Index** |
| 0 | -1/4 | 0000 |
| 1 | 3/8 | 001 |
| 2 | 1/2 | 1 |
| 3 | 5/8 | 01 |
| 4 | 5/4 | 0001 |

For advanced motion vector prediction (AMVP) mode, the weight selection in GBi is explicitly signalled at CU-level if this CU is coded by bi-prediction. For merge mode, the weight selection is inherited from the merge candidate. In this proposal, GBi supports DMVR to generate the weighted average of template as well as the final predictor for BMS-1.0.

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| 4.4.1 | Generalized bi-prediction | JVET-K0248 (MediaTek) |

Random access results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 4.4.1 | -0.86% | -0.96% | -0.95% | 114% | 102% | -0.56% | -0.74% | -0.80% | 108% | 102% |

Low delay B results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 4.4.1 | -1.07% | -0.37% | -0.26% | 114% | 102% | -0.40% | -0.54% | -0.20% | 112% | 100% |

Has some commonality with weighted prediction. However, WP is only invoked at slice level, whereas the proposal invokes different weights at CU level. The tools could probably coexist, WP would better be suited for global illumination changes such as fade.

It is pointed out that potentially this could potentially also be achieved by using WP together with ref picture indexing. This would however be more complicated as it always requires signalling two selev´cted indices.

The gain drops significantly in case of BMS, in particular for LDB. The proponents report this is mainly due to interference with ALF.

On the other hand, the encoder complexity increase is not large, and decoder complexity increase is negligible.

Decision: Adopt JVET-K0248 to BMS

**CE4.5: Reference Picture Boundary Padding**

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| 4.5.1 | Multi-directional boundary padding (MDBP) | JVET-K0195 (HHI) |
| 4.5.2 | MC Padding | JVET-K0363 (Qualcomm) |
| 4.5.3 | Boundary pixel padding using motion compensation | JVET-K0117 (Samsung) |

Random access results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 4.5.1 | -0.07% | -0.05% | -0.07% | 102% | 101% | -0.04% | -0.10% | -0.08% | 102% | 102% |
| 4.5.2 | -0.22% | -0.20% | -0.17% | 100% | 104% | -0.22% | -0.22% | -0.19% | 107% | 108% |
| 4.5.3 | -0.18% | -0.19% | -0.16% | 100% | 107% | -0.18% | -0.17% | -0.17% | 100% | 103% |

Low delay B results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 4.5.1 | -0.07% | -0.07% | -0.11% | 104% | 103% | -0.13% | -0.08% | -0.07% | 103% | 100% |
| 4.5.2 | -0.02% | 0.04% | 0.10% | 102% | 105% | -0.08% | -0.11% | 0.11% | 104% | 105% |
| 4.5.3 | -0.01% | 0.12% | 0.01% | 101% | 109% | -0.07% | -0.15% | 0.03% | 99% | 105% |

Padding width is either 64 (K0363) or 128 (other two proposals).

May be difficult to implement on the fly, in particular the motion compensated method. If it is done in ref picture memory, the increase would be large. Intra based method could be implemented on the fly, but also gives lowest gain.

Relative low gain compared to the increase of decoder complexity; however this may be caused by the fact that it improves only in boundary areas. For low resolutions, the gain would be higher, but also the complexity increase would be higher.

Questions are raised with regard to subjective quality; does it improve the visual quality at boundaries? Is visual quality at boundaries relevant, as often observers tend to look in center?

Further study with regard to impact on decoder memory, possible complications of on-the-fly processing, and subjective quality.

**CE4.6: Local illumination compensation** (Track B, Fri 13th 1000-1030, chaired by JRO)

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| 4.6.1 | Bi-directional illumination compensation (BIC) | Withdraw |
| 4.6.2 | Combination of BIC, generalized OBMC and Simplified BIC parameters derivation | Withdraw |
| 4.6.3 | Inter prediction refinement | JVET-K0118 (Samsung) |

Random access results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 4.6.1 | / | / | / | / | / | / | / | / | / | / |
| 4.6.2 | / | / | / | / | / | / | / | / | / | / |
| 4.6.3 | -0.37% | -0.32% | -0.29% | 127% | 112% | -0.31% | -0.10% | -0.17% | 113% | 106% |

Low delay B results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 4.6.1 | / | / | / | / | / | / | / | / | / | / |
| 4.6.2 | / | / | / | / | / | / | / | / | / | / |
| 4.6.3 | -0.31% | -0.09% | -0.16% | 146% | 111% | -0.30% | -0.24% | -0.11% | 125% | 107% |

JVET-K0118

At the end of processing of each CU, linear models (Ax + B = y) are computed by comparing the prediction pixels (x) and the reconstructed pixels (y). The models are computed using linear regression. The lower edge pixels are used for model computation. Two sets of models are computed, one for each colour component. Each set contains models for the two directions (if available). These parameters (sets of A and B) are then stored in the motion information buffer for later access.

The conditions under which IPR is applicable is kept same as that of JEM LIC. These conditions depend on the size of the current CU, type of prediction mode and whether the CU uses merge mode or not. If applicable, a flag is signalled to indicate whether IPR is to be applied to a given inter coded CU or not. If the tool is to be applied, the motion compensation pixels (x) are refined using the application of the linear model (Ax + B) to obtain the new set of prediction pixels. The model used is derived from the current CU’s neighbor. Top and top left neighbour are considered for borrowing the parameter. Since both the neighbors may be available, pruning is performed to obtain a single model. Intra coded neighbours are discarded in the pruning process. Next, the neighbours which have trivial model [x = y] are discarded in the pruning process. If no model is remaining, a trivial model is used. If model(s) still remain, the neighbours with the same reference index as the current CU are given higher priority. If multiple neighbours have the same reference index, the top neighbour is given higher priority compared to the top-left neighbour.

Additionally, when performing RDO check for determining whether IPR will be used or not, mean-removed SAD (Sum of Absolute Difference) is used instead of the regular SAD for the ME (Motion Estimation) process.

The linear regression method used is same as that in JEM. Also, using two sides for parameter derivation improves the gain.

**Key points**

* LIC model pre-computation and storage after processing a CU
* LIC model pruning to get the only one model applied to a CU

Additional line buffer is necessary to store the parameters of IC or the prediction samples reconstructed samples (not all, only subsampled set is used) of the previous CU row.

Gain (RA) close to 0.4% in VTM; still 0.3% in BMS. Decoding time increase is 12%/6%

Implementation of linear regression uses lookup table and shift operation, similar as LM chroma. Gain would likely be higher for sequences with more illumination changes. Gain in current test set is up to 1% (Ritual Dance).

Illumination compensation is beneficial for some sequences. It also seems to have low interference with other tools as shown by the BMS results. However, the proposed method has relative high memory requirements and computation needs (e.g. by comparing models from two blocks) versus the relative moderate gain. Further study in particular w.r.t. memory needs.

**CE4.7: Chroma interpolation filter** (Track B, Fri 13th 1000-1030, chaired by JRO)

|  |  |  |
| --- | --- | --- |
| **Test#** | **Description** | **Document#** |
| 4.7.1 | Modified 4-tap chroma interpolation filter | JVET-K0207 (Huawei) |

Random access results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 4.7.1 | -0.01% | -0.57% | -0.62% | 102% | 99% | 0.02% | -0.33% | -0.41% | 100% | 99% |

Low delay B results

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **VTM\_tool\_test** | | | | | **BMS\_tool\_test** | | | | |
| **Test#** | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 4.7.1 | 0.03% | 0.31% | 0.31% | 100% | 98% | -0.03% | -0.22% | -0.08% | 100% | 103% |

Why does it have worse performance for LDB in VTM? Not known.

The filters are constructed as combination of bilinear and sharpening filter. This might cause ringing at edges or other artifacts for case of chroma discontinuities. Besides the fact that the gain is low, the impact on visual should be carefully considered.

No action.

[JVET-K0047](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3548) CE4: Affine flexing (Test 1.2) [J. Lainema (Nokia)]

The approach performs block-based motion compensation followed by a resampling or "flexing" filtering of the prediction block. The original proposal is based on 4x4 luma block motion compensation as in JEM 7.0. In addition, a version with 8x8 luma block compensation was tested. The test was using the 16-position JEM motion compensation filter for the flexing operation.

The affine flexing tool calculates the horizontal and vertical motion vector differences on the affine sub-PU boundaries. Based on whether the nature of motion is more rotational or zooming, the method decides to apply either lateral compensation (moving the sub-PU boundary samples laterally with respect to the neighboring sub-PU samples) or perpendicular compensation (moving the sub-PU boundary samples away or closer to the samples of the neighboring sub-PU), respectively. The remapping offsets applied to different lines are calculated based on motion vector difference and distance from the sub-PU border. A look-up table implementation is provided in the CE software.

Further details:

* Flexing is applied to both horizontally and vertically, and the result of the flexing in first dimension is used as an input to the second dimension
* Flexing is applied only for the luma channel in the case of 4:2:0 video
* Line ends are padded with the last available samples on the line

[JVET-K0079](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3582) CE4-2.6: Simplified ATMVP [H. Jang, J. Lim, J. Nam, S. Kim (LGE)]

The proposed simplified ATMVP is basically similar to ATMVP as place holder for the temporal motion vector derivation. However, in the proposed method, the usage of memory is effectively reduced because the reference picture to find the corresponding block is restricted to the collocated picture (designated in the slice segment header), while 4 reference pictures are required in the place holder method. To find the corresponding block, a temporal vector is derived from one of spatial candidates in a scanning order. If the current candidate has the reference picture same with the collocated picture, the search is finished. If all of candidates do not refer the collocated picture then zero motion is applied.

In CE4-2.6(a), it is applied that Simplified ATMVP refers only one fixed collocated picture to derive temporal motion information, and the sub-block size is decided between 4x4 and 8x8 based on slice level syntax. This sub-block size decision algorithm is identical with CE4-2.5(b). The sub-block size decision algorithm decides sub-block size for ATMVP based on statistic of the prediction unit size and pre-defined threshold value for each slice and each temporal layer in encoder side. In decoder side, adaptive sub-block decision on/off flag is signalled and sub-block size is also signalled when the on/off flag is set as true in slice header.

* CE4-2.6(a): one fixed collocated picture is used to derive temporal motion information
* CE4-2.6(b): CE4-2.6(a) plus the slice level adaptive sub-block decision in CE4-2.5(b).

[JVET-K0094](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3597) CE4: Affine prediction modification (CE4.1.7 and CE4.2.7) [J. Lee, J. Nam, N. Park, H. Jang, J. Lim, S. Kim (LGE)]

***Test 4.1.7.a: Affine motion vector prediction***

Affine motion vector prediction constructs up to two candidates derived from the affine motion models of neighboring affine coded blocks. Neighbouring blocks, A, B, C, D, and E are checked whether it is coded as the affine prediction and its reference frame is same as the reference frame of the current block. Neighboring blocks with affine prediction and same reference frame are considered firstly, and then neighboring blocks with affine prediction and different reference frame are considered to generate the affine candidate. If the number of generated affine candidates is less than two, JEM affine process is performed until the number of candidate is two.

* Checking neighbour blocks in the same order as HEVC AMVP, candidate number is two
* Firstly derive affine MVP from the affine model of neighboring blocks
* Secondly construct affine MVP from motion vector of neighboring blocks as BMS affine

***Test 4.1.7.b: Four and six parameter motion model***

In the proposed method, four parameter affine model with two control point motion vectors (CPMVs) and six parameter affine model with three CPMVs are used.

***Test 4.1.7.c: Adaptive four and six parameter motion model***

Difference compared with test 4.1.7.b

* Alternative 6-param affine model is allowed only when a neighboring block is in affine mode
* When both 4-param and 6-param is allowed, an indication flag is by-pass coded

In the proposed algorithm, affine merge candidate is derived from the affine motion model of neighboring affine-coded block and up to two candidates can be considered in the affine merge list. Affine merge mode is skipped when there is no neighboring affine coded block. Affine merge index is signalled only when the available neighboring affine coded block is more than one.

* Max candidate number is 2
* Affine merge candidate signaling depends on whether neighboring block is in affine mode

[JVET-K0115](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3620) CE4 Ultimate motion vector expression in J0024 (Test 4.2.9) [S. Jeong, M. W. Park, C. Kim (Samsung)]

[JVET-K0116](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3621) CE4: Adaptive Motion Vector Resolution in JVET-J0024 (Test 4.3.5) [A. Tamse, S. Jeong, M. W. Park, C. Kim (Samsung)]

[JVET-K0117](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3622) CE4: Reference Picture Boundary Padding in JVET-J0025 (Test 4.5.3) [M. Park, M. W. Park, W. Choi, C. Kim (Samsung)]

[JVET-K0118](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3623) CE4: Inter Prediction Refinement in JVET-J0024 (Test 4.6.3) [A. Tamse, C. Kim (Samsung)]

[JVET-K0124](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3629) CE4: Adaptive three and four parameter motion model (Test 4.1.8) [K. Kondo, T. Suzuki (Sony)]

An encoder can choose four modes which are a conventional translate, scaling, rotation and Affine. A number of parameters are 2 (one motion vector), 3, 3 and 4 (two MVs) for translate mode, scaling mode, rotation mode and Affine mode.

The motion model index is defined and it is sent to the decoder. The number in parenthesizes shows the number of parameters of differential motion vector that are sent to a decoder.

When the scaling mode is chosen, three parameters are used to predict. It is one motion vector () and one element (). For the prediction process, the frame work is used the same as for Affine transform in JEM software. The parameter () is implicitly assumed as the same as ().

[JVET-K0135](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3641) CE4.2.14: Planar Motion Vector Prediction [N. Zhang, Y. Lin, J. Zheng (HiSilicon)]

Planar motion vector prediction is achieved by averaging a horizontal and vertical linear interpolation on 4x4 block basis as follows.

*W* and *H* denote the width and the height of the block. *(x,y)* is the coordinates of current sub-block relative to the above left corner sub-block. All the distances are denoted by the pixel distances divided by 4. is the motion vector of the current sub-block.

calculated as follows:

where and are the motion vectors of the 4x4 blocks to the left and right of the current block. and are the motion vectors of the 4x4 blocks to the above and bottom of the current block.

***Key points***

* Separately signalled (not in the current merge list)
* Motion in the L-shaped area is padded in the same manner as intra reference pixel padding, separately for L0 and L1.
* Bi-directional or uni-directional: decided based on the prediction direction of the motion in the L-shaped area. Once there is a L0 or L1 MV, L0 or L1 direction is activated.
* Ref\_idx in L0 and L1 is set to 0.

[JVET-K0184](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3693) CE4: Affine motion compensation with fixed sub-block size (Test 1.1) [H. Chen, H. Yang, J. Chen (Huawei)]

[JVET-K0185](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3694) CE4: Affine inter prediction (Test 1.5) [H. Chen, H. Yang, M. Sychev, J. Chen (Huawei)]

This contribution reports the results of integrating affine inter prediction in JVET-J0024 [1] to BMS1.0. In this contribution, affine inter prediction is changed by 4 aspects:

1) Model based affine candidates are inserted into affine AMVP candidates with the same scan order of AMVP in HEVC; the candidate reorder operation in BMS affine is removed.

2) Affine MVD zero flag is signalled to indicate whether affine MVDs is zero or not;

3) Additional 6-parameter affine model is added and different models are adaptively selected in CU level;

4) Enhanced bi-linear Interpolation Filter (EIF) is used for affine motion compensation, when the width or height of affine sub-block is less than 8.

For affine AMVP improvement, experimental results reportedly show on average 0.26%/0.06% luma BD-rate gain in RA/LB configurations over to VTM with affine on. For EIF, experimental results reportedly show on average 0.25%/0.41% luma BD-rate gain with 9%/6% decoding time reduction in RA/LB configurations over VTM with affine on.

[JVET-K0186](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3695) CE4: Affine merge enhancement (Test 2.10) [H. Chen, H. Yang, J. Chen (Huawei)]

UMVE is used for either skip or merge modes with a proposed motion vector expression method.

UMVE re-uses merge candidate as same as using in VVC. Among the merge candidates, a candidate can be selected, and is further expanded by the proposed motion vector expression method.

UMVE provides a new motion vector expression with simplified signaling. The expression method includes prediction direction information, starting point, motion magnitude, and motion direction.

Prediction direction information indicates a prediction direction among L0, L1 and L0 and L1 predictions. In B slice, the proposed method can generate bi-prediction candidates from merge candidates with uni-prediction by using mirroring technique. For example, if a merge candidate is uni-prediction with L1, a reference index of L0 is decided by searching a reference picture in list 0, which is mirrored with the reference picture for list 1. If there is no corresponding picture, the nearest reference picture to current picture is used. L0’ MV is derived by scaling L1’s MV. The scaling factor is calculated by POC distance.

If the prediction direction of the UMVE candidate is the same with one of the original merge candidate, the index with value 0 is signalled as an UMVE prediction direction. But, if not the same, same with one of the original merge candidate, the index with value 1 is signalled. After sending first bit, remaining prediction direction is signalled based on the pre-defined priority order of UMVE prediction direction. Priority order is L0/L1 prediction, L0 prediction and L1 prediction.

If the prediction direction of merge candidate is L1, signaling ‘0’ is for UMVE’ prediction direction L1. Signaling ‘10’ is for UMVE’ prediction direction L0 and L1. Signaling ‘11’ is for UMVE’ prediction direction L0.

If L0 and L1 prediction lists are same, UMVE’s prediction direction information is not signalled.

Base candidate index defines the starting point. Base candidate index indicates the best candidate among candidates in the list as follows.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Base candidate IDX** | 0 | 1 | 2 | 3 |
| **Nth MVP** | 1st MVP | 2nd MVP | 3rd MVP | 4th MVP |

If the number of base candidate is equal to 1, Base candidate IDX is not signalled.

Distance index is motion magnitude information. Distance index indicates the pre-defined distance from the starting point information. Pre-defined distance is as follows:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Distance IDX** | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| **Pixel distance** | 1/4-pel | 1/2-pel | 1-pel | 2-pel | 4-pel | 8-pel | 16-pel | 32-pel |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Direction IDX** | 00 | 01 | 10 | 11 |
| **x-axis** | + | – | N/A | N/A |
| **y-axis** | N/A | N/A | + | – |

In terms of Test a, one of merge candidate indexes is used as an UMVE flag. MRG\_MAX\_NUM\_CANS is increased by 1. In this test, 3rd index of merge candidate is used as an UMVE flag. In decoder side, if skip/merge index is 3, decoder starts to parse UMVE syntaxs for UMVE information. If skip/merge index is over 3, Actual skip/merge index is going to be skip/merge index minus 1. If received index is 4, 3rd candidate is selected as a candidate of original skip/merge mode.

In terms of Test b, UMVE flag is singnaled after sending a skip flag and merge flag. If skip and merge flag is true, UMVE flage is parsed. If UMVE flage is equal to 1, UMVE syntaxes are parsed. But, if not 1, skip/merge index is parsed for VTM/BMS’s skip/merge mode.

Additional line buffer due to UMVE candidates is not needed. Because a skip/merge candidate of software is directly used as a base candidate. Using input UMVE index, the supplement of MV is decided right before motion compensation. There is no need to hold long line buffer for this.

The NUM\_MRG\_SATD\_CAND is changed according to test option. It will be described with the performance at next chapter.

To reduce the encoder complexity, block restriction is applied. If either width or height of a CU is less than 4, UMVE is not performed.

* Test 4.2.9 (a) UMVE candidates as additional merge candidates with additional information signalled,
* Test 4.2.9 (b) UMVE candidates as independent merge candidates with additional information signalled.
* Test 4.2.9 (b) is consist of 3 sub-tests with changing parameters.
  + The number of NUM\_MRG\_SATD\_CAND is set to 120 at Sub-CE 4.2.9 b-1.
  + The number of NUM\_MRG\_SATD\_CAND is set to 384 at Sub-CE 4.2.9 b-2.
  + The number of NUM\_MRG\_SATD\_CAND is set to 32 at Sub-CE 4.2.9 b-3.

|  |  |  |  |
| --- | --- | --- | --- |
| **Test #** | **Description** | **Tester** | **Crosschecker** |
| 4.2.9 | a) UMVE candidates as additional merge candidates with additional information signalled  b) UMVE candidates as independent merge candidates (using new merge candidate list) with additional information signalled (including conventional merge candidates, “best“ set of candidate from this subCE, etc.) | Seungsoo Jeong (Samsung) | a) Ruoyang Yu (Ericsson)  b)Byeongdoo Choi (Sharp) |

**CE4.2.10 in JVET-K0186 (Huawei)**

The affine merge candidate list is constructed as following steps:

1. Insert model based affine candidates

Model based affine candidate means that the candidate is derived from the valid neighbor reconstructed block coded with affine mode. As shown in the figure, the scan order for the candidate block is from left, above, above right, left bottom to above left. The same derived method in BMS is used [2].

1. Insert control point based affine candidates

Control points based candidate means the candidate is constructed by combining the neighbor motion information of each control point.

The motion information for the control points is derived firstly from the specified spatial neighbors and temporal neighbor shown in the figure. CPk (k=1, 2, 3, 4) represents the k-th control point. A, B, C, D, E, F and G are spatial positions for predicting CPk (k=1, 2, 3); H is temporal position for predicting CP4.

The motion information of each control point is obtained according to the following priority order:

* For CP1, the checking priority is A🡪B🡪C, A is used if it is available. Otherwise, if B is available, B is used. If both A and B are unavailable, C is used. If all the three candidates are unavailable, the motion information of CP1 cannot be obtained.
* For CP2, the checking priority is E🡪D;
* For CP3, the checking priority is G🡪F;
* For CP4, H is used.

Secondly, the combinations of controls points are used to construct the motion model.

Motion vectors of two control points are needed to compute the transform parameters in 4-parameter affine model. The two control points can be selected from one of the following six combinations ({CP1, CP4}, {CP2, CP3}, {CP1, CP2}, {CP2, CP4}, {CP1, CP3}, {CP3, CP4}). For example, use the CP1 and CP2 control points to construct 4-parameter affine motion model, denoted as Affine (CP1, CP2).

Motion vectors of three control points are needed to compute the transform parameters in 6-parameter affine model. The three control points can be selected from one of the following four combinations ({CP1, CP2, CP4}, {CP1, CP2, CP3}, {CP2, CP3, CP4}, {CP1, CP3, CP4}). For example, use CP1, CP2 and CP3 control points to construct 6-parameter affine motion model, denoted as Affine (CP1, CP2, CP3).

All of these models will be converted to 6-parameter affine model represented by top-left, top-right, and bottom-left control point. During MC and motion vector derivation for sub-block, unified 6-parameter affine model is used.

***Key points***

* Two type of affine merge candidates: 1) 4-param affine merge candidates represented by 2 CPMV, 2) 6-param affine merge candidates represented by 3 CPMV.
* Two ways of candidate derivation: 1) derive 3 CPMV using the affine model of neighboring affine coded block, 2) derive 3 CPMV from the motion vector of neighboring blocks.
* A separate merge list is used.
* Max candidate number is 5.

|  |  |  |  |
| --- | --- | --- | --- |
| **Test #** | **Description** | **Tester** | **Crosschecker** |
| 4.2.10 | a) Model based affine candidates (separate test for different number)  b) Additional control point based affine candidates (separate test for different number) | Huanbang Chen (Huawei) | Fangdong Chen (Hikvision) |

[JVET-K0188](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3697) CE4: Symmetrical mode for bi-prediction (Test 3.2) [H. Chen, H. Yang, J. Chen (Huawei)]

In this contribution, a mode is proposed for motion information coding in bi-prediction. A symmetrical mode flag indicating whether symmetrical mode is used or not is explicitly signalled if the prediction direction is bi-prediction. When the flag is true, the reference index and MVD for list 0 are derived from motion information of list 1. More ever, the reference index of list 1 is inferred as 0 for this mode.

***Key points***

* A symmetrical mode flag is explicitly signalled for each CU in inter bi-prediction mode
* Ref\_idx in L1 is enforced to be 0, Ref\_idx in L1 is inferred to be symmetrical of L0 reference
* MVD in L0 is mirrored from MVD in L1

[JVET-K0198](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3707) CE4: Enhanced Merge Mode (Test 4.2.15) [X. Chen, N. Zhang, J. Zheng (HiSilicon)]

Test A : More spatial candidates

More spatial positions are checked. The extended spatial positions from 6 to 27 are checked according to their numerical order after the temporal candidate. In order to save the MV line buffer, all the spatial candidates are restricted within two CTU lines. That is, the spatial candidates beyond the CTU line above the current CTU line are excluded.

***Key points***

* Additional 6 merge candidates are added
* New candidates are appended to the end of the current merge list
* Redundancy checking is performed for the added merge candidates

**Test B: Merge offset extension**

In Merge offset extension, more candidates are checked based on existed candidate. New candidates has mv offset to the mv of existed candidate (except ATMVP candidate). New extended mv offset candidates are constructed only based on the first candidate of merge candidate list.

***Key points***

* Additional 8 merge candidates are added
* Ordering of the candidates : cross direction offset is checked firstly, and then X direction offsets
* New candidates are appended to the end of the current merge list

**Test C:** **Combined Average Merge Candidates**

* The first 4 candidates in the merge list is used for deriving the new candidates
* The reference picture in L0 and L1 is decided by voting, separately
* Combination of the 4 candidates is performed to derive 6 pair candidates
* For each pair of candidates, L0 and L1 motion vectors are averaged separately

[JVET-K0207](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3716) CE4: Enhanced Chroma Interpolation Filter (Test 7.1) [M. Sychev, G. Zhulikov, T. Solovyev, J. Chen (Huawei)]

[JVET-K0208](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3717) CE4.3: Unicity in motion information candidate lists (tests 4.3.6) [A. Robert, T. Poirier, F. LeLeannec (Technicolor)]

[JVET-K0218](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3728) CE4.1: Affine mode enhancement (Test 4.1.4) [A. Robert, T. Poirier, F. LeLeannec (Technicolor)]

Test 4.1.4.a: Affine motion seed storage

In the JEM, the affine model seeds are stored in the top-left, top-right and bottom-left 4x4 sub-blocks in the considered CU. In the proposed solution, the affine model seeds are stored separately as a motion information associated to the whole CU. The motion model is thus decoupled from the motion vectors used for actual motion compensation at the 4x4 block level. This storage allows preserving the complete motion vector field at the 4x4 sub-block level. It also allows using affine motion compensation for block of size 4 in width or height.

Test 4.1.4.b: Affine motion vector prediction

|  |  |  |
| --- | --- | --- |
|  | JEM | Proposed |
| Motion information predictor composition | inherited MV triplet | inherited MV triplet |
| Candidate list | 2 candidates, among:  - constructed affine models from spatial neighbors {a1, b1, a0, b0, b2}  - classical AMVP candidates | 2 candidates, among:  - spatial Affine positions {b1, a1, a0, b0, b2, **a2, b3**}  - improved constructed affine models from spatial neighbors  - classical AMVP candidates  - zero MVs |
| Signaling | For each reference picture:  - reference picture idx  - predictor idx  - 2 MVDs Affine flag Residual is coded | For each reference picture:  - reference picture idx  - predictor idx  - 2 MVDs Affine flag Residual is coded |

The number of affine merge candidates is up to 8.

[JVET-K0219](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3729) CE4.2: Merge mode enhancement (Test 4.2.4) [F. Galpin, T. Poirier, F. LeLeannec, A. Robert (Technicolor)]

Candidates are added to the list in a pre-defined order:

1. Spatial candidates for blocks 1-4.
2. *Extrapolated affine candidates for blocks 1-4.*
3. *Re-ordering, the bi-prediction ones are inserted before the ones with uni-prediction.*
4. ATMVP (BMS test).
5. *Virtual affine candidate*.
6. Spatial candidate (block 5) (used only when the number of the available candidates is smaller than 6 (or 4 for VTM test)).
7. *Extrapolated affine candidate (block 5).*
8. Temporal candidate (derived as in HEVC).
9. *Non-adjacent spatial candidate followed by extrapolated affine candidate, and extrapolated affine candidate (blocks 6 to 49).*
10. Combined candidates.
11. Zero candidates.

Moreover, for the first four spatial candidates (and extrapolated affine candidates in Test4.2.3.e), the bi-prediction ones are inserted before the ones with uni-prediction.

[JVET-K0228](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3738) CE4-2.1: Adding non-adjacent spatial merge candidates [R. Yu, P. Wennersten, R. Sjöberg (Ericsson)]

***Key points***

* Firstly check vertical positions and then horizontal positions, duplication checking is performed
* Checking in each direction stop if one candidate is found or max distance is reached
* Max 2 new candidates in total
* New candidates are added before the TMVP candidate in the merge candidate list when the number of already added spatial candidates is less than a threshold (3 in currently implementation)

[JVET-K0234](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3744) CE4-3.1: MV prediction between two directions for AMVP mode [R. Yu, P. Wennersten, R. Sjöberg (Ericsson)]

As described in JVET-J0012 section 2.1.7.4, the proposed tool exploits correlation between the two motion vectors for bi-prediction when the current block is coded using the AMVP mode. The tool allows a motion vector in a bi-prediction pair to be scaled and used as a predictor for the second vector. Furthermore, the tool makes the reconstruction order of the bi-prediction pair adaptively depending on how promising the L1 motion vector prediction candidate list is. A promising candidate list here means a list containing a non-scaled prediction candidate.

No change is made to the parsing of the motion vector differences. The reconstruction process of the bi-prediction pair starts with deriving the two candidate lists for L0 and L1 as normal. If the L1 candidate list does not contain a scaled candidate, the L1 vector is determined to be promising and therefore get reconstructed first. After that, if the L0 prediction list contains a scaled candidate, the first scaled candidate in the L0 candidate list is replaced by a scaled version of the L1 vector. Otherwise, if the L1 candidate list does contain a scaled candidate, the L0 vector is reconstructed first. After that, the first scaled candidate in the L1 candidate list is replaced with a scaled version of the L0 vector.

The tool is not enabled when the two reference picture lists contains the same reference pictures with mvd\_l1\_zero\_flag being enabled.

***Key points***

* MV in one direction is used as MVP in the other direction
* The prediction direction (L0->L1 or L1->L0) depends on whether an MVP is a scaled one or not
* Enabled only when mvd\_l1\_zero\_flag being is disabled

[JVET-K0244](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3754) CE4.1.6: MVP pair list construction for affine inter mode [Z.-Y. Lin, T.-D. Chuang, C.-Y. Chen, Y.-W. Huang, S.-M. Lei (MediaTek)]

There are two kinds of MVP pair candidates. One kind is SMVP-pair candidate, and the other kind is CMVP-pair candidate. If the neighboring CUs are coded with affine, SMVP-pair candidates can be generated.

For SMVP-pair candidates, the searching order is A1🡪B1🡪B0🡪A0🡪B2, as shown in Figure 1. If the neighboring CU is coded with affine but the reference picture is different from the target reference picture, the control points’ MVs are scaled to current target picture to derive current CU’s affine model.

If the number of MVP pair candidates is smaller than two after searching SMVP-pair candidates, CMVP-pair candidates will be searched. The neighboring MVs at A0, A1, A2, B0, B1, C0, and C1, as shown in Figure 2, are used to derive CMVP-pair candidates. The first available MV in set A (A0, A1, and A2) and the first available MV in set B (B0 and B1) are used to calculate the first CMVP-pair candidate. The first available MV in set A and the first available MV in set C (C0 and C1) are used to calculate the second CMVP-pair candidate.

***Key points***

* Number of candidates is 2
* Two kinds of candidates : SMVP-pair candidate, and CMVP-pair candidate
* MV scaling is supported in the derivation of both type of candidates

[JVET-K0245](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3755) CE4.2.8: Merge mode enhancement [Y.-L. Hsiao, Z.-Y. Lin, T.-D. Chuang, C.-C. Chen, C.-Y. Chen, C.-W. Hsu, Y.-W. Huang, S.-M. Lei (MediaTek)]

***CE4.2.8.A: Affine merge candidates***

* Two kinds of affine candidates: spatial inherited one, corner derived.
* Spatial inherited affine candidate
  + Only use 6-parameter affine model
  + If available, regular merge candidate in the same position will not be used
* Corner derived affine candidates
  + Only use 6-parameter affine model
  + At most four candidates
  + Inserted in regular merge list after spatial candidate at location B2.
* Number of all candidates is 6/7 for VTM/BMS

***CE4.2.8.B: Middle spatial and multiple temporal candidates***

* Two middle spatial candidates located at ML and MT are added after temporal candidates
* Four temporal candidates are added in the order {RB, CT, LB, RT}
* Number of all candidates is 10

***CE4.2.8.C: Pairwise-average candidates as replacement of HEVC combined candidates***

* Predefined pairs are defined as {(0, 1), (0, 2), (1, 2), (0, 3), (1, 3), (2, 3)}
* Rule for MV averaging
  + If both MVs are available in one list, the MV with the larger merge index is scaled to the reference picture of the merge candidate with the smaller merge index
  + If only one MV is available, use the MV directly
  + If no MV is available, keep this list invalid
* Number of all candidates is 10

***CE4.2.8.D: Double ATMVP candidates with uniformity check and decoder-side speed-up***

* Two MVs from merge candidate list instead of one are used as temporal MVs to generate two ATMVP candidates
* Uniformity check at four corners
  + If the same for the two ATMVP candidates, the second one is removed from the list
  + If the same as the center motion, ATMVP is treated as TMVP
* Number of all candidates is 7

[JVET-K0247](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3757) CE4.3.4: Removal of AMVR flag constraint [C.-Y. Lai, T.-D. Chuang, C.-Y. Chen, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-K0248](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3758) CE4.4.1: Generalized bi-prediction for inter coding [Y.-C. Su, T.-D. Chuang, C.-Y. Chen, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-K0279](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3792) CE4: Restricted merge (Test 4.2.2) [M. Winken, H. Schwarz, D. Marpe, T. Wiegand (HHI)]

In merge mode (but not SKIP mode) in B slices, three types of merge list is constructed

* Regular merge list
* Merge list with uni-prediction candidate in L0 direction
* Merge list with uni-prediction candidate in L1 direction

The syntax element merge\_ref\_pic\_list\_idc can take three different values for the indication.

[JVET-K0286](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3799) CE4: Additional merge candidates (Test 4.2.13) [J. Ye, X. Li, S. Liu (Tencent)]

[JVET-K0337](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3854) CE4.1.3: Affine motion compensation prediction [Y. Han, H. Huang, Y. Zhang, C.-H. C.-H. Hung, C.-C. Chen, W.-J. Chien, M. Karczewicz (Qualcomm)]

***Test 4.1.3.a:*** ***Affine mvp construction***

Three types of affine mv predictor set

* Inherited affine mv predictor set from neighboring blocks in affine mode, if the same reference picture is referred to.

Up to two different affine MV predictor sets are derived from affine motion of the neighboring blocks. Neighboring blocks A0, A1, B0, B1, and B2. If the neighboring block is coded using affine motion model and its reference frame is same as the reference frame of the current block, MVs at two (for 4-parameter affine model) or three (for 6-parameter affine model) control points of the current block are derived from the affine model of this neighbor.

* Constructed virtual affine mv predictor set from neighboring mv, if the same reference picture is referred to.

The neighboring MVs are divided into three groups: , and . is the first MV in *S*0 that refers to the same reference picture as the current block; is the first MV in *S*1 that refers to the same reference picture of the current block; and is the first in *S*2 that refers to the same reference picture of the current block.

* If only and can be found, is derived as:

where the current block size is *W*×*H*.

* If only and can be found, is derived as:

* HEVC affine mv predictor

***Test 4.1.3.b: Affine mv prediction***

Two candidate sets with two (three) candidates are used to predict two (three) control points of the affine motion model. Given motion vector difference vectors, , the control points are calculated:

***Test 4.1.3.c&d: 4-param/6-param affine model switching at CU/slice level***

The affine models are adaptively selected at both slice level and block level. Slice header flags affine4\_flag and affine6\_flag indicate whether the 4-parameter affine and 6-parameter affine model are applied. If one of them is equal to 1, affine\_flag signalled at block level indicates whether the allowed affine model is used for this block. If both are equal to 1, there is an affine\_type flag signalled at block level. If affine\_type is equal to 1, the 6-parameter affine model is used, otherwise the 4-parameter affine model is applied. To determine the affine4\_flag and affine6\_flag for the current slice the statistics from the previous coded slice are used.

[JVET-K0338](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3855) CE4.2.16: Sub-block merge candidates in BMS and JEM [Y. Han, Y. Zhang, C.-H. Huang, C.-C. Chen, W.-J. Chien, M. Karczewicz (Qualcomm)]

[JVET-K0339](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3856) CE4.2.3: Improvement on Merge/Skip mode [Y. Han, H. Huang, Y. Zhang, C.-H. Huang, C.-C. Chen, W.-J. Chien, M. Karczewicz (Qualcomm)]

Candidates are added to the list in a pre-defined order:

1. Spatial candidates for blocks 1-4.
2. *Extrapolated affine candidates for blocks 1-4.*
3. *Re-ordering, the bi-prediction ones are inserted before the ones with uni-prediction.*
4. ATMVP (BMS test).
5. *Virtual affine candidate*.
6. Spatial candidate (block 5) (used only when the number of the available candidates is smaller than 6 (or 4 for VTM test)).
7. *Extrapolated affine candidate (block 5).*
8. Temporal candidate (derived as in HEVC).
9. *Non-adjacent spatial candidate followed by extrapolated affine candidate, and extrapolated affine candidate (blocks 6 to 49).*
10. Combined candidates.
11. Zero candidates.

Moreover, for the first four spatial candidates (and extrapolated affine candidates in Test4.2.3.e), the bi-prediction ones are inserted before the ones with uni-prediction.

[JVET-K0486](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4016) Cross-check of JVET-K0339: CE4.2.3 related: Improvement on Merge/Skip mode (4.2.3f, 4.2.3g, 4.2.3h) [T. Ikai (Sharp)] [late]

[JVET-K0341](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3858) CE4.2.5: Simplifications on advanced temporal motion vector prediction (ATMVP) [X. Xiu, Y. He, Y. Ye (InterDigital)]

***CE4.2.5.a Simplified collocated block derivation with one fixed collocated picture***

* Using the same collocated picture as in HEVC for ATMVP derivation is signalled at the slice header
* The scaled MV is used in ATMVP if the original MV from a neighboring block points to a reference picture other than the collocated picture.

***CE4.2.5.b Adaptive ATMVP sub-block size***

Slice-level adaptation of the sub-block size for the ATMVP motion derivation

* One default sub-block size is signalled at sequence level
* One flag is signalled at slice-level to indicate if the default sub-block size is used for the current slice
* If the flag is false, the corresponding ATMVP sub-block size is further signalled in the slice header for the slice.

[JVET-K0349](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3866) CE4-2.11: MVPlanar prediction [S. Iwamura, S. Nemoto, A. Ichigaya (NHK)]

***Key points***

* The interpolation is carried out in a similar way as intra planar prediction.
* In the case that neighboring CU is intra coded, the closest neighboring MV is substituted in a similar way as intra reference sample substitution.
* If the mv\_palanr\_flag is equal to 1, inter prediction index “inter\_pred\_idc” syntax and reference index “ref\_idx” syntaxes for both L0 and L1 lists are additionally signalled.

[JVET-K0355](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3873) CE4.2.12 Affine merge mode [Y. Wang, X. Fan, D. Zhao, Y. Li, D. Liu, F. Wu (USTC)]

***CE4.2.12.a BMS affine merge modification***

Instead of finding the first neighboring block with affine mode, the affine model from the neighboring coding unit with largest size is used.

***CE4.2.12.b Complex merge mode***

Three types of motion model, all are constructed from MV of neighboring blocks

* Bilinear model
* 6-param affine model
* 4-param affine model

Candidates are added to the list in the order

1. Affine (CP2, CP3)
2. Affine (CP1, CP3)
3. Affine (CP1, CP2, CP3)
4. Affine (CP1, CP2)
5. Affine (CP2, CP4);
6. Affine (CP3, CP4);
7. Affine (CP1, CP4);
8. Bilinear
9. Affine (CP1, CP2, CP4);
10. Affine (CP2, CP3, CP4);
11. Affine (CP1, CP3, CP4);

The reference index with the highest utilization rate among all the reference indices is selected as the final reference index. The control points with different reference indices are scaled to the final reference index.

Redundancy checking is performed.

All candidates are put in a separate list.

Number of candidates are 8, and FL 3 bins are used for index coding.

[JVET-K0357](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3875) CE4.3.3: Locally adaptive motion vector resolution and MVD coding [Y. Zhang, Y. Han, C.-C. Chen, C.-H. Hung, W.-J. Chien, M. Karczewicz (Qualcomm)]

Motion vector differences (MVDs) can be signalled in different precision to allow flexible MVD coding for video sequences with different resolution. In JEM and BMS, MVD could be signalled either in units of quarter luma sample, integer luma sample or four luma samples.

A variable length MVD resolution flag (0 to 2 bits) is conditionally signalled in CU level for those that have at least one non-zero MVD components with the first bit identifying whether quarter luma sample MVD precision is used. When the first bit (equal to 1) indicates that quarter luma sample MVD precision is not used, a second bit is signalled to indicate if integer luma sample MVD precision or four luma samples MVD precision is used.

When a zero is signalled for the first bit of the MVD resolution flag, quarter luma sample MVD resolution is used. When the MVD resolution flag is not signalled (which means both MVDs for reference list 0 and reference list 1 are zero), quarter luma sample MVD resolution is inferred. In the other cases when integer-luma sample MVD precision or four luma samples MVD precision is adopted, the MVP candidates in the AMVP candidate list is rounded to the corresponding precision.

The scheme of MVD coding context is modified in the proposed MVD coding method so that the binarization and context modeling are dependent on the MVD precision and the POC distance between the current frame and the reference frame.

[JVET-K0363](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3881) CE4.5.2: Motion compensated boundary pixel padding [Y. Zhang, Y. Han, C.-C. Chen, C.-H. Hung, W.-J. Chien, M. Karczewicz (Qualcomm)]

## CE5: Arithmetic coding engine (9)

Contributions in this category were discussed Wednesday 11 July 1820–2000 (chaired by GJS).

[JVET-K0025](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3653) CE5: Summary report of CE on Arithmetic Coding Engine [T. Nguyen, A. Said]

This report summarizes the experimental results and the findings for the Core Experiment 5 on Arithmetic Coding Engine. Twelve experiments have been conducted: two experiments in the main category (experiments 5.1.1 and 5.1.2), four experiments in Subset A (experiments 5.2.1 – 5.2.4), four experiments in Subset B (experiments 5.3.1 – 5.3.4), and two experiments in Subset C (experiments 5.4.1 and 5.4.2). The experimental results indicate that further analysis is necessary on the topic of probability estimators and its memory requirements. Furthermore, the results show that a final rLPS design that has a maximum size equal to or less than 2048 bit is sufficient to achieve the compression efficiency.

**Main Category**

The proposals in the main category indicate the best performance in compression efficiency that can be achieved for the time of this document, considering the increase in complexity. There were only two experiments in this sub-category, and they share the same basic design: (a) high-precision probability estimation; (b) double window probability estimation using different window pairs per context; (c) small multiplication tables; and (d) Context probability initialization from previous frames.

The results below summarize the performance of Experiment 5.1.1 (JVET-K0381).

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **5.1.1** | **Over VTM-1.0** | | | | | **Over BMS-1.0** | | | | |
|  | Y | U | V | EncT | DecT | Y | U | V | EncT | DecT |
| **AI** | -0.97% | -0.45% | -0.41% | 103% | 103% | -1.04% | -0.30% | -0.36% | 103% | 99% |
| **RA** | -1.02% | -0.14% | -0.28% | 102% | 101% | -1.17% | -0.21% | -0.44% | 103% | 100% |
| **LB** | -0.87% | 1.06% | 1.11% | 102% | 103% | -1.23% | 0.93% | 1.47% | 101% | 100% |

and for experiment 5.1.2 (JVET-K0283)

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **5.1.2** | **Over VTM-1.0** | | | | | **Over BMS-1.0** | | | | |
|  | Y | U | V | EncT | DecT | Y | U | V | EncT | DecT |
| **AI** | -0.98% | -0.66% | -0.68% | 106% | 103% | -0.84% | -0.47% | -0.58% | 105% | 107% |
| **RA** | -1.00% | -0.51% | -0.48% | 103% | 101% | -0.98% | -0.39% | -0.64% | 104% | 104% |
| **LB** | -1.17% | 0.44% | 0.55% | 103% | 102% | -1.32% | 0.57% | 0.82% | 102% | 100% |

**Performance:** Both experiments yield an improvement in coding efficiency of about 1% in BD-rate in the AI and RA configurations. Slight differences can be observed, but with deltas between the two proposals smaller than 0.2%. Both the encoding and the decoding run times are similar with an increase of about 3% for the encoder and 1-2% for the decoder. Experiment 5.1.2 introduces new context initialization values. The difference in the performance is not significant, i.e., an improvement of about 0.05% in BD-rate can be observed with new initialization values.

**Notes:** Both experiments include the usage of context model states from the previously coded frame.

**Observation:** An improvement in BD-rate of about 0.9 – 1.3% can be achieved by modifying all parts of the arithmetic coding engine. A slight increase in run times can be observed. The performance can be achieved with a rLPS table size similar to the reference (compare experiments 5.1.1 and 5.1.2, and 5.2.1). The new initialization values do not improve the performance significantly (see detailed results for experiment 5.1.2).

**Memory:** As the reference, CABAC, as specified in HEVC, has a 64×4×8 (2048 bit) rLPS table, and each context model requires 7 bits. Moreover, a 64×6 (384 bit) transition table is necessary. Experiment 5.1.1 employs a 16×16×8 (2048 bits) rLPS table and custom window sizes for each context model. Each context model requires 2×15+4 (34 bits). The experiment 5.1.2 employs a 32×8×8 (2048 bits) rLPS table and custom window sizes for each context model, resulting in 2×15+4 (34 bits). For both experiments, each context model requires about 5.14 times the memory of the reference. The following table summarizes the memory requirements with all values in number of bit. The numbers are derived as follows:

* Number of context models: 359, by analyzing the BMS-1.0 software implementation
* Initialization values: 8×3×359=8616 (8-bit initialization values, three slice types, 359 context models)
* Total number of context memory: number of context models multiplied by memory per context
* CABAC in HEVC: Transition table for the states 64×6 (384 bit)
* Experiment 5.1.1: 4 bit for each context model specifying custom window size (1436 bit in total)
* Experiment 5.1.2: same as 5.1.1, but separately for each slice type (4308 bit in total)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Configuration** | **rLPS** | **stateTable** | **cwInit** | **ROM** | **perCtx** | **RAM** |
| HEVC | 2048 | 384 | 0 | 11048 | 7 | 2513 |
| 5.1.1 | 2048 | 0 | 1436 | 12100 | 34 | 12206 |
| 5.1.2 | 2048 | 0 | 4308 | 14972 | 34 | 12206 |

**Subset A and B: Probability Estimation and Derivation of Sub-Interval Range for LPS**

The idea of these two subsets is the evaluation of the proposed probability estimators (PE) and the modified rLPS tables separately. It is however not always possible to decouple the derivation of the sub-interval range for LPS from the probability estimation. The table below, therefore, summarizes the results of the experiments of both subsets.

|  |  |  |  |
| --- | --- | --- | --- |
| **Experiment** | **Document / Proponent** | **Description** | **Results (AI/RA/LB-VTM/BMS)** |
| CE5-2.1 | JVET-K0380  Qualcomm | PE fix, rLPS variation PE: 2×15+4 (34)  **Observation:** Smaller rLPS table sizes can be realized, but requires more operations for the access.  **Memory:** 1. Configuration: 10564 ROM / 12206 RAM 2. Configuration: 12100 ROM / 12206 RAM | **rLPS: 8×8×8 (512)** -0.95% -0.86% -0.61% -1.02% -X.XX% -0.92%  **rLPS: 16×16×8 (2048)** -0.97% -0.88% -0.64% -1.04% -1.00% -0.95% |
| CE5-2.2 | JVET-K0283 HHI | rLPS fix, PE variation rLPS: 32×8×8 (2048)  **Observation:** The memory requirement of multi-parameter PE with custom window sizes can be further reduced (see the second set of results).  **Memory:** 1. Configuration: 10664 ROM / 8616 RAM 2. Configuration: 14972 ROM / 10052 RAM 3. Configuration: 10664 ROM / 10770 RAM 4. Configuration: 14972 ROM / 12206 RAM | **PE: 10+14 (24)** -0.74% -0.64% -0.50% -0.86% -0.81% -0.89%  **PE: 10+14+4 (28)** -1.02% -1.00% -0.81% -0.94% -1.03% -1.10%  **PE: 15+15 (30)** -0.76% -0.66% -0.55% -0.86% -0.80% -0.86%  **PE: 15+15+4 (34)** -1.03% -0.99% -0.80% -0.91% -1.00% -1.11% |
| CE5-2.3 | JVET-K0282 DJI and Peking University | rLPS: 512×64×9 (294912) PE: the introduction of a counter, initial values of the MP are kept when the counter is below a threshold, equal to 2×15+10 (40) per context model.  **Observation:** Results worse when compared to 5.1.1, especially for RA and LB.  **Memory:** 14360 RAM | -0.66% 0.13% 0.22% -0.53% 0.05% 0.04% |
| CE5-2.4 | JVET-K0170 Samsung | rLPS: 512×64×9 (294912) PE: Different update strategy for the MP. Introduction of a counter, only short window is updated when the counter is below a threshold, equal to per 2×15+5 (35) context model.  **Observation:** Performs similar to the configuration without custom window sizes but requires 1 bit less memory per context model.  **Memory:** 12565 RAM | -0.81% -0.56% -0.39% -0.89% -0.64% -0.54% |
| CE5-3.1 | JVET-K0383 Qualcomm | PE fix, rLPS variation PE: 2×15 (CEM1)  **Observation:** In comparison to 5.2.1, the “virtual” performance of custom window sizes can be derived. The improvement is about 0.5% in BD-rate for RA and of about 0.4% in LB. Not clear whether the drop in performance is due to initialization. For AI, the delta is 0.3% in BD-rate.  **Memory:** 1. Configuration: 12100 ROM / 10770 RAM 2. Configuration: 10564 ROM / 10770 RAM | **Relative to CABAC engine 1 of BMS-1.0**  **rLPS: 16×16×8 (2048)**  0.00% 0.03% 0.03%  0.02% 0.0x% 0.0x%  **rLPS: 8×8×8 (512)**  0.03% 0.06% 0.07%  0.04% 0.0x% 0.0x% |
| CE5-3.2 | JVET-K0283 HHI | PE fix, smaller rLPS table tested. PE: 10+14+2x3 (30 bit)  **Observation:** Similar to experiments 5.1.1 and 5.3.1, an even smaller rLPS table performs close to larger rLPS table sizes.  **Memory:** 13948 ROM / 10770 RAM | **Relative to BMS-1.0**  **rLPS: 32×8×8 (2048)** -0.99 -0.96 -0.77 -0.90 -0.98 -1.05  **rLPS: 16×8×8 (1024)** -0.99 -0.96 -0.77 -0.90 -0.98 -1.05 |
| CE5-3.3 | JVET-K0249 MediaTek | PE fix, smaller rLPS table tested. PE: 2×15 (CEM1)  **Observation:** Performs similar to experiment 5.3.1 with a slightly lower performance for RA and LB.  **Memory:** 1. Configuration: 9640 ROM / 10770 RAM 2. Configuration: 10664 ROM / 10770 RAM | **Relative to CABAC engine 3 of BMS-1.0**  **rLPS: 16×8×8 (1024)**  0.04% 0.03% 0.09%  0.05% 0.03% 0.06%  **rLPS: 32×8×8 (2048)**  0.01% 0.00% 0.04%  0.02% -0.01% 0.02% |
| CE5-3.4 | JVET-K0249 MediaTek | PE fix, smaller rLPS tables, derivation requires multiplications PE: 2×15 (CEM1)  **Observation:** Performs similar to experiment 5.3.3 with even smaller rLPS table sizes.  **Memory:** 1. Configuration: 8816 ROM / 10770 RAM 2. Configuration: 8904 ROM / 10770 RAM | **Relative to CABAC engine 3 of BMS-1.0**  **rLPS: 5×5×8 (200)**  0.04% 0.03% 0.09%  0.05% 0.03% 0.06%  **rLPS: 6×6×8 (288)**  0.01% 0.00% 0.04%  0.02% -0.01% 0.02% |

The conclusion from the rLPS configuration tests is that we can use tables that are about as small as those used for HEVC without a significant loss.

For the probability estimator, double windows were proposed. Most of the gain comes from using a double window without a customized window size.

It was suggested to use a double window with a fixed window size and to use a high-precision range computation as the basic thing to compare against.

A suggestion was a 32×8 table. 512x64 was used in the JEM – use that.

Further study in CE was requested, with a straightforward double window anchor.

This was further discussed Thursday 1445 (chaired by GJS).

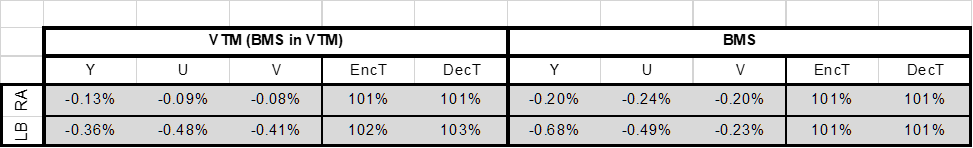
We would like a specific design that is straightforward, well understood and at least reasonably tested, in which we’re confident that there are no unnecessary elements, before acting.

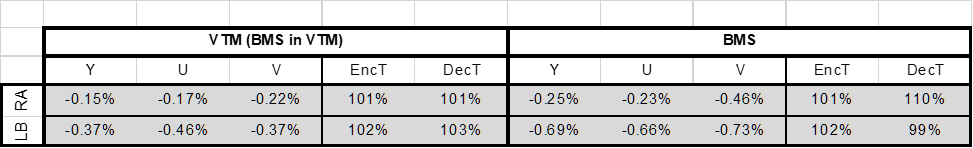
It was commented that we need to make sure the design has high throughput, not just good coding efficiency.

**Subset C: Context Initialization and Parameterization**

This subset of the CE was discussed at 1400-1445 on Thursday 12 July (chaired by GJS).

The final subset evaluates two experiments employing techniques that allow the inheritance of context model states from previously coded frames. The difference between the two proposals is the granularity level of the state inheritance, either slice level (7.4.1) or CTU-line level (7.4.2). The latter requires more memory but has benefit for parallel processing applications.





Summary: The improvement in compression efficiency is relatively small for RA configuration, whereas the improvement for LB-BMS is almost 0.7% for both schemes. It seems that the LB configuration is generally prone to initialization values.

It was noted that the method of context probability initialization might differ, and that these might have been trained on the test set, or even on the low-resolution class (which would appear to have the most benefit). It was suggested that, properly, there should be a specific method of initializing the values.

It was commented that the overall benefit of initialization is pretty small, especially for high-resolution video. A participant said that initializing the coefficient coding probabilities to 0.5 results in a loss of only 0.15% in RA and 0.05 in AI and 0.54% in LB (esp. higher in LB because that average does not include the UHD sequences, and the average number of coded bins per frame is lower due to having fewer bits per frame).

It was asked how it was determined where to inherit the contexts from. For the experiments the last coded frame with the same QP was used and only one slice per frame was used. No proposals had been made about how this would work in the actual standard.

No memory analysis had been provided for the proposals.

Further study was needed to determine how this feature would really work in the standard.

[JVET-K0170](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3678) CE5: Counter-based probability estimation (Test 2.4) [K. Choi, Y. Piao, C. Kim (Samsung)]

[JVET-K0249](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3759) CE5.3.3 & CE5.3.4: CABAC range sub-interval derivation [T.-D. Chuang, C.-Y. Chen, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-K0282](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3795) CE5: Context adaptive counter-based probability estimation (test 2.3.0) [J. Cui, S. Wang, S. Ma (Peking University), X. Zheng (DJI)]

[JVET-K0283](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3796) CE5: Counter-based probability estimation and changes to the arithmetic coding engine (Tests 1.2, 2.2, 3.2 and 4.2) [J. Stegemann, H. Kirchhoffer, D. Marpe, H. Schwarz, T. Wiegand (HHI)]

[JVET-K0379](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3898) CE5: CABAC probability initialization from previous inter frames (test C1) [A. Said, H. Egilmez, Y.H. Chao, M. Karczewicz, V. Seregin (Qualcomm)]

[JVET-K0380](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3899) CE5: Per-context CABAC initialization with double-windows (test A1) [A. Said, H. Egilmez, Y-H. Chao, M. Karczewicz, V. Seregin (Qualcomm)]

[JVET-K0381](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3900) CE5: Combined Arithmetic Coding Tools (test CE 5.1) [A. Said, H. Egilmez, Y-H. Chao, M. Karczewicz, V. Seregin (Qualcomm)]

[JVET-K0383](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3902) CE5: Binary Arithmetic Coding Range Update with Small Table or Short Multiplications (test B1) [A. Said, H. Egilmez, Y-H. Chao, M. Karczewicz, V. Seregin (Qualcomm)]

## CE6: Transforms and transform signalling (26)

Contributions in this category were discussed Thursday 12 July 1500–XXXX (chaired by GJS).

[JVET-K0026](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3776) CE6: Summary Report on Transforms and Transform Signalling [A. Said, X. Zhao]

This contribution summarizes the activities of Core Experiment (CE) on Transforms and Transform Signalling, including list of participants and experiments, summary of test results, brief experiment descriptions, and conclusions. The CE studies were divided into three categories: (1) Primary transforms, with 14 proposals; (2) Secondary transform, with 6 proposals; and (3) and transform combinations and signalling, which has 4 proposals. Unlike other CEs, in this CE all experiments were done using software based on the BMS-1.1 version, and results are reported using the corresponding anchors.

Core experiments (CEs) are organized according to the following three categories:

* Primary transforms: CEs 6.1.1-6.1.14 test proposals that introduce changes for primary transforms only.
* Secondary transforms: CEs 6.2.1-6.2.6 test proposals that introduce changes for secondary transforms only.
* Combinations and Signalling: CEs 6.3.1-6.3.4 test proposals that introduce changes for both primary and secondary transforms, and how they can be combined.

The following table lists all the experiments in each category, and the corresponding input document to the Ljubljana meeting.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **CE 6.1 – Primary Transforms** | | | | |
| **CE #** | **Proponent** | **Related Docs.** | **Summary of the tool** | **Cross-checkers** |
| 6.1.1 | A. Said, H.E. Egilmez, Y.-H. Chao, V. Seregin, M. Karczewicz (Qualcomm) | JVET-K0381 | **Primary transformation:** In addition to transform types of AMT in JEM7, DST-4 and Identity transform are also used for intra prediction residual. | 1. Orange  2. Samsung  3. Brightcove  4. Wilus |
| 6.1.2 | A. Said, H.E. Egilmez, Y.-H. Chao, V. Seregin, M. Karczewicz (Qualcomm) | JVET-K0272 | **AMT complexity reduction:** The AMT transforms are implemented using the DCT-2 family of transforms and low-complexity adjustment stages. | 1. Orange  2. Technicolor  3. ETRI  4. Samsung  5. Sony (withdrawn)  6. Tencent  7. Brightcove |
| 6.1.3 | K. Naser, F. Le Léannec, E. François (Technicolor) | JVET-K0264 | **Transform Reduction:** It is proposed to reduce transform types in AMT to DCT-8, DST-4 and DST-7. DST-7 is implemented based on the DCT-8 with flipping and sign changes. | 1. Brightcove  2. FastVDO (withdrawn) |
| 6.1.4 | S.-C. Lim, J. Kang, H. Lee, J. Lee, S. Cho, H. Y. Kim (ETRI), N.-U. Kim, Y.-L. Lee (Sejong Univ.), D.-Y. Kim, W. J. Jeong (Chips&Media) | JVET-K0167 | **Residual flipping**: It is proposed to replace the AMT with DCT-II and DST-VII using four types of residual flipping. | 1. Qualcomm  2.  3. |
| 6.1.5 | K. Choi, M. Park, C. Kim (Samsung) | JVET-K0171 | **Reduced AMT transform types:** The AMT transforms in JEM7 are replaced with a reduced set of transforms using DCT-8 and DST-7. | 1. Qualcomm  2. Brightcove  3. |
| 6.1.6 | K. Choi, M. Park, C. Kim (Samsung) | JVET-K0173 | **Primary transformation restriction:** When block width or height is over 32 pixels, AMT is not used or signalled. | 1. Qualcomm  2. Tencent  3. |
| 6.1.7 | T. Tsukuba, M. Ikeda, T. Suzuki (Sony) | JVET-K0121 | **Adaptive Multiple core Transforms for Luma and Chroma:** It is proposed to apply AMT, which includes transform matrix replacement, for both luma and chroma. | 1.Tencent |
| 6.1.8 | T. Tsukuba, M.I keda, T. Suzuki (Sony) | JVET-K0123 | **Transform Matrix Replacement:** It is proposed to replace transform matrix of DCT-8, DST-1 and DCT-5 used in JEM with that of alternate transforms as below:   * Flipped DST-7, which is derived by flipping DST-7 without sign changes, is used instead of DCT8. * DST-6, which is derived by transposing DST-7, is used instead of DST-1.   DCT-2 is used instead of DCT-5. | 1.Brightcove  2. Sejong Univ. |
| 6.1.9 | K. Kawamura, Y. Kidani, S. Naito (KDDI) | JVET-K0397 | **Shrink Transform:** When the transform block size is equal or larger than 64, inverse transform is realized by two steps: inverse transform with half size and up-sampling. | 1. MediaTek  2. Panasonic  3. |
| 6.1.10 | M.-S. Chiang, Z.-Y. Lin, T.-D. Chuang, C.-Y. Chen, C.-W. Hsu, Y.-W. Huang, S.-M. Lei (MediaTek) | JVET-K0250 | **Adaptive multi-transform (AMT) with constraints:** AMT transform type is reduced to DCT-II, DST-VII and flipped DST-VII. AMT is only allowed when both CU width and height are smaller than or equal to 64. For some block sizes, the number of AMT candidates is reduced. | 1.Brightcove  2.  3. |
| 6.1.11 | M. Koo, M. Salehifar, J. Lim, S. Kim (LGE) | JVET-K0096 | **Primary transform:** Only DST7 and DCT8 are used for AMT. AMT is applied up to 32. DST7 and DCT8 are implemented using FFT | 1. Brightcove 2. Samsung |
| 6.1.12 | Y. Zhao, H. Yang, J. Chen (Huawei) | JVET-K0139 | **Spatial varying transform:** For an inter-predicted CU, a TU of half size of the CU may be used, and in this case the TU shape and position information are signalled. | 1. MediaTek  2. bcom |
| 6.1.13 | Y. Lin, M. Mao, S. Song, J. Zheng, J. An, C. Zhu, (HiSilicon) | JVET-K0125 | **Prediction dependent transform**   1. Intra prediction mode dependent transform for luma and chroma 2. Position dependent transform for residual of HEVC merge mode | 1. Brightcove |
| 6.1.14 | X. Zhao, X. Li, S. Liu (Tencent) | JVET-K0083 | **Block size dependent zero-out transform**: If the coding block size is 128, only the first 16 coefficients are used (anchor uses 32) the remaining coefficients are considered as 0. Otherwise the same as the anchor. | 1. Sony |
| **CE 6.2 – Secondary Transforms** | | | | |
| **CE #** | **Proponent** | **Related Docs.** | **Summary of the tool** | **Cross-checkers** |
| 6.2.1 | Fabrice Urban, Fabrice Le Léannec, Edouard Francois (Technicolor) | JVET-K0271 | **Secondary transformation:** NSST can be applied only when AMT transform flag or index is equal to 0. NSST transform choice is separately signalled for Cb and Cr chroma components. | 1. Samsung  2. HHI  3. |
| 6.2.2 | A. Said, H.E. Egilmez, Y.-H. Chao, V. Seregin, M. Karczewicz (Qualcomm) | JVET-K0374 | **Hierarchically Structured Matrix-based Transforms:** HyGT based NSST is replaced by hierarchically structured matrix-based transforms (HSMTs). | 1. Technicolor  2. Samsung  3. HHI  4. Tencent  5. Brightcove  6. bcom |
| 6.2.3 | K. Choi, M. Park, C. Kim (Samsung) | JVET-K0174 | **Secondary transformation:** NSST is modified with reduced sets based on grouping intra prediction modes and signalling with on/off flag with 4 modes. | 1. Qualcomm  2. HHI  3. Tencent  4. |
| 6.2.4 | X. Zhao, X. Li, S. Liu (Tencent) | JVET-K0084 | **Matrix multiplication based NSST**: It is proposed to use matrix multiplication based NSST. | 1. LG Electronics |
| 6.2.5 | M. Koo, M. Salehifar, J. Lim, S. Kim (LGE) | JVET-K0098 | **4x4 Secondary transform:** 4x4 Layered Givens Transform (LGT) is tested. | 1. Brightcove |
| 6.2.6 | M. Salehifar, M. Koo, J. Lim, S. Kim (LGE) | JVET-K0099 | **8x8 Secondary transform:**  8x8 Reduced Secondary Transform (RST) keeping 16 bases and 32 of the 64-dimensional space (16x64 and 32x64 matrix) with NSST index modification based on RST modification. | 1. Huawei  2. bcom |
|  |  |  |  |  |
| **CE 6.3 – Combinations and Signalling** | | | | |
| **CE #** | **Proponent** | **Related Docs.** | **Summary of the tool** | **Cross-checkers** |
| 6.3.1 | X. Zhao, X. Li, S. Liu  (Tencent) | JVET-K0085 | **Coupled primary and secondary transform**: AMT and NSST is coupled and signalled by only one transform index and applied for both luma and chroma components. | 1. Technicolor  2. Samsung  3. Sony  4. Brightcove |
| 6.3.2 | M. Siekmann, C. Bartnik, H. Schwarz, D. Marpe, T. Wiegand (HHI) | JVET-K0305 | **Set of transforms:** A set of 5 predefined transform candidates, and each candidate specifies a primary horizontal transform, a primary vertical transform, and a non-separable secondary transform. The transform candidates depend on the block size and the prediction mode. | 1. Qualcomm  2.Brightcove |
| 6.3.3 | K. Abe, T. Toma (Panasonic) | JVET-K0127 | **AMT and NSST complexity reduction:** (1) NSST is used only when DCT2 is used as primary transform. (2) DST7 is used as the horizontal and vertical transform without signalling when width or height is less than or equal to 4. | 1. KDDI |
| 6.3.4 | M.-S. Chiang, Z.-Y. Lin, T.-D. Chuang, C.-Y. Chen, C.-W. Hsu, Y.-W. Huang, S.-M. Lei (MediaTek) | JVET-K0250 | **AMT and NSST syntax signalling:** The NSST index signalling depends on luma/chroma, number of nonzero DC/AC coefficient of upper-left 8x8 or 4x4. | 1. Bcom |

**Primary transforms**

It was noted that the BMS has transform design differences relative to the VTM, i.e., adaptive primary transform selections and also secondary transforms. It may be difficult to consider them all together.

Key aspects:

* Transform size
* How many transforms
* Computation of transform (esp. not a large DST7)

Focused discussions:

* 6.1.11.1 (combined aspects, not 64-length additional transform, and DCT2+DCT8+DST7, fast form of DST7),
* 6.1.5.1 (reduction of transform types - BMS has 5, this has 3: DCT2+DCT8+DST7)
* 6.1.6.1 (reduces length but not transform types)
* 6.1.4.1 (only DCT2 and DST7)
* 6.1.8.1 (DCT2 and DST7)
* 6.1.10.1 (DCT2 and DST7)

DCT8 is similar in concept to flipped DST7.

Note that all these mix and match horizontal and vertical types

Decision:

* Adopt AMT (both intra and inter, each controlled by an SPS flag) as follows (approx 3.3%, RA 2.0%, LB 1.3%):
* No 64-length DST7 and DCT8 (no AMT syntax sent when either dimension is larger then 32)
* No 128-length DCT2
* Only DCT2, DST7 and DCT8
* All transforms are to have 10 bit coefficients
* Uses the syntax that has been in the BMS. AMT is applied only for luma. There are separate enabling flags for intra and inter at SPS level. When AMT is enabled, then
  + If CBF=1, then
    - A flag for DCT2 in both directions; if not then
      * If (intra and the number of nonzero coefficients is greater than two) or inter (regardless of the number of nonzero coefficients)
        + Flag for horizontal is DST7 vs. DCT8
        + Flag for vertical is DST7 vs. DCT8
      * Otherwise (intra block with only 1 or 2 nonzero coefficients), DST7 is used both horizontally and vertically
* Rather than AMT, suggested name is multiple transform selection (MTS)

Further discussion was planned to be held regarding whether to disable inter AMT in CTC (40% runtime, 0.5% coding gain).

This CTC issue was further discussed in plenary (see section 12.1) and in Track A on Tuesday 0900 (chaired by GJS). The runtime impact was said to be about 15%, not 40%, when considering only the inter aspect and the block size restriction, as reported for a CE test in K0173. The syntax description above was corrected. However, there was some doubt about the runtime impact, and it was agreed to disable AMT for inter CUs in the CTC.

6.1.12 (K0139) has a scheme where inter AMT operates such that the encoder chooses whether the transform covers the whole block with a DCT2 or covers only part of the block and indicates a selection among 6 positions which part of the block is covered by a DST7 or DCT8 or DST1. (It was commented that this seems similar to a further CU split or to introducing a TU segmentation.) The tested encoder uses a fast optimization to determine. There is also a syntax customization to skip syntax in cases the encoder does not evaluate. Further study of that is needed.

6.1.14 ( JVET-K0083) Block size dependent zero-out transform: If the coding block size is 128, only the first 16 coefficients are used (anchor uses 32) the remaining coefficients are considered as 0. Otherwise the same as the anchor. This contribution was further discussed Friday 13 July 1820 (chaired by GJS). No reduction in encoder and decoder runtimes was reported. It did not seem desirable to introduce this special case.

Further study on the method of computing the transforms and potential other aspects like tiling large regions with small transforms.

**Secondary transforms**

It was reported that the expected gain from a 4×4 nonseparable secondary transform on top of AMT is roughly 1.0% gain in RA, about 1.6% in AI, with roughly 1.3× encoder time. (Going to 8x8 would provide more gain, but seems impractical.)

Most of the tested methods used ~105 selectable 4x4 secondary transform matrices.

6.2.3 has a somewhat smaller number of matrices.

6.2.4.2 has about 35 matrices.

A participant remarked that doing this on 4x4 blocks is seems highly questionable. This would be 16 multiply-accumulates per sample.

Overall it seemed questionable whether the gain is worth the added stage, memory, and decoder processing (and encoding time).

Plan to continue CE work to consider what can be done.

Secondary transforms were further discussed Saturday 14 July (GJS) 0920.

The CE summary report included measures of the relative gain of methods of modifying the secondary transform relative to the 4x4 NSST that was in the BMS, not on the gain of the secondary transform on top of AMT.

It was agreed to keep secondary transforms in the BMS and conduct a CE to study ways of improving its complexity-coding efficiency tradeoff.

The training method for designing the secondary transforms included training on the test set. For the CE, the proposals will not be trained on the CTC sequences.

**Combinations and signalling**

6.3.1 proposes to couple a primary transform combination to a specific secondary transform. This is said to provide about 0.9% gain over AMT in AI and 0.5% gain in RA. (Going to 8x8 would provide more gain.) The number of selectable secondary transforms is the same.

6.3.2 proposes a somewhat different coupling.

6.3.3 proposes to use the secondary transform only when a DCT2 is chosen and does not use DCT2 in small blocks.

6.2.1 is somewhat similar to 6.3.3.

6.3.4 proposes to disallow AMT if the number of nonzero coeffs is large.

Variations are proposed depending on whether 8x8 is included.

Plan to continue CE work to consider what can be done.

[JVET-K0083](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3586) CE6: Block size dependent zero-out transform (Test 1.14) [[X. Zhao](mailto:xinzzhao@tencent.com), [X. Li](mailto:xlxiangli@tencent.com), [S. Liu (Tencent)](mailto:shanl@tencent.com)]

[JVET-K0084](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3587) CE6: Matrix multiplication based NSST (Test 2.4.1 and Test 2.4.2) [X. Zhao, X. Li, S. Liu (Tencent)]

[JVET-K0085](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3588) CE6: Coupled primary and secondary transform (Test 3.1.1 and Test 3.1.2) [X. Zhao, X. Li, S. Liu (Tencent)]

[JVET-K0096](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3599) CE 6.1.11: AMT replacement and restriction [M. Koo, M. Salehifar, J. Lim, S. Kim (LGE)]

[JVET-K0098](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3601) CE 6.2.5: Layered Givens Transform (LGT) [M. Koo, M. Salehifar, J. Lim, S. Kim (LGE)]

[JVET-K0099](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3602) CE 6.2.6: Reduced Secondary Transform (RST) [M. Salehifar, M. Koo, J. Lim, S. Kim (LGE)]

[JVET-K0121](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3627) CE6: Chroma AMT (CE6.1.7.1) [T. Tsukuba, M. Ikeda, T. Suzuki (Sony)]

[JVET-K0123](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3628) CE6: Transform Matrix Replacement (CE6.1.8.1) [T. Tsukuba, M. Ikeda, T. Suzuki (Sony)]

[JVET-K0125](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3630) CE6.1.13: Implicit transform design for intra and inter residual coding [Y. Lin, M. Mao, S. Song, J. Zheng (HiSilicon), C. Zhu (UESTC)]

[JVET-K0127](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3632) CE6: AMT and NSST complexity reduction (CE6-3.3) [K. Abe, T. Toma (Panasonic)]

[JVET-K0139](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3645) CE6: Spatially Varying Transform (Test 6.1.12.1) [Y. Zhao, H. Yang, J. Chen(Huawei)]

[JVET-K0167](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3675) CE6: DST-VII with residual flipping (Test 1.4) [S.-C. Lim, J. Kang, H. Lee, J. Lee, S. Cho, H. Y. Kim (ETRI), N.-U. Kim, Y.-L. Lee (Sejong Univ.), D.-Y. Kim, W. J. Jeong (Chips&Media)]

[JVET-K0171](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3679) CE6: AMT with reduced transform types (Test1.5) [K. Choi, M. Park, C. Kim (Samsung)]

[JVET-K0173](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3681) CE6: AMT with block size restriction (Test1.6) [K. Choi, M. Park, C. Kim (Samsung)]

[JVET-K0174](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3682) CE6: NSST with modified NSST sets and signalling (Test2.3) [K. Choi, M. Park, C. Kim (Samsung)]

[JVET-K0250](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3760) CE6.1.10 & CE6.3.4: AMT simplification and improvement [M.-S. Chiang, Z.-Y. Lin, T.-D. Chuang, C.-Y. Chen, C.-W. Hsu, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-K0264](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3774) CE6: Report of CE6.1.3 (Transform reduction in AMT), with further reduction via DST-4 inheritance [K. Naser, F. Le Léannec, E. François (Technicolor)] [late]

[JVET-K0271](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3783) CE6.2 Secondary Transformation: NSST Signalling (test 6.2.1.1) [F. Urban, F. Le Léannec, E. Francois (Technicolor)]

[JVET-K0272](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3892) CE6.1.2: Efficient Implementations of AMT with Transform Adjustment Stages [A. Said, Y-H. Chao, H. Egilmez, M. Karczewicz, V. Seregin (Qualcomm)]

[JVET-K0305](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3819) CE6: Set of Transforms (Tests 3.2.1 and 3.2.2) [M. Siekmann, C. Bartnik, H. Schwarz, D. Marpe, T. Wiegand (HHI)]

[JVET-K0374](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3893) CE6.2.2: Hierarchically Structured Matrix-based Transforms for NSST [A. Said, H. Egilmez, Y-H. Chao, M. Karczewicz, V. Seregin (Qualcomm)]

[JVET-K0375](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3894) CE6.1.1: Extended AMT [A. Said, H. Egilmez, Y-H. Chao, M. Karczewicz, V. Seregin (Qualcomm)]

[JVET-K0397](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3917) CE6: Shrink Transform (CE6.1.9) [Kei Kawamura, Yoshitaka Kidani, Sei Naito (KDDI)] [late]

[JVET-K0356](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3874) Cross check of proposals for CE6 [Y. Reznik (Brightcove)] [late]

## CE7: Quantization and coefficient coding (12)

[JVET-K0027](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3572) CE7 summary report on quantization and coefficient coding [H. Schwarz, M. Coban]

The CE report summarizes the test results and cross check reports for the CE7 on quantization and coefficient coding. The following subtests are included in the CE:

**Subtest 7.1: Comparison of 4 proposals for entropy coding of transform coefficient levels**

This subtest was discussed 0900 Friday 13 July (chaired by GJS).

Overall gain relative to the VTM for these proposals was about 1.5/0.9/0.6 for AI/RA/LB. All proposals seemed roughly in that ballpark. All of them used about the same number of contexts (123–156).

The Samsung proposal 7.1.3 had a difference of initialization method, using 0.5 probability initialization. Variation (1) in the summary table tries to compensate for this by changing the reference to also use that initialization. The difference between the two was 0.0/0.2/0.6 for AI/RA/LB in the VTM.

There was also a difference in the VTM testing in the coding of remaining coefficients, where the Samsung proposal did not use additional features found in the BMS but others did. The proponent suggested focusing on the BMS results to avoid that difference of ~0.25%.

A proponent remarked that the method of training the context initialization was not known in general in this test for most technologies tested.

Test 7.1.2 is a scheme compatible with the K0071 trellis quantization method, but with that aspect disabled. A participant remarked that this scheme might have throughput issues due to the way that scheme could have a high number of context coded bins in the worst case.

**Subtest 7.2: Comparison of dependent quantization and sign data hiding**

This subtest was discussed 1920-2000 Thursday 12 July (chaired by GJS).

K0072 This has two sets of quantization reconstruction levels and a state machine to choose between them. The parity of the coefficient is used in the state machine. From the encoder perspective, it is suggested to be basically trellis coded quantization.

This uses double the number of context models for the significance flag and the absolute level greater than 1 flag.

The gain over the VTM is 5.0%/3.4%/2.7% for AI/RA/LD. The gain over the BMS is 2.5%/1.9%/1.6%. The encoder impact is about 10-13%.

This effectively has a combination of quantization and entropy coding together.

It was commented that were several relevant non-CE contributions which should be taken into account.

It was commented that also we need a fall-back mode that does not require encoder trellis search.

The decoding process is a bit more complicated.

This should be considered for testing with non-CE contributions in a CE. It was commented that at least one of the non-CE approaches is better and may be considered instead.

**Subtest 7.3: Investigation of 3 approaches related the derivation and signalling of quantization step sizes**

This subtest was discussed 0945 Friday 13 July (chaired by GJS).

There were three proposals in this area.

* K0140 (with two variants) has a scaling based on neighbouring samples. This is perceptually motivated, and detrimental to PSNR, but shows some gain in MSSIM.
  + Variant 1 uses reconstructed neighbours to control inverse quantization. It appeared that this has an unacceptable impact on decoding complexity.
  + Variant 2 scales the spatial-domain residual signal using reconstructed samples.
  + It was commented that encoder-side tricks can alternatively be used – e.g., using MS-SSIM for mode decisions or using adaptive delta-QP control.
  + This “bakes in” a specific criterion for the QP control.
  + Encoder tricks can also be combined with this. The proponent said that a combination of this scheme and encoder perceptual R-D control could provide roughly a MOS difference of 0.3.
  + There might be a different desired behaviour for PQ vs. HLG vs. SDR, 4:2:0 vs. 4:4:4, YCbCr vs. RGB, and other application-specific circumstances.
  + Very similar concepts have arisen in the HDR studies.
  + Hypothetically, we could end up with several different selectable automatic QP adjustment schemes, if we want to build in such automatic schemes.
  + This does QP control at a finer granularity than what an encoder would typically use, which may or may not be desirable.
  + It was commented that restricting such a scheme to only larger block regions may be desirable.
  + This adds some processing (a scaling of the residual signal) and has a serialization dependency (using neighbour samples to control current block processing).
  + As tested, the deblocking is not accounting for the scaling change.
  + Further study was encouraged, with consideration of the above-identified issues.
* K0251 proposed increasing the upper bound on QP by 6 (no effect on CTC)
  + The upper limit actually was encountered in CfP high QP use for RA conditions
  + Several participants commented that this had been encountered in a product, where quantization matrices need to be used to get around the issue.
  + For deblocking, an extrapolation of the current behaviour is used (specifics are in software)
  + The chroma QP derivation change is straightforward
  + Decision: Adopt – extending the range by 12 (pending spec text)
  + It was commented that in the future we should also think about the granularity of step sizes.
* K0252 proposed a different way of deriving chroma QP from luma QP
  + This proposes a change of the mapping function for deriving chroma QP by establishing a maximum difference between the input and output QPs of the matching function.
  + The cross-checker commented that there were other ways to adjust chroma QP
  + It ws noted that RExt has an ability to change the chroma QP on a block basis
  + A participant remarked that the chroma QP derivation function can cause strange effects in rate control since the R-D behaviour of the chroma is different from that for the luma
  + It was commented that QP also affects deblocking and that relationship should be studied.
  + Further study was encouraged, with a desire to establish a less ad-hoc manner of dealing with chroma QP.

**Subtest 7.4: Investigation of an approach for modifying the scanning order of transform coefficients**

This subtest was discussed 1042 Friday 13 July (chaired by GJS).

This proposal modifies the scanning order of transform coefficients based on block shapes. The reported gain was 0.3%/0.1%/0.1% for AI/RA/LB relative to the VTM and basically no gain over BMS.

HEVC has a dependency on the intra prediction mode for the scan order. That was removed when designing the first draft of VVC.

It was commented that this has some interaction with NSST and other aspects of coefficient coding.

This should be further studied together with secondary transforms and other aspects of coefficient coding

**Subtest 7.5: Comparison of two configurations for transform domain sign prediction**

This subtest was discussed 1050 Friday 13 July (chaired by GJS).

Test 7.5.1 (K0044) performs residual sign prediction in the transform domain, predicting up to 5 signs per transform block. The inverse transform requires the reconstructed samples of the neighbour blocks. This has a serious complexity impact (as per above with K0140). This also has a very significant impact on decoder runtime. The gain over the VTM is reported as 1.3%/1.0%/0.7% for AI/RA/LB.

Test 7.5.2 is the same, but only used for intra.

Further study was recommended. A way to avoid the serial dependency is especially needed.

[JVET-K0044](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3545) CE7: Residual sign prediction in transform domain (Tests 7.5.1 and 7.5.2) [A. Filippov, A. Karabutov, V. Rufitskiy, J. Chen (Huawei)]

[JVET-K0069](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3570) CE7: Coefficient Coding (Test 1.1) [M. Coban, J. Dong, T. Hsieh, M. Karczewicz (Qualcomm)]

[JVET-K0071](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3573) CE7: Transform coefficient coding and dependent quantization (Tests 7.1.2, 7.2.1) [H. Schwarz, T. Nguyen, D. Marpe, T. Wiegand (Fraunhofer HHI)]

[JVET-K0138](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3644) CE7.1.3: Scan Region-based Coefficient Coding [Y. Piao, W. Choi, C. Kim (Samsung)]

[JVET-K0140](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3646) CE7: Adaptive quantization step size scaling (Test 7.3.1) [Y. Zhao, H. Yang, J. Chen (Huawei)]

[JVET-K0251](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3761) CE7.3.2: Extension of quantization parameter value range [S.-T. Hsiang, S.-M. Lei (MediaTek)]

[JVET-K0252](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3762) CE7.3.3: Derivation of chroma QP from luma QP [S.-T. Hsiang, S.-M. Lei (MediaTek)]

[JVET-K0321](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3835) CE 7.1.4: JEM 7.0 coefficient coding with complexity reduction [C. Auyeung, J. Chen (Huawei)]

[JVET-K0398](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3918) CE7: Block size dependent coefficient scanning (CE7.4.1) [Y. Kidani, K. Kawamura, S. Naito (KDDI)] [late]

[JVET-K0457](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3983) Crosscheck for CE7-1.2 [M. Gao, W. Zhang (Hulu)] [late]

[JVET-K0459](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3985) Crosscheck for CE7-5.1 [M. Gao, W. Zhang (Hulu)] [late]

## CE8: Current picture referencing (6)

Contributions in this category were discussed Friday 13 July 1200–1300 (chaired by GJS).

[JVET-K0028](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3579) CE8 summary report on current picture referencing [X. Xu, K. Müller, L. Wang]

This contribution provides a summary report of Core Experiment 8 on current picture referencing. Four tests have been agreed to carry out in CE8 in between JVET-J and JVET-K meetings, to study and evaluate technologies related to current picture referencing. In this report, coding performance and complexity of these tests are reported and analyzed. In particular, test results against VTM anchor are provided to show the coding efficiency and complexity trade-off of each proposed approach. Test results against BMS anchor are also provided to show the interaction with BMS coding tools. Crosschecking results for the performed tests are integrated in this contribution.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Test** | **Tester** | **Document** | **Tool description** | **Cross checker** |
| 8.1 | [G. Venugopal](mailto:gayathri.venugopal@hhi.fraunhofer.de)  (HHI) | JVET-J0039 | Intra region-based template matching | F. Racape  (Technicolor) |
| 8.2.1 | X. Xu  Tencent | JVET-J0050 | Current picture referencing, for intra and inter pictures, using CPR flag | [G. Venugopal](mailto:gayathri.venugopal@hhi.fraunhofer.de)  (HHI)  Xiaozhen Zheng  (DJI) |
| 8.2.2 | X. Xu  (Tencent) | JVET-J0050 | Current picture referencing for intra and inter pictures, using refIdx approach | W. Zhang  (Hulu) |
| 8.3 | X. Zuo  (Hikvision) | JVET-J0042 | Current picture referencing for intra pictures | X. Ma  (Huawei) |

Note that 8.3 is only using CPR for I slices. The others are using it for both.

Note that 8.1 is only using CPR for luma; an extension to support chroma is in a non-CE contribution.

It was remarked that the HEVC scheme supports CPR in biprediction. These do not.

These schemes all provided substantial gain on CTC as well as SCC content.

The schemes provided about as much gain in the BMS context as in VTM.

It was suggested that the 8.2.2 approach is more mature, as it is using the same method as HEVC.

There is a substantial complexity impact. It was suggested that a baseline profile would need some constraints on the design.

Decision: Adopt 8.2.2 approach (JVET-K0076) into BMS. Regarding whether to include in CTC or not, it will be included; per section 12.2. Further study is needed to determine appropriate constraints and profiling implications. The current version seems too complex for a “baseline profile”, but some variation of this seems needed in the standard, and with some constraints it could become appropriate for a “baseline profile”.

[JVET-K0048](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3549) CE8: Intra Region-based Template Matching (Test 8.1) [G. Venugopal, K. Müller, H. Schwarz, D. Marpe, T. Wiegand (HHI)]

[JVET-K0075](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3577) CE8-2.1: Current picture referencing using block level flag signalling [X. Xu, X. Li, S. Liu (Tencent)]

[JVET-K0076](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3578) CE8-2.2: Current picture referencing using reference index signalling [X. Xu, X. Li, G. Li, S. Liu (Tencent)]

[JVET-K0436](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3959) Crosscheck for CE8-2.2 [W. Zhang (Hulu)] [late]

[JVET-K0450](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3975) CE8-3.1: Current picture referencing for intra pictures [L. Wang, F. Chen (Hikvision)] [late]

## CE9: Decoder side motion vector derivation (25)

Contributions in this category were discussed Friday 13 July 1100–1230 (chaired by JRO).

[JVET-K0029](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3721) CE9: Summary Report on Decoder Side MV Derivation [S. Esenlik, Y.-W. Chen]

The tools in the scope of this CE include bi-directional optical flow, template matching and bilateral matching based techniques for motion vector derivation and refinement at the decoder side.

The core experiment is organized into 5 sub-tests as follows:

* CE9.1 - Decoder Side Motion Vector Refinement (DMVR): 5 tests are performed in this subcategory.
* CE9.2 - Bilateral Matching: 8 tests.
* CE9.3 - Template Matching: 7 tests.
* CE9.4 - MV Candidate List Reordering by Template Matching: 3 tests.
* CE9.5 - BIO: 3 tests.

This report summarises the status of each experiment. Crosscheck results are integrated in the document.

CE9.1: Decoder Side Motion Vector Refinement (DMVR)

|  |  |  |
| --- | --- | --- |
| # | Test | Input Documents/Tester |
| CE9.1.1 | * Search Range is 1 * Adaptive search pattern (6 points instead of 9) * Mean removed SAD as cost function * Early termination: if motion vector is not changed after an iteration | [JVET-K0199](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3708)  [X. Chen](mailto:anderson.chen@hisilicon.com) (Hisilicon, Huawei) |
| CE9.1.2 | * Early termination after L0 search if motion vector is not changed after one iteration | [JVET-K0253](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3763)  Yu-Chi Su  (MediaTek) |
| CE9.1.3 | * Early termination based on initial SAD cost between prediction L0 and prediction L1 * High precision SAD (no clip and round) | JVET-K0342  Xiaoyu Xiu  (InterDigital) |
| CE9.1.5 | * DMVR not applied if MV difference between the selected candidate and any of the previous candidates in the merge list is less than a pre-defined threshold in both horizontal and vertical directions, where the thresholds are ¼-pel, ½-pel and 1-pel for blocks with less than 64, less than 256 and more than 256 pixels, respectively. | JVET-K0358  Chun-Chi Chen  (Qualcomm) |
| CE9.1.6 | * MV difference mirroring. * Results are to be provided for number of iterations 4, 2, 1, and half-pel on/off. * 6 point corner selective integer search and 4 point half pel search. * Results are to be provided by switching off spatial MV prediction from refined motion vectors in 32x32 grid. | JVET-K0216  Semih Esenlik (Huawei , USTC) |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Tester** | **VTM** | | | | | **BMS** | | | | |
|  | | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| AHG 13 | DMVR in BMS according to AHG13  (Test is DMVR off) | -2.65% | -2.54% | -2.67% | 109% | 131% | 1.47% | 1.55% | 1.64% | 99% | 92% |
| 9.1.1 | Xu Chen (Hisilicon, Huawei,) | -1.46% | -1.44% | -1.47% | 106% | 116% | 0.59% | 0.58% | 0.65% | 99% | 95% |
| -2.80% | -2.58% | 2.69% | 107% | 117% | -0.16% | -0.07% | -0.11% | 99% | 95% |
| 9.1.2 | Yu-Chi Su (MediaTek), only RA | -2.65% | -2.52% | -2.65% | 108% | 127% | -0.01% | -0.01% | -0.02% | 101% | 99% |
| 9.1.3 | Xiaoyu Xiu (InterDigital) | -2.60% | -2.52% | -2.64% | 109% | 119% | 0.03% | 0.00% | 0.02% | 100% | 96% |
| 9.1.5 | Chun-Chi Chen (Qualcomm) | -2.66% | -2.54% | -2.67% | 108% | 127% | -0.02% | 0.00 | -0.01% | 100% | 99% |
| 9.1.6 | Semih Esenlik (Huawei , USTC) | -2.91% | -2.62% | -2.74% | 103% | 114% | -0.23% | -0.10% | -0.13% | 98% | 95% |
| -3.50% | -2.92% | -3.07% | 106% | 120% | -0.58% | -0.31% | -0.34% | 99% | 96% |
| -3.62% | -3.30% | -3.43% | 104% | 118% | -0.68% | -0.53% | -0.58% | 98% | 96% |
| -4.39% | -3.72% | -3.90% | 107% | 127% | -1.17% | -0.86% | -0.93% | 99% | 98% |
| -3.94% | -3.58% | -3.75% | 105% | 121% | -0.90% | -0.78% | -0.84% | 99% | 97% |
| -4.71% | -4.04% | -4.22% | 109% | 131% | -1.40% | -1.07% | -1.17% | 99% | 99% |
| -2.48% | -2.18% | -2.27% | 105% | 117% | -0.04% | 0.11% | 0.12% | 99% | 97% |
| -2.96% | -2.47% | -2.55% | 106% | 123% | -0.34% | -0.09% | -0.06% | 100% | 98% |
| -3.20% | -2.86% | -2.93% | 105% | 121% | -0.49% | -0.32% | -0.31% | 99% | 98% |
| -3.87% | -3.25% | -3.39% | 108% | 130% | -0.92% | -0.59% | -0.64% | 100% | 100% |
| -4.26% | -3.60% | -3.75% | 110% | 134% | -1.16% | -0.83% | -0.89% | 101% | 101% |

The following table shows properties of the different methods

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Tester** | **Initial MV signalled** | **Sub-CU refinement** | **Neighboring recon. samples used** | **Max # of SAD calculation** | **Max. SR** | **Cost Function** | **Interpolation filter/tap no** | **Note** |
| AHG 13 | DMVR in BMS according to AHG13\*  \*(Anchor is BMS-DMVR) | yes | no | no | 18 | 1 | SAD | DCTIF/8 | SIMD = SSE42 anchor&test |
| 9.1.1 | Xu Chen (Hisilicon, Huawei,) | yes | no | no | 12 | 1 | MRSAD | DCTIF/8 | No SIMD for MRSAD calculation  SIMD = AVX2 anchor&test  Prediction from refined MV disabled |
| yes | no | no | 12 | 1 | MRSAD | DCTIF/8 | No SIMD for MRSAD calculation  SIMD = AVX2 anchor&test |
| 9.1.2 | Yu-Chi Su (MediaTek), only RA | yes | no | no | 18 | 1 | SAD | DCTIF/8 | SIMD = SSE42 anchor&test  MAX # of SAD for L0 is 9; Max # of SAD for L1 is 9 but its optional |
| 9.1.3 | Xiaoyu Xiu (InterDigital) | yes | no | no | 19 | 1 | SAD | DCTIF/8 | SIMD = AVX anchor&test |
| 9.1.5 | Chun-Chi Chen (Qualcomm) | yes | no | no | 18 | 1 | SAD | DCTIF/8 | SIMD = AVX anchor&test |
| 9.1.6 | Semih Esenlik (Huawei , USTC) | yes | no | no | 6 | 1 | MRSAD | DCTIF/8 | No SIMD for MRSAD calculation,  SIMD = AVX2 anchor&test |
| yes | no | no | 10 | 1 | MRSAD | DCTIF/8 |
| yes | no | no | 12 | 2 | MRSAD | DCTIF/8 |
| yes | no | no | 16 | 2 | MRSAD | DCTIF/8 |
| yes | no | no | 24 | 4 | MRSAD | DCTIF/8 |
| yes | no | no | 28 | 4 | MRSAD | DCTIF/8 |
| yes | no | no | 6 | 1 | MRSAD | DCTIF/8 | No SIMD for MRSAD calculation,  SIMD = AVX2 anchor&test,  No reference to refined MV inside 32x32 grid |
| yes | no | no | 10 | 1 | MRSAD | DCTIF/8 |
| yes | no | no | 12 | 2 | MRSAD | DCTIF/8 |
| yes | no | no | 16 | 2 | MRSAD | DCTIF/8 |
| yes | no | no | 28 | 4 | MRSAD | DCTIF/8 |

Important complexity aspects are number of SADs, memory access (search range) in general, and latency (due to dependency between spatial neighbors, pipelining is complicated). The latter aspect is addressed in 9.1.1.a, however it loses 1.2% in VTM; and 0.6% in BMS. For the other aspects, it can be seen that increasing SAD number or SR improves quality.

It is agreed that DMVR is not mature enough to be moved into VTM.

It is agreed that the next version of BMS should include a DMVR that resolves the latency problem.

It is agreed that upcoming CEs should not include any approach that has a latency problem.

The only proposal from CE9.1 that resolves the latency problem is 9.1.1a.

It was initially agreed to adopt JVET-K0199 (as per CE9.1.1.a), i.e. do not use refined motion vectors for anything but the MC of the current block. This is asserted to be the simplest solution for the latency problem, no additional storage requirements, no additional rules, etc. This decision was later revisited due to the adoption of 9.2.9.l. However, in the context of 9.2.9.l, still the aspect of using the non-refined MV in deblocking (as initially suggested in K0199) was retained.

Question: Do we know what is the impact on worst case memory bandwidth for SR1/2/4? Compared to SR0 = DMVR off ? SR1: 140%; SR2: 186%; SR4: 298%; SR8: 600%.

Note: SR up to 2 with bilinear interpolation is claimed to be still 100%

Note these Numbers are preliminary, need more check – to be done in upcoming CE (there are some further notes under CE9 related section).

CE9.2: Bilateral Matching

|  |  |  |
| --- | --- | --- |
| # | Test | Tester |
| CE9.2.1 | * Explicitly signalled initial MV candidate. * Sub-CU search on * For the sub-CU-level search, only the MV determined from the CU-level search is evaluated. * Bounding window for Sub-CU search * Disabled for 4x4, 4x8, and 8x4 CUs. | JVET-K0254  Tzu-Der Chuang  (MediaTek) |
| CE9.2.2 | * Clustering of initial candidates * The initial MV is rounded to the integer precision * Max number of iterations 12 * Half pel refinement: ½, 1/4, 1/8 pel in order * Early termination based on SAD cost | JVET-K0343  Xiaoyu Xiu  (InterDigital) |
| CE9.2.3 | * SubCU level process is removed. * Candidate list size reduced * Predefined memory access windows relative to the current CTU (dependent on number of reference frames). * Adaptive search pattern to simplify search | JVET-K0177  Jingya Li  (Panasonic) |
| CE9.2.5 | * Merge index is signalled * Adaptively apply Bilateral Matching when the following conditions are met   + Uni-directional, ATMVP, STMVP, affine and candidates using IC mode are excluded   + (POCref0 – POCcur)\*(POCref1 – POCcur) values is negative   + The MV difference between the selected candidate and any of the previous candidates is not less than a pre-defined threshold (i.e. ¼-pel, ½-pel and 1-pel for blocks with less than 64, less than 256 and other larger blocks, respectively) * Sub-block refinement is removed * Mean removed SAD is applied adaptively based on CU size (i.e. MRSAD for blocks with more than 64 pixels) | JVET-K0359  Chun-Chi Chen  (Qualcomm) |
| CE9.2.6 | Based on CE9.2.5 two modifications are tested:   * The DCTIF for search is replaced by bi-linear filter; * The Search range is reduced from 8 to 2 | JVET-K0359  Chun-Chi Chen  (Qualcomm) |
| CE9.2.7 | * Implementation based on Bilateral Matching code in BMS1.0 software. * Motion vector difference is mirrored for forward and backward MVPs (bilateral matching disabled otherwise). * Sub-CU refinement off. * Merge candidates are used as origin MVs. | JVET-K0303  Byeongdoo Choi (Sharp) |
| CE9.2.8 | Implemented on top of CE9.2.9.   * 4 points half-pel search is replaced by 2 point adaptive half-pel search pattern. | JVET-K0378  Yue Li (USTC) |
| CE9.2.9 | * Bilateral matching cost function instead of generating template. * MVD mirroring with initial MV candidate signalled as merge index. * Results are to be provided for search ranges 4, 2, 1, and half-pel off. * Results are to be provided by disabling spatial prediction from refined motion vectors within 32x32 grid. * Results are to be provided for disabling spatial prediction from refined MV completely. | JVET-K0217  Semih Esenlik (Huawei , USTC) |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Tester** | **VTM** | | | | | **BMS** | | | | |
|  | | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 9.2.1 | Tzu-Der Chuang (MediaTek), RA/LB | -5.72% | -4.64% | -4.82% | 172% | 272% | -2.35% | -1.90% | -2.00% | 128% | 168% |
| -0.43% | -0.71% | -0.86% | 236% | 151% | -0.34% | -0.64% | -0.64% | 149% | 132% |
| (supplementary), no sub-block search  RA/LB | -4.76% | -3.92% | -4.06% | 155% | 168% | -1.58% | -1.27% | -1.33% | 118% | 163% |
| -0.46% | -0.51% | -0.66% | 172% | 131% | -0.41% | -0.47% | -0.59% | 125% | 118% |
| 9.2.2 | Xiaoyu Xiu (InterDigital)  RA/LB | -2.45% | -2.25% | -2.22% | 125% | 162% | -1.02% | -1.10% | -1.08% | 107% | 112% |
| -1.26% | -1.36% | -1.48% | 138% | 148% | -0.57% | -0.73% | -0.72% | 113% | 117% |
| 9.2.3 | Jingya Li (Panasonic)  RA/LB | -1.90% | -1.76% | -1.81% | 116% | 154% | -0.62% | -0.71% | -0.78% | 106% | 116% |
| -0.64% | -0.86% | -1.10% | 119% | 135% | -0.64% | -0.66% | -0.73% | 108% | 121% |
| 9.2.5 | Chun-Chi Chen (Qualcomm) | -5.28% | -4.70% | -4.88% | 116% | 164% | -1.76% | -1.55% | -1.62% | 102% | 112% |
| 9.2.6 | Chun-Chi Chen (Qualcomm) | -4.76% | -4.15% | -4.32% | 112% | 150% | -1.40% | -1.13% | -1.22% | 100% | 108% |
| 9.2.7 | Byeongdoo Choi (Sharp)  RA/LB | -6.68% | -6.11% | -6.37% | 157% | 277% | -3.32% | -3.10% | -3.26% | 125% | 162% |
| -2.98% | -3.18% | -3.37% | 218% | 338% | -2.41% | -2.67% | -2.56% | 150% | 232% |
| 9.2.8 | Yue Li (USTC) | -3.61% | -3.15% | -3.27% | 107% | 119% | -0.64% | -0.46% | -0.54% | 99% | 97% |
| -4.48% | -3.92% | -4.11% | 109% | 125% | -1.22% | -0.98% | -1.05% | 100% | 98% |
| 9.2.9 | Semih Esenlik (Huawei , USTC) | -3.18% | -2.95% | -3.09% | 103% | 115% | -0.38% | -0.33% | -0.35% | 97% | 96% |
| -3.68% | -3.17% | -3.33% | 106% | 124% | -0.68% | -0.48% | -0.52% | 98% | 98% |
| -3.80% | -3.57% | -3.70% | 105% | 119% | -0.79% | -0.73% | -0.78% | 99% | 97% |
| -4.59% | -3.97% | -4.14% | 107% | 132% | -1.29% | -1.02% | -1.10% | 98% | 101% |
| -4.07% | -3.86% | -4.03% | 105% | 121% | -0.97% | -0.96% | -0.98% | 99% | 97% |
| -4.92% | -4.31% | -4.50% | 109% | 136% | -1.50% | -1.28% | -1.35% | 99% | 102% |
| -2.72% | -2.48% | -2.56% | 103% | 118% | -0.18% | -0.06% | -0.05% | 98% | 97% |
| -3.11% | -2.67% | -2.76% | 105% | 125% | -0.41% | -0.22% | -0.21% | 99% | 100% |
| -3.38% | -3.13% | -3.23% | 106% | 121% | -0.59% | -0.49% | -0.50% | 99% | 98% |
| -4.08% | -3.47% | -3.62% | 109% | 133% | -1.03% | -0.78% | -0.78% | 100% | 101% |
| -4.47% | -3.88% | -4.02% | 111% | 138% | -1.28% | -1.04% | -1.06% | 100% | 103% |
| -2.80% | -2.57% | -2.63% | 109% | 132% | -0.30% | -0.23% | -0.20% | 100% | 102% |

The following table shows properties of the different methods

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Tester** | **Initial MV signalled** | **Sub-CU refinement** | **Neighboring recon. samples used** | **Max # of SAD calculation** | **Max. SR** | **Cost Function** | **Interpolation filter/tap no** | **Note** |
| 9.2.1 | Tzu-Der Chuang (MediaTek), RA/LB | yes | Yes  (optional) | no | CU-level:  9+32\*5+8  = 177  Sub-CU-level: 5 + 16\*3 + 4 = 57 | 8 | SAD | ME: Bilinear filter/2  MC: DCTIF/8  (same as JEM) | SIMD = SSE42 anchor&test |
| (supplementary), no sub-block search | yes | no | no | 9+32\*5+8  = 177 | 8 | SAD | ME: Bilinear filter/2  MC: DCTIF/8  (same as JEM) | SIMD = SSE42 anchor&test |
| 9.2.2 | Xiaoyu Xiu (InterDigital) | no | Yes  (optional) | no | 16 for CU-level | 12 | SAD | ME: Bilinear filter/2  MC: DCTIF/8 | SIMD = AVX anchor&test |
| 9.2.3 | Jingya Li (Panasonic) | no | no | no | Not  defined | Within pre-determined memory block | SAD | ME: Bilinear filter/2  MC: DCTIF/8  (same as JEM) | SIMD = SSE42 anchor&test |
| 9.2.5 | Chun-Chi Chen (Qualcomm) | yes | no | no | A loose upper bound:  9+25\*5+8  = 142 | 8 | >64 pixels:  MRSAD  ≤64 pixels:  SAD | DCTIF/8 | SIMD for MRSAD calculation,  SIMD = AVX anchor&test |
| 9.2.6 | Chun-Chi Chen (Qualcomm) | yes | no | no | A loose upper bound:  9+5+11\*2+8  = 44 | 2 | >64 pixels:  MRSAD  ≤64 pixels:  SAD | ME: Bilinear filter/2  MC: DCTIF/8 | SIMD for MRSAD calculation,  SIMD = AVX anchor&test |
| 9.2.7 | Byeongdoo Choi (Sharp) | no | Yes  (optional) | yes | same as JEM | 8 | SAD | ME: Bilinear filter/2  MC: DCTIF/8  (same as JEM) | SIMD = AVX2 anchor&test |
| 9.2.8 | Yue Li (USTC) | yes | no | no | 10 | 1 | MRSAD | DCTIF/8 | No SIMD for MRSAD calculation,  SIMD = AVX2 anchor&test |
| yes | no | no | 13 | 2 | MRSAD | DCTIF/8 |
| 9.2.9 | Semih Esenlik (Huawei , USTC) | yes | no | no | 6 | 1 | MRSAD | DCTIF/8 | No SIMD for MRSAD calculation,  SIMD = AVX2 anchor&test |
| yes | no | no | 10 | 1 | MRSAD | DCTIF/8 |
| yes | no | no | 9 | 2 | MRSAD | DCTIF/8 |
| yes | no | no | 13 | 2 | MRSAD | DCTIF/8 |
| yes | no | no | 15 | 4 | MRSAD | DCTIF/8 |
| yes | no | no | 19 | 4 | MRSAD | DCTIF/8 |
| yes | no | no | 6 | 1 | MRSAD | DCTIF/8 | No SIMD for MRSAD calculation,  SIMD = AVX2 anchor&test,  No reference to refined MV inside 32x32 grid |
| yes | no | no | 10 | 1 | MRSAD | DCTIF/8 |
| yes | no | no | 9 | 2 | MRSAD | DCTIF/8 |
| yes | no | no | 13 | 2 | MRSAD | DCTIF/8 |
| yes | no | no | 19 | 4 | MRSAD | DCTIF/8 |
| yes | no | no | 13 | 2 | MRSAD | DCTIF/8 | No SIMD for MRSAD calculation,  SIMD = AVX2 anchor&test,  No reference to refined MV inside the whole frame |

The only proposal from CE9.2 that resolves the latency problem is 9.2.9l. This also gives 0.3% bit rate reduction, has less SAD computations, but is slightly worse in terms memory accesses (SR2 vs SR1 of current BMS-DMVR).

Could be an interesting candidate of next BMS, depending on report of complexity/memory.

Further discussion track B Monday afternoon: Though the memory bandwidth is not optimized yet, this is currently the best available solution and should replace the previous bilateral matching in BMS.

Decision (BMS): Adopt JVET-K0217 (variant 9.2.9l) and and aspect from JVET-K0199. Modification: Non refined MV to be used for deblocking, i.e. use the method of 9.1.1.a here. (Note: Some implementations might do the deblocking right after reconstruction, such that using a refined MV would again cause a latency problem. Refined MV to be used for TMVP as in original 9.2.9.l)

Note: This will be used as reference for comparison in the upcoming CE, whereas it is known that further reduction of memory bandwidth and complexity is needed, and there are other proposals in the CE which might be better in that regard.

CE9.3: Template Matching

|  |  |  |
| --- | --- | --- |
| # | Test | Tester |
| CE9.3.1 | * Additional merge list is constructed which is different from regular merge list (Max 7 candidates). * Index signalled. * Sub-PU search off * Search range is restricted to 8 samples | JVET-K0168  Hahyun Lee  (ETRI) |
| CE9.3.2 | * Sub-CU refinement process is removed. * Candidate list size reduced (unilateral candidates not inserted) * Adaptive pattern search   + the search only one point in the direction of smallest cost,   + if the cost does not get lower in the selected direction, other directions are searched. * Predefined memory access windows relative to the current CTU (dependent on number of reference frames).   + 374x374 samples if 1 refPic   + 264x264 samples if 2 refPic   + … | JVET-K0178  Jingya Li  (Panasonic) |
| CE9.3.4 | * Applied on top of JVET-J0021 * Unilateral candidates are not inserted in merge list. | JVET-K0214  Antoine Robert (Technicolor) |
| CE9.3.5 | * Applied on top of JVET-J0021 * MVP list is derived from candidates in merge list * Max 2 candidates are checked by template matching | JVET-K0214  Antoine Robert (Technicolor) |
| CE9.3.6 | * Merge: MV refinement applied only if first candidate is selected. * AMVP: refinement applied to both candidates * AM mode: applied only on the merge direction * Sub-CU search disabled | JVET-K0200  Xu Chen  (HiSilicon) |
| CE9.3.7 | * Applied only to the first candidate in the merge list. * First candidate MV in merge list is refined using template matching. * Not applied to ATMVP candidate. | JVET-K0088  Naeri Park  (LGE) |
| CE9.3.8 | * Combined results of CE9-4.1 and CE9-3.7 | JVET-K0088  Naeri Park  (LGE) |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Tester** | **VTM** | | | | | **BMS** | | | | |
|  | | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| AHG 13 | DMVR in BMS according to AHG13  (Test is DMVR off) | -2.65% | -2.54% | -2.67% | 109% | 131% | 1.47% | 1.55% | 1.64% | 99% | 92% |
| 9.3.1 | Hahyun Lee (ETRI)  RA/LB | -4.25% | -4.20% | -4.44% | 257% | 172% | -2.22% | -1.97% | -2.20% | 148% | 113% |
| -1.23% | -1.31% | -1.24% | 275% | 141% | -1.10% | -1.00% | -0.92% | 163% | 117% |
| 9.3.2 | Jingya Li (Panasonic)  RA/LB | -4.80% | -4.57% | -4.71% | 126% | 180% | -2.62% | -2.37% | -2.58% | 115% | 127% |
| -2.12% | -1.78% | -1.68% | 152% | 159% | -1.78% | -1.45% | -1.22% | 132% | 135% |
| 9.3.4 | Antoine Robert (Technicolor)  RA/LB | -6.57% | -6.22% | -6.44% | 171% | 295% | -3.65% | -3.52% | -3.68% | 133% | 173% |
| -3.39% | -3.62% | -3.94% | 219% | 227% | -2.76% | -3.02% | -2.98% | 166% | 186% |
| 9.3.5 | Antoine Robert (Technicolor) | Combined results (CE9.3.4 and CE9.3.5) are reported by the proponent (see 9.3.4) | | | | | | | | | |
| 9.3.6 | Xu Chen (HiSilicon)  RA/LB | -1.70% | -1.86% | -1.90% | 156% | 122% | -0.71% | -0.77% | -0.81% | 115% | 108% |
| -0.94% | -1.13% | -1.26% | 182% | 126% | -0.66% | -0.92% | -0.73% | 126% | 120% |
| -3.12% | -2.92% | -2.98% | 159% | 119% | -1.60% | -1.42% | -1.50% | 116% | 109% |
| -1.60% | -1.54% | -1.55% | 183% | 120% | -1.34% | -1.07% | -0.95% | 126% | 110% |
| 9.3.7 | Naeri Park (LGE)  RA/LB | -3.42% | -3.28% | -3.46% | 115% | 167% | -1.51% | -1.42% | -1.52% | 103% | 124% |
| -1.42% | -1.26% | -1.27% | 112% | 160% | -1.24% | -0.82% | -0.88% | 103% | 130% |
| additional test (Search range 2)  RA/LB | -2.91% | -2.60% | -2.77% | 112% | 155% | -1.29% | -1.12% | -1.20% | 103% | 121% |
| -1.21% | -0.86% | -0.97% | 111% | 153% | -1.05% | -0.45% | -0.60% | 103% | 127% |
| additional test (Search range 1)  RA/LB | -2.36% | -2.02% | -2.12% | 110% | 145% | -1.13% | -0.89% | -0.97% | 103% | 118% |
| -0.96% | -0.61% | -0.76% | 109% | 146% | -0.92% | -0.53% | -0.34% | 103% | 125% |
| 9.3.8 | Naeri Park (LGE)  RA/LB | -4.01% | -3.90% | -4.08% | 117% | 193% | -2.10% | -2.08% | -2.17% | 103% | 141% |
| -1.83% | -1.50% | -1.61% | 114% | 196% | -1.52% | -1.08% | -1.00% | 103% | 152% |

Properties of proposals:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Tester** | **Initial MV signalled** | **Sub-CU refinement** | **Neighboring recon. samples used** | **Max # of SAD calculation** | **Max. SR** | **Cost Function** | **Interpolation filter/tap no** | **Note** |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 9.3.1 | Hahyun Lee (ETRI) | Yes | No | Yes | Not  defined | 8 | SAD | Bilinear filter/2 | SIMD = AVX anchor&test (Encoder)  = SSE42 anchor&test (Decoder) |
| 9.3.2 | Jingya Li (Panasonic) | no | no | yes | Not  defined | Within pre-determined memory block | SAD | ME: Bilinear filter/2  MC: DCTIF/8  (same as JEM) | SIMD = SSE42 anchor&test |
| 9.3.4 | Antoine Robert (Technicolor) | no | yes | yes | Not  defined | 8 | SAD | Bilinear filter/2 | SIMD = SSE42 anchor&test |
| 9.3.5 | Antoine Robert (Technicolor) | Combined results (CE9.3.4 and CE9.3.5) are reported by the proponent (see 9.3.4) | | | | | | | |
| 9.3.6 | Xu Chen (HiSilicon) | no | no | yes | 9 | 1 | SAD | DCTIF/8 | SIMD = AVX2 anchor&test |
| no | no | yes | Not  defined | 8 | SAD | Bilinear filter/2 | SIMD = AVX2 anchor&test |
| 9.3.7 | Naeri Park (LGE) | no | no | yes | Not  defined | 8 | SAD | Bilinear filter/2 | SIMD = AVX2 anchor&test |
|  | additional test (Search range 2) | no | no | yes | Not defined | 2 | SAD | Bilinear filter/2 | SIMD = AVX2 anchor&test |
|  | additional test (Search range 1) | no | no | yes | Not defined | 1 | SAD | Bilinear filter/2 | SIMD = AVX2 anchor&test |
| 9.3.8 | Naeri Park (LGE) | no | no | yes | Not  defined | 8 | SAD | Bilinear filter/2 | SIMD = AVX2 anchor&test |

These methods show interesting in both cases of VTM and BMS. DMVR is turned off (except for 9.3.2 and 9.3.6) when TM is run. The BMS results are against a normal anchor, i.e. gains over DMVR are shown in the other cases.

However, due to the fact that TM requires reconstructed samples from the neighbour blocks, the latency issue is even more severe than in BMS-DMVR. Therefore, it is unlikely that any of these approaches would be acceptable for standardization.

No further action on any of these proposals.

CE9.4: MV Candidate List Reordering by Template Matching

|  |  |  |
| --- | --- | --- |
| # | Test | Tester |
| CE9.4.1 | * Merge list is constructed according to VTM and BMS rules. * The candidates are reordered according to template matching cost. | JVET-K0088  Naeri Park  (LGE) |
| CE9.4.2 | * Applied to AMVP mode * MVD sign is signalled as an index, where table entries are ranked according to template matching cost | JVET-K0067  Sergey Ikonin  (Huawei) |
| CE9.4.3 | * Merge list is constructed according to VTM and BMS rules. * The candidates are reordered according to template matching cost. * Sub-CU candidates are also considered in sorting. * Sorting is performed within groups (not full reorder). | JVET-K0143  Na Zhang (HiSilicon) |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Tester** | **VTM** | | | | | **BMS** | | | | |
|  | | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| 9.4.1 | Naeri Park (LGE)  RA/LB | -0.71% | -0.74% | -0.74% | 104% | 122% | -0.62% | -0.73% | -0.76% | 101% | 116% |
| -0.66% | -0.71% | -0.55% | 103% | 130% | -0.65% | -0.67% | -0.40% | 101% | 122% |
| 9.4.2 | Sergey Ikonin (Huawei)  RA/LB | -0.49% | -0.44% | -0.38% | 102% | 104% | -0.29% | -0.28% | -0.28% | 100% | 101% |
| -0.28% | -0.17% | -0.32% | 102% | 104% | -0.33% | -0.41% | -0.37% | 100% | 101% |
| 9.4.3 | Na Zhang (HiSilicon)  RA/LB | -0.56% | -0.51% | -0.48% | 103% | 115% | -0.71% | -0.81% | -0.80% | 101% | 113% |
| -0.81% | -0.46% | -0.63% | 104% | 119% | -0.90% | -0.67% | -0.83% | 101% | 113% |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Tester** | **Initial MV signalled** | **Sub-CU refinement** | **Neighboring recon. samples used** | **Max # of SAD calculation** | **Max. SR** | **Cost Function** | **Interpolation filter/tap no** | **Note** |
| 9.4.1 | Naeri Park (LGE) | no | no | yes | 5 | N.A | SAD | Bilinear filter/2 | SIMD = AVX2 anchor&test |
| 9.4.2 | Sergey Ikonin (Huawei) | yes | no | yes | 4 | N.A. | SAD | Bilinear filter/2 | SIMD = AVX2 anchor&test |
| 9.4.3 | Na Zhang (HiSilicon) | no | no | yes | 2 | N.A | SAD | DCTIF/8 | SIMD = AVX2 anchor&test |
| no | no | yes | 2 | N.A | SAD | DCTIF/8 | SIMD = AVX2 anchor&test |

These approaches construct a merge list based on template matching. The number of SAD computations depends on the number of candidates that are checked. Though no search is performed around the candidate positions, worst case memory access increases linearly with the number of candidates that need to be reordered, if they would pointing to non-overlapping areas in the reference picture.

Otherwise, the latency problem mentioned for CE9.3 is also existing here.

No further action on any of these proposals.

CE9.5: BIO

|  |  |  |
| --- | --- | --- |
| # | Test | Tester |
| CE9.5.2 | * Picture resolution dependent block size 2x2 and 4x4 based on the video resolution. * gradients are directly calculated based on interpolated prediction signal using a 3-tap filter ([1 0 -1]) * BIO applied to chroma. Luma displacement vectors are reused after scaling. * Reference block size is (w+7)x(h+7), padding is applied. * Division is replaced by shift operations | JVET-K0255  Tzu-Der Chuang  (MediaTek) |
| CE9.5.3 | * Bio is conditionally disabled based on template matching cost on CU level, thr = 2(BDepth-9) * Same applied at Sub-CU level, thr = 3x2(BDepth-10) | JVET-K0344  Xiaoyu Xiu  (InterDigital) |
| CE9.5.4 | * Reference block size is extended from (w+7)x(h+7) to (w+11)×(h+11) by using boundary padding. * Gradients are directly calculated based on interpolated prediction signal using a single 5-tap filter ([2, -9, 0, 9, 2]) | JVET-K0119  Chao-Hsiung Hung  (Qualcomm) |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Tester** | **VTM** | | | | | **BMS** | | | | |
|  | | **Y** | **U** | **V** | **EncT** | **DecT** | **Y** | **U** | **V** | **EncT** | **DecT** |
| AHG 13 | DMVR in BMS according to AHG13  (Test is DMVR off) | -2.65% | -2.54% | -2.67% | 109% | 131% | 0% | 0% | 0% | 100% | 100% |
| 9.5.2 | Ching-Yeh Chen (MediaTek) | -2.86% | -0.94% | -0.65% | 108% | 146% | -1.28% | -0.49% | -0.42% | 104% | 118% |
| With adaptive unit | -2.85% | -0.92% | -0.66% | 111% | 156% | -1.28% | -0.48% | -0.41% | 105% | 122% |
| Apply on chroma | -2.90% | -2.23% | -2.30% | 110% | 155% | -1.31% | -1.15% | -1.25% | 105% | 122% |
| With adaptive unit and apply on chroma | -2.90% | -2.22% | -2.30% | 113% | 168% | -1.31% | -1.16% | -1.23% | 106% | 127% |
| 9.5.3 | Xiaoyu Xiu (InterDigital) | -2.61% | -1.02% | -0.74% | 111% | 137% | -1.26% | -0.55% | -0.49% | 103% | 108% |
| 9.5.4 | Chao-Hsiung Hong(Qualcomm) | -2.94% | -1.07% | -0.76% | 116% | 166% | -1.49% | -0.58% | -0.49% | 106% | 130% |

Complexity characteristics:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Tester** | **Initial MV signalled** | **Sub-CU refinement** | **Neighboring recon. samples used** | **Max # of SAD calculation** | **Max. SR** | **Cost Function** | **Interpolation filter/tap no** | **Note** |
| 9.5.2 | Ching-Yeh Chen (MediaTek) | yes | yes | no | 0 | N.A | N.A | 3 tap filter  [-1, 0, 1]  for gradient,  DCTIF/8  for MC | SIMD = SSE42 anchor&test |
| With adaptive unit | yes | yes | no | 0 | N.A | N.A | 3 tap filter  [-1, 0, 1]  for gradient,  DCTIF/8  for MC | SIMD = SSE42 anchor&test |
| Apply on chroma | yes | yes | no | 0 | N.A | N.A | 3 tap filter  [-1, 0, 1]  for gradient,  DCTIF/8  for MC | SIMD = SSE42 anchor&test |
| With adaptive unit and apply on chroma | yes | yes | no | 0 | N.A | N.A | 3 tap filter  [-1, 0, 1]  for gradient,  DCTIF/8  for MC | SIMD = SSE42 anchor&test |
| 9.5.3 | Xiaoyu Xiu (InterDigital) | yes | yes | no | 1 | N.A. | SAD | 6 tap filter for gradient,  6 tap filter for MC  (as in JEM7.0) | SIMD = AVX anchor&test |
| 9.5.4 | Chao-Hsiung Hung (Qualcomm) | yes | yes | no | 0 | N.A. | N.A. | 5 tap filter  [2, -9, 0, 9, 2]  for gradient  DCTIF/8  for MC | SIMD = AVX anchor&test |

The test of BMS is with DMVR on. This demonstrates that BIO still has additive gain in the range of 1.3% when combined with other tools targeting the same aspects.

Proposal 9.5.3, when put on top of VTM, has very similar performance, and also very similar encoder and decoder run time as the current BMS-DMVR. However, it does not have the latency problem that the latter has (where we know from CE9.2 that this could be managed without losing gain or increasing run time).

On the other hand, a detailed worst case complexity analysis on BIO has never been made. What is the worst number of computations (including divisions), and worst case additional memory accesses.

It was later reported (see under JVET-K0485) that the worst case complexity of 9.5.3 is the same as “JEM BIO”, which is unacceptable high according to the analysis in JVET-K0485. Based on that information, the proposal should not be adopted to BMS.

[JVET-K0067](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3568) CE9: Motion vector difference signs derivation (Test 4.2) [S. Ikonin, J. Chen (Huawei)]

[JVET-K0088](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3591) CE9: Template matching based reordering and refining (CE9-3.7, CE9-3.8 and CE9-4.1) [N. Park, J. Nam, H. Jang, J. Lee, S. Kim (LGE)]

[JVET-K0119](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3625) CE9: BIO gradient calculation improvement (Test 9.5.4) [C.-H. Hung, W.-J. Chien, M. Karczewicz (Qualcomm)]

[JVET-K0143](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3649) CE9.4.3: Template Matching based Adaptive Merge Candidate Reorder [N. Zhang, X. Chen, Y. Lin, J. Zheng (HiSilicon)]

[JVET-K0168](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3676) CE9: Template Matched Merge (Test 9.3.1) [H. Lee, J. Kang, S.-C. Lim, J. Lee, H. Y. Kim (ETRI)]

[JVET-K0177](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3685) CE9: Bilateral matching (Test 9.2.3) [J. Li, C. Lim (Panasonic)]

[JVET-K0178](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3686) CE9: Template matching (Test 9.3.2) [J. Li, C. Lim(Panasonic)]

[JVET-K0199](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3708) CE9: Simplification of DMVR (Test 9.1.1) [X. Chen (HiSilicon), Semih Esenlik (Huawei), J. Zheng (HiSilicon)]

[JVET-K0487](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4017) Cross-check of JVET-K0199: CE9: Simplification of DMVR(Test 9.1.1) [T. Zhou, T. Ikai (Sharp)] [late]

[JVET-K0200](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3709) CE9: DMVR extension based on template matching(Test 9.3.6) [X. Chen, J. Zheng (HiSilicon)]

[JVET-K0214](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3724) CE9.3: Template matching from J0022 (Test 9.3.4 and 9.3.5) [A. Robert, T. Poirier, F. LeLeannec (Technicolor)]

[JVET-K0216](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3726) CE9: DMVR with Motion Vector Difference Mirroring (Test 1.6) [S. Esenlik, I. Krasnov, Z. Zhao, J. Chen (Huawei), Y. Li (USTC)]

[JVET-K0217](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3727) CE9: DMVR with Bilateral Matching (Test 2.9) [S. Esenlik, I. Krasnov, Z. Zhao, M. Xiang, H. Yang, J. Chen (Huawei), Y. Li (USTC)]

[JVET-K0253](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3763) CE9.1.2: DMVR early termination [Y.-C. Su, T.-D. Chuang, C.-Y. Chen, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-K0254](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3764) CE9.2.1: Bilateral matching merge mode [T.-D. Chuang, C.-Y. Chen, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-K0255](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3765) CE9.5.2: BIO with simplified gradient calculation, adaptive BIO granularity, and applying BIO to chroma components [C.-Y. Chen, C.-Y. Lai, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-K0303](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3817) CE9.2.7: Asymmetric Bilateral Matching for FRUC merge mode [B. Choi, F. Bossen, K. Misra, A. Segall (Sharp)]

[JVET-K0342](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3859) CE9.1.3: Complexity reduction on decoder-side motion vector refinement (DMVR) [X. Xiu, Y. He, Y. Ye (InterDigital)]

[JVET-K0343](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3860) CE9.2.2: Simplifications on bilateral matching mode [X. Xiu, Y. He, Y. Ye (InterDigital)]

[JVET-K0344](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3861) CE9.5.3: Bi-directional optical flow (BIO) simplification [X. Xiu, Y. He, Y. Ye (InterDigital)]

[JVET-K0358](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3876) CE9.1.5: MVD-based Early-skip Condition for DMVR [C.-C. Chen, W.-J. Chien, M. Karczewicz (Qualcomm)]

[JVET-K0359](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3877) CE9.2.5/9.2.6: DMVR with Template-free Bilateral Matching [C.-C. Chen, W.-J. Chien, M. Karczewicz (Qualcomm)]

[JVET-K0378](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3897) CE9-Test 2.8: DMVR with bilateral matching and 2 half-pel points search [Y. Li, D. Liu (USTC)]

[JVET-K0437](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3960) Crosscheck for CE9-2.6 [W. Zhang (Hulu)] [late]

[JVET-K0438](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3961) Crosscheck for CE9-2.7 [W. Zhang (Hulu)] [late]

## CE10: Combined and multi-hypothesis prediction (9)

Contributions in this category were discussed Friday 13 July 1600–XXXX (chaired by JRO except otherwise noted)..

[JVET-K0030](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3848) CE10: Summary report on combined and multi-hypothesis prediction [C.-W. Hsu, M. Winken, X. Xiu]

A summary of Core Experiment 10 (CE10) on combined and multi-hypothesis prediction is reported. Four sub CEs are created to test different methods of combined predictions, including CE10.1: multi-hypothesis prediction, CE10.2: overlapped block motion compensation, CE10.3: non-rectangular partitions and CE10.4: diffusion filtering of inter- and intra-prediction signals. In CE10.1, one out of 10 tests was withdrawn and in CE10.2, one out of 3 tests was withdrawn. So there are 9, 2, 3 and 6 tests for each sub CE, respectively. All tests are evaluated based on the common test conditions defined in JVET-J1010. All tests and crosscheck results are integrated in this report.

**CE10.1 Multi-hypothesis prediction**

In CE10.1, the goal is to test prediction to be combined coming from multiple hypotheses, where one hypothesis refers to prediction from inter mode or from intra mode. The tests and corresponding results are summarized as follows,

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| # | Proposal | Tester | Supported modes | Hypothesis type | Signaling of hypothesis | # of extra hypothesis | Block constraint in luma samples |
|
| CE10.1.1 | JVET-K0257 | Chih-Wei Hsu (MediaTek) | AMVP uni prediction | inter | merge index | 1 | >= 8x8 |
|  |  |
|  |  |
| CE10.1.2 | JVET-K0257 | Chih-Wei Hsu (MediaTek) | skip | inter | implicitly derived | 1 or 2 |  |
|  | merge |  |
|  |  |  |
| CE10.1.3 | JVET-K0257 | Chih-Wei Hsu (MediaTek) | merge | intra | intra mode index | 1 |  |
|  |  |
|  |  |
| CE10.1.4 | JVET-K0257 | Chih-Wei Hsu (MediaTek) | skip | inter | merge index for uni-prediction + implicitly derived + intra mode index | 1 or 2 | >= 8x8 |
|  | merge | intra |  |
|  | AMVP uni prediction |  |  |
| CE10.1.5 | JVET-K0269 | Martin Winken (HHI) | merge | inter | ref index + MVP index + MVDs +  weights | 1 |  |
|  | AMVP |  |
|  |  |  |
| CE10.1.6 | JVET-K0269 | Martin Winken (HHI) | merge | inter | ref index + MVP index + MVDs +  weights | 1 or 2 |  |
|  | AMVP |  |
|  |  |  |
| CE10.1.7 | JVET-K0269 | Martin Winken (HHI) | merge | inter | ref index + MVP index + MVDs +  weights | 1 | > 8x8 |
|  |  |  |
|  |  |  |
| CE10.1.8 | JVET-K0269 | Martin Winken (HHI) | merge | inter | ref index + MVP index + MVDs +  weights | 1 or 2 | > 8x8 |
|  | AMVP |  |
|  |  |  |
| CE10.1.9 |  | ***withdrawn*** |  |  |  |  |  |
|  |  |
|  |  |
| CE10.1.10 | JVET-K0147 | Na Chang (HiSilicon) | merge | inter | implicitly derived | 1 |  |
|  |  |
|  |  |

By restricting multi-hypothesis to block sizes >=8, the multi-hypothesis prediction does not have worse memory bandwidth requirements than VTM with 4x4.

For each additional hypothesis, another prediction needs to be generated, i.e. the computational complexity would e.g. double in case of uni prediction with 1 additional hypotheses, or bi prediction with 2 additional hypotheses. Each additional hypothesis is then superimposed (with weighted superposition)

The superposition weights are fixed for test 1-4, and can be varied for tests 5-8.

Major differences are:

* Test 1-4 uses fixed weighting 5/8 and 3/8, test 5-8 switches between 2 different weights
* Test 1-4 allows combining inter and intra

Test 10 generates a second reference for LDP, for which the MV is derived. This requires same number of reference computations as LDB. According to proponents, this performs worse than LDB as such, but has faster encoder.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| # | Config. | VTM | | | | | BMS | | | | |
| Y | U | V | EncT | DecT | Y | U | V | EncT | DecT |
| CE10.1.1 | RA | -0.26% | -0.26% | -0.20% | 109% | 102% | -0.28% | -0.33% | -0.29% | 105% | 101% |
| LB | -0.12% | -0.16% | -0.16% | 112% | 102% | -0.17% | -0.39% | -0.23% | 107% | 102% |
| LP |  |  |  |  |  |  |  |  |  |  |
| CE10.1.2 | RA | -0.76% | -0.75% | -0.71% | 110% | 105% | -0.68% | -0.87% | -0.87% | 105% | 107% |
| LB | -0.40% | -0.38% | -0.52% | 115% | 104% | -0.51% | -0.67% | -0.73% | 109% | 102% |
| LP |  |  |  |  |  |  |  |  |  |  |
| CE10.1.3 | RA | -0.68% | -0.68% | -0.51% | 112% | 104% | -0.62% | -0.30% | -0.33% | 104% | 102% |
| LB | -0.58% | -1.14% | -1.14% | 113% | 103% | -0.56% | -0.97% | -1.02% | 103% | 102% |
| LP |  |  |  |  |  |  |  |  |  |  |
| CE10.1.4 | RA | -1.29% | -1.45% | -1.29% | 119% | 106% | -1.13% | -1.15% | -1.15% | 107% | 105% |
| LB | -0.76% | -1.32% | -1.55% | 126% | 105% | -0.77% | -1.18% | -1.13% | 111% | 101% |
| LP |  |  |  |  |  |  |  |  |  |  |
| CE10.1.5 | RA | -1.63% | -1.07% | -1.07% | 127% | 104% | -0.92% | -0.83% | -0.83% | 107% | 100% |
| LB | -2.05% | -0.39% | -0.44% | 145% | 107% | -0.93% | -0.47% | -0.39% | 114% | 99% |
| LP |  |  |  |  |  |  |  |  |  |  |
| CE10.1.6 | RA | -1.91% | -1.24% | -1.19% | 130% | 104% | -1.07% | -0.91% | -0.95% | 108% | 100% |
| LB | -2.29% | -0.39% | -0.43% | 147% | 108% | -1.13% | -0.36% | -0.11% | 115% | 99% |
| LP |  |  |  |  |  |  |  |  |  |  |
| CE10.1.7 | RA | -1.55% | -1.07% | -1.02% | 122% | 104% | -0.89% | -0.78% | -0.84% | 106% | 100% |
| LB | -2.04% | -0.36% | -0.54% | 135% | 107% | -0.94% | -0.35% | -0.09% | 112% | 98% |
| LP |  |  |  |  |  |  |  |  |  |  |
| CE10.1.8 | RA | -1.84% | -1.29% | -1.21% | 125% | 105% | -1.03% | -0.88% | -0.91% | 107% | 100% |
| LB | -2.27% | -0.32% | -0.48% | 138% | 108% | -1.07% | -0.29% | -0.12% | 114% | 100% |
| LP |  |  |  |  |  |  |  |  |  |  |
| CE10.1.10 | RA | 0.00% | 0.00% | 0.00% | 101% | 100% | 0.00% | 0.00% | 0.00% | 100% | 100% |
| LB | 0.00% | 0.00% | 0.00% | 101% | 100% | 0.00% | 0.00% | 0.00% | 100% | 97% |
| LP | -3.41% | -2.07% | -1.78% | 110% | 99% | -1.67% | -1.33% | -1.02% | 100% | 95% |

Question: What would be the effect if only intra and inter are combined? This would be

10.1.3 (combining bi pred and one intra pred), which gives 0.6% gain.

Generally, this experiment provides interesting gain, but requires additional computations (depending on variant), where some of the variants require more memory bandwidth than others. Gain decreases when used in BMS. Should be further studied in combination with other methods of improving motion comp, e.g. improved merge. Further reduction of encoder run time would be desirable as well.

**Test 10.2: OBMC**

In CE10.2, the goal is to test prediction to be combined from using motions of neighboring coding units (CUs). The tests and corresponding results are summarized as follows,

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| # | Proposal | Tester | # of blending lines | Blending order | Blending order (sequential/parallel) | Runtime reduction technique |
|
| CE10.2.1 |  | ***withdrawn*** |  |  |  |  |
|  |
| CE10.2.2 | JVET-K0345 | Xiu, Xiaoyu (InterDigital) | 2: CU area <64 or 4x4 sub CU 4: otherwise | Phase 1 : T->L (CU boundary) | Sequential | MV merge skip similar MVs |
| Phase 2 : T->L->B->R (other sub CU boundaries) |
| CE10.2.3 | JVET-K0213 | Antoine Robert (Technicolor) | 2: CU width or height < 8 4 for one side or 2 for both sides: otherwise | (T,B)->(L,R) | Sequential |  |
|
| JEM OBMC | JVET-K0258 |  | 2: CU area <64 or 4x4 sub CU 4: otherwise | T->L->B->R | Sequential | N/A |
|

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| # | Config. | VTM | | | | | BMS | | | | |
| Y | U | V | EncT | DecT | Y | U | V | EncT | DecT |
| CE10.2.2 | RA | -0.96% | -1.91% | -1.98% | 107% | 111% | -1.14% | -2.14% | -2.24% | 102% | 108% |
| LB | -1.37% | -1.80% | -1.90% | 110% | 112% | -1.80% | -2.92% | -2.60% | 105% | 115% |
| CE10.2.3 | RA | -1.01% | -2.03% | -2.09% | 114% | 125% | -1.24% | -2.22% | -2.37% | 107% | 131% |
| LB | -1.40% | -1.93% | -2.07% | 117% | 120% | -1.94% | -2.98% | -2.77% | 109% | 136% |
| JEM OBMC | RA | -1.03% | -2.03% | -2.08% | 113% | 123% | -1.22% | -2.18% | -2.28% | 105% | 128% |
| LB | -1.36% | -1.80% | -1.90% | 116% | 124% | -1.87% | -3.03% | -2.73% | 108% | 136% |

The worst case number of computations (e.g. for interpolation) in the prediction is likely more than doubled, also the memory accesses are likely more than doubled. Considering that, no direct action follows from CE proposals. There are CE related proposals (K0259, K0258) which target reduction of memory accesses and computations by using padding.

**Test 10.3: Non-rectangular partitions**

(chaired by C.-W. Hsu)

In CE10.3, the goal is to test prediction to be combined from non-rectangular prediction partitions within one CU. The tests and corresponding results are summarized as follows,

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| # | Proposal | Tester | Supported modes | Prediction type | Partitioning | Block constraint in luma samples | Note |
|
| CE10.3.1 | JVET-K0144 | Ru-Ling Liao (Panasonic) | skip | inter | Diagonal and inverse diagonal triangular |  |  |
| merge |  |
|  |  |
| CE10.3.2 | JVET-K0144 | Ru-Ling Liao (Panasonic) | skip | inter | Diagonal and inverse diagonal triangular | >= 8x8 |  |
| merge |
|  |
| CE10.3.3 | JVET-K0146 | Max Blaeser (RWTH Aachen University) | all, no skip | inter | wedge shaped |  |  |
| intra |
|  |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| # | Config. | VTM | | | | | BMS | | | | |
| Y | U | V | EncT | DecT | Y | U | V | EncT | DecT |
| CE10.3.1 | RA | -1.00% | -1.52% | -1.53% | 119% | 103% | -0.96% | -1.34% | -1.44% | 109% | 99% |
| LB | -1.70% | -2.03% | -2.21% | 125% | 103% | -1.44% | -2.12% | -1.84% | 113% | 100% |
| LP |  |  |  |  |  |  |  |  |  |  |
| CE10.3.2 | RA | -0.95% | -1.44% | -1.43% | 118% | 103% | -0.91% | -1.25% | -1.33% | 108% | 100% |
| LB | -1.63% | -1.94% | -1.89% | 122% | 102% | -1.41% | -1.74% | -1.63% | 112% | 101% |
| LP |  |  |  |  |  |  |  |  |  |  |
| CE10.3.3 | RA | -0.80% | -1.32% | -1.19% | 250% | 137% | -0.73% | -1.17% | -1.08% | 142% | 114% |
| LB | -0.27% | -0.40% | -0.46% | 545% | 160% | -0.31% | -0.53% | -0.43% | 239% | 134% |
| LP |  |  |  |  |  |  |  |  |  |  |

Normal transforms are used in all proposals

It is mentioned that there are CE related contributions (e.g. K0148, which combines 3.2 with sub experiment 1)

It is mentioned that diagonal partitions duplicate the memory access, as the external reference picture memory is usually accessed in 2D rectangular structures, and 2 rectangles would need to be fetched for the 2 adjacent diagonal partitions. A possible solution could be restriction to uni prediction. It was mentioned that it might be useful to investigate in a CE what benefit is achieved if diagonal partitioning is restricted to uni pred.

Further study in CE on memory impact, possible solutions to this, and interdependency with other tools.

**Test 10.4: Diffusion filtering**

In CE10.4, the goal is to test prediction to be combined using filtering, where two types of diffusion filters (uniform and signal dependent) with two iteration parameters are included. The tests and corresponding results are summarized as follows,

|  |  |  |  |
| --- | --- | --- | --- |
| # | Proposal | Tester | Description |
|
| CE10.4.1 | JVET-K0323 | Jennifer Rasch (HHI) | • Fast Encoder Decisions and restrictions for All Intra |
|
| CE10.4.2 | JVET-K0323 | Jennifer Rasch (HHI) | • Fast Encoder Decisions and Restrictions • Merging diffusion parameters |
|
|
| CE10.4.3 | JVET-K0323 | Jennifer Rasch (HHI) | • Fast Encoder Decisions and restrictions • Additionally sending diffusion parameters in merge case |
|
|
| CE10.4.4 | JVET-K0323 | Jennifer Rasch (HHI) | • More Extensive Search and released restrictions • Additionally sending diffusion parameters in merge case |
|
|
|
| CE10.4.5 | JVET-K0323 | Jennifer Rasch (HHI) | • Fast Encoder Decisions and restrictions • Additionally sending diffusion parameters in merge case • No neighboring block samples used |
|
|
| CE10.4.6 | JVET-K0323 | Jennifer Rasch (HHI) | • Fast Encoder Decisions and restrictions • Additionally sending diffusion parameters in merge case  • Restrict application of diffusion filter in inter mode |
|
|

Test 6 is used only for AMVR when it did integer or 4-pel accuracy, therefore with BMS only.

Mainly used for larger blocks.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| # | Config. | VTM | | | | | BMS | | | | |
| Y | U | V | EncT | DecT | Y | U | V | EncT | DecT |
| CE10.4.1 | AI | -0.92% | -1.15% | -1.19% | 157% | 116% | -0.74% | -1.63% | -1.56% | 140% | 104% |
|  |  |  |  |  |  |  |  |  |  |  |
| CE10.4.2 | RA | -1.13% | -1.50% | -1.49% | 112% | 109% | -0.67% | -1.33% | -1.31% | 120% | 98% |
| LB | -0.58% | 0.01% | -0.15% | 113% | 110% | -0.37% | -0.16% | -0.18% | 116% | 98% |
| LP |  |  |  |  |  |  |  |  |  |  |
| CE10.4.3 | RA | -1.26% | -1.67% | -1.69% | 126% | 107% | -0.79% | -1.46% | -1.51% | 130% | 98% |
| LB | -0.67% | 0.54% | 0.23% | 138% | 109% | -0.51% | -0.09% | 0.14% | 136% | 99% |
| LP |  |  |  |  |  |  |  |  |  |  |
| CE10.4.4 | AI | -1.03% | -1.29% | -1.31% | 200% | 118% | -0.75% | -1.68% | -1.64% | 162% | 104% |
| RA | -1.41% | -1.89% | -1.87% | 142% | 109% | -0.83% | -1.62% | -1.62% | 140% | 99% |
| LB | -0.69% | -0.08% | 0.17% | 148% | 109% | -0.50% | 0.23% | 0.07% | 141% | 99% |
| LP |  |  |  |  |  |  |  |  |  |  |
| CE10.4.5 | RA | -1.13% | -1.43% | -1.48% | 126% | 106% | -0.67% | -1.47% | -1.50% | 130% | 96% |
| LB | -0.48% | 0.22% | 0.13% | 136% | 107% | -0.30% | -0.08% | 0.01% | 137% | 97% |
| LP |  |  |  |  |  |  |  |  |  |  |
| CE10.4.6 | RA | n/a | n/a | n/a | n/a | n/a | -0.72% | -1.47% | -1.42% | 127% | 98% |
| LB | -0.53% | -0.07% | 0.09% | 130% | 100% |
| LP |  |  |  |  |  |

More analysis is requested about the worst case number of computations (filter size, also considering symmetries of the non-adaptive filter, and need for filter adaptation, etc.)

It is pointed out that replacing the adaptive filter by a switchable variant might be beneficial for complexity reduction.

Further study recommended.

[JVET-K0144](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3650) CE10: Triangular prediction unit mode (CE10.3.1 and CE10.3.2) [R.-L. Liao, C. S. Lim (Panasonic)]

[JVET-K0146](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3652) CE10: Results on Geometric block partitioning (Test 3.3) [M. Bläser, J. Sauer (RWTH Aachen)]

[JVET-K0147](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3654) CE10.1.10: Dual Merge Mode [N. Zhang, Y. Lin, Q. Yu, J. Zheng (HiSilicon)]

[JVET-K0213](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3723) CE10.2: Generalized OBMC (Test 10.2.3) [A. Robert, T. Poirier, F. LeLeannec (Technicolor)]

Non-adjacent spatial candidates are added.

* Derivation of new candidates
  + The search grid is based on block width and block height, with maximum search range 96
  + Total (96 / max (width, height)) search points are checked. The detailed search pattern for each round is described in the figure above.
  + When max (width, height) is greater than the threshold (64), the search grid is 32x32
* Redundancy checking (in a different way for other merge candidates pruning) is performed for the added merge candidates.
* New candidates are added after TMVP candidates in the merge candidate list.
* Maximum merge candidate number is 10 or 8 in VTM 1.0 and 11 or 9 in BMS 1.0.

[JVET-K0257](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3767) CE10.1: Combined and multi-hypothesis prediction [M.-S. Chiang, C.-W. Hsu, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-K0269](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3780) CE10: Multi-hypothesis inter prediction (Tests 1.5-1.8) [M. Winken, H. Schwarz, D. Marpe, T. Wiegand (HHI)]

[JVET-K0323](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3837) CE 10: Signal Adaptive Diffusion Filters For Video Coding (Test 10.4.1-10.4.5) [J. Rasch, J. Pfaff, M. Schäfer, A. Henkel, H. Schwarz, M. Siekmann, M. Winken, P. Helle, D. Marpe, T. Wiegand (Fraunhofer HHI)]

[JVET-K0345](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3862) CE10.2.2: Complexity reduction for over-lapped block motion compensation (OBMC) [X. Xiu, Y. He, Y. Ye (InterDigital)]

## CE11: Composite reference pictures (4)

Contributions in this category were discussed Friday 13 July 1820–1940 (Track B chaired by JRO).

[JVET-K0031](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3658) CE11: Summary report on composite reference pictures [X. Zheng, G. Li, Y. Li]

This contribution is a summary report of Core Experiment 11 on composite reference picture. Three tests categories and 11 subtests were agreed to carry out in CE11 in between JVET-J and JVET-K meeting cycle, to study and evaluate technologies related to composite reference picture.

Follow by common test condition recommend from J meeting, BMS1.0 and BMS1.0 with VTM configurations are used to evaluate CE11 technologies. Test conditions are specified for each category. The corresponding coding performance of each coding tool under evaluated in CE11 are summarized in this contribution. To further evaluate CE11 tools, crosschecking reports are also integrated in this contribution.

There are three test categories to be evaluated in CE11. Test 1 targets at the evaluation of the performance when CTU-level refresh rate is set to unlimited, 1/2, 1/8 and 1/10 per frame. Since composite reference might have higher coding performance when composite reference is replaced with IDR at a longer GOP, test 2 targets at the test of larger random access GOP whose IDR period is set to two seconds and five seconds. The best block update refresh rate conducted at test 1 is used as the default refresh rate at test 2. Test 3 explores the coding efficiency by using HEVC long-term reference mechanism.

Note: J0011=K0156; J0032=K0370

|  |  |  |  |
| --- | --- | --- | --- |
| Test # | Description | Tester | Cross-checker |
| CE11.1.1 | No limitation on block update refresh rate at J0011 | Xiaozhen Zheng (DJI) | Wenhao Zhang (Hulu) |
| CE11.1.2 | 1/2 block update refresh rate at J0011 | Xiaozhen Zheng (DJI) | G. Li (Tencent) |
| CE11.1.3 | 1/8 block update refresh rate at J0011 | Xiaozhen Zheng (DJI) | G. Li (Tencent) |
| CE11.1.4 | 1/10 block update refresh rate at J0011 | Xiaozhen Zheng (DJI) | Yue Li (USTC) |
| CE11.1.5 | No limitation on block update refresh rate at J0011 | Yue Li (USTC) | G. Li (Tencent) |
| CE11.1.6 | 1/2 block update refresh rate at J0032 | Yue Li (USTC) | G. Li (Tencent) |
| CE11.1.7 | 1/8 block update refresh rate at J0032 | Yue Li (USTC) | G. Li (Tencent) |
| CE11.1.8 | 1/10 block update refresh rate at J0032 | Yue Li (USTC) | Xiaozhen Zheng (DJI) |
| CE11.2.1 | Intra period as two seconds at J0011 | Xiaozhen Zheng (DJI) | G. Li (Tencent) |
| CE11.2.2 | Intra period as five seconds at J0011 | Xiaozhen Zheng (DJI) | Yue Li (USTC) |
| CE11.2.3 | Intra period as two seconds at J0032 | Yue Li (USTC) | G. Li (Tencent) |
| CE11.2.4 | Intra period as five seconds at J0032 | Yue Li (USTC) | Xiaozhen Zheng (DJI) |
| CE11.3.1 | HEVC encoder only long-term reference mechanism (K0157) | Xiaozhen Zheng (DJI) / Yue Li (USTC) | Wei-Jung Chien (Qualcomm) |

The followings are a summary table of the tests in this CE.

Table 1: CE11 test results against VTM/BMS anchor (lowdelay B main10)



Table 2: CE11 test results against VTM/BMS anchor (random access main10)



Question: Why worse for RA? Answer: The reference would need to be newly generated for each IDR period, and at least for the first GOP of B pictures it cannot be used.

It is pointed out by one expert that strategies exist which would still allow this to some extent.

Test 3 uses the HEVC long-term reference mechanism, in combination with signaling of no output coded pictures (pic\_output\_flag = 0). This brings comparable ore even higher gain as with the other two methods. Only one additional reference picture is generated.

The average bit rate reduction is even slightly higher for BMS than it is for VTM.

The picture is built using a mechanism for static background detection, and put areas that are likely from static background. Therefore, gain is highest for sequences with static background and occlusions, e.g. Cactus, Basketball, and class E.

Even though this is currently specific for a certain type of sequences, and only has benefit for LDB, the approach of JVET-K0157 is interesting as a non-normative add-on in the encoder.

Decision (SW): Add software from JVET-K0157 as non-normative tool in VTM (non CTC). Disable motion scaling part. The proponents should also be asked to provide software for HM.

From the results of CE11, the other two proposals which would require block-level signalling do not show substantial benefit over the “long-term reference + no-output picture” solution from HEVC. They also provide most gain for sequences with static background and non-moving camera. There are also CE-related contributions that suggest additional enhancements.

Generally, it would be interesting to have the benefit of composite reference pictures extended to other cases, in particular moving cameras.

From CE related, no superior methods compared to CE11. Discontinue CE11

[JVET-K0156](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3664) CE11: Results on composite reference picture (test 11.1.1, 11.1.2, 11.1.3, 11.1.4, 11.2.1 and 11.2.2) [W. Li, X. Zheng (DJI)]

Usually, background areas have few motions in a long temporal window. Therefore, blocks with minor difference between the background and the current frame are picked up, and are used to replace the co-located blocks in a long-term reference. The proposed update method targets at the renewal of the background information. An indication flag is signalled at CTU level to indicate whether current CTU is used to update the long-term reference at decoder side. After a picture is decoded and reconstructed completely, the process of updating the long-term reference will be performed. For every CTU marked to update the long-term reference, its luma and chroma reconstructed pixels will be used to replace the co-located pixels in the long-term reference.

In the decoding process, if the long-term reference is used as reference, motion vector scaling and decoder motion refine operation that use motion trajectory are invalid because the distance between the long-term reference and the current slice is not available and motion trajectory model doesn’t work for long-term reference. Therefore, the tools like BIO、DMVR、FRUC are set to disable if any of the motion vector is referred to the long-term reference.

[JVET-K0157](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3665) CE11: HEVC-like encoder only solution for composite reference picture [W. Li, X. Zheng (DJI)]

Different to CE11 test 1 and test 2 that introduces a new composed virtual reference frame, CE11 test 3 targets at evaluating the use of the HEVC long-term reference mechanism, potentially in combination with signaling of no output coded pictures (pic\_output\_flag = 0). This combination could theoretically achieve similar functionality as that provided by composite reference pictures, e.g. by synthesizing and signaling a no-output reference picture that only contains background information.

the frames I0, B1, B2, B3 and B4 are coded as short-term reference. During the encoding of those short-term frames, an alternative frame is composed by exploiting previous coded frames’ content. When an encoder determines the alternative frame has been completely constructed, such frame will be coded as long-term reference with pic\_output\_flag=0. To harmonize current coding tools in VTM and BMS, the tools with motion vector scaling and motion vector refinement at decoder side are modified when they use reference data from such long-term reference.

In order to implement encoder only solution, RPS cfg setting is changed.

Meanwhile, HEVC syntax elements long\_term\_ref\_pics\_present\_flag, num\_long\_term\_ref\_pics\_sps, Output\_flag\_present\_flag, deblocking\_filter\_control\_present\_flag, pps\_deblocking\_filter\_disabled\_flag, short\_term\_ref\_pic\_set\_sps\_flag, deltaRPS, ref\_idcs, inter\_ref\_pic\_set\_prediction\_flag, delta\_rps, used\_by\_curr\_pic\_flag[j], use\_delta\_flag[j], num\_long\_term\_pics, used\_by\_curr\_pic\_lt\_flag, num\_ref\_idx\_active\_override\_flag and num\_ref\_idx\_l0\_active\_minus1are exploited and modified.

[JVET-K0370](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3888) CE11: Block-composed Background Reference (BCBR) [C. Ma, D. Liu, Y. Li, F. Wu (USTC)]

In the decoder side, if BCBR\_enable\_flag is true, a long-term reference picture, which is also the synthesized background reference picture, is appended into the reference picture list, i.e the number of reference picture list is increased by 1. The reference index for the added reference picture is equal to the number of reference picture minus 1, and all the motion compensation processes remain unchanged. The background reference picture is initialized using the reconstructed I frame. Then, a reconstructed CTU will be substitute the collocated one in the background reference picture if its background\_flag\_ctu is true.

In the encoder side, the flowchart is shown as follows, which contains three parts : background block selection, coding parameter decision, and background reference updating. Temporal and spatial correlation constrains are used to select the background block. The coding parameter for the background CTU is decided according the following equation:

where is the coding parameter of the decided background CTU, is the coding parameter of I picture, is the sequence length for LDB and LDP configurations, the interval between adjacent I pictures for RA configurations, is the number of the encoded pictures after encoding one I picture. After encoding current picture, the background reference picture in the location of the selected CTUs will be updated with the reconstructed ones. A flag signalling the background CTU is transmitted to the decoder. Besides, the coding parameter of the background CTU is transmitted to the decoder through the DQP technology.

[JVET-K0439](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3962) Crosscheck for CE11-1.1 [W. Zhang (Hulu)] [late]

## CE12: Mapping for HDR content (4)

Contributions in this category were discussed XXday XX July XXXX–XXXX (chaired by XXX).

[JVET-K0032](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3810) CE12: Summary report on HDR coding [E. Francois, D. Rusanovskyy, P. Yin]

[JVET-K0298](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3812) CE12: Report of dynamic range adaptation (DRA) and DRA refinement [E. Francois (Technicolor), D. Rusanovskyy (Qualcomm)] [late]

[JVET-K0308](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3822) CE12: HDR In-loop Reshaping (CE12-5, 12-6, 12-7 and 12-8) [T. Lu, F. Pu, P. Yin, W. Husak, S. McCarthy, T. Chen (Dolby)]

[JVET-K0392](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3912) Cross-check for CE12.6.1 and CE12.6.2 [J. Zhao, K. Misra] [late]

## CE13: Projection formats (8)

Contributions in this category were discussed XXday XX July XXXX–XXXX (chaired by XXX).

[JVET-K0033](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3853) CE13: Summary report on projection formats [P. Hanhart, J.-L. Lin]

[JVET-K0131](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3637) CE13: Modified Cubemap Projection in JVET-J0019 (Test 5) [Y.-H. Lee, J.-L. Lin, S.-K. Chang, C.-C. Ju (MediaTek)]

[JVET-K0182](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3691) CE13: Parallel-to-Axis Uniform cubemap projection (PAU) in JVET-J0033 (Test 7) [Y. Sun, X. Huangfu, B. Wang, L. Yu (Zhejiang Univ.)] [late]

[JVET-K0328](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3842) CE13: Cubemap projection (Tests 2.1 and 2.2) [P. Hanhart, Y. He, Y. Ye (InterDigital)]

[JVET-K0329](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3843) CE13: Equi-angular cubemap projection (Tests 3.1 and 3.2) [P. Hanhart, Y. He, Y. Ye (InterDigital)]

[JVET-K0330](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3844) CE13: Hybrid angular cubemap projection (Tests 4.1 and 4.2) [P. Hanhart, Y. He, Y. Ye (InterDigital)]

[JVET-K0331](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3845) CE13: Adaptive frame packing (Tests 4.3 and 4.4) [P. Hanhart, Y. He, Y. Ye (InterDigital)]

[JVET-K0387](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3906) CE13: Rotated Sphere Projection (Tests 8.1, 8.2 and 8.3) [C. Pujara, A. Singh, A. Konda (Samsung)] [late]

# Non-CE Technology proposals

## CE1 related – Partitioning (9)

Contributions in this category were discussed XXday XX July XXXX–XXXX (chaired by XXX).

[JVET-K0145](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3651) Non-CE1: On Transform Unit Partition-Uniform Transform Unit Structure [J. Zhu, J. Yao, W. Cai, K. Kazui (Fujitsu)]

This contribution was discussed Saturday 14 July 1610 (GJS).

This contribution proposes a “uniform TU” (UTU) structure. A CU is proposed to be partitioned into TUs uniformly, i.e. each TU in a CU has same size. A syntax element, utu\_mode, would be signalled in CU syntax. When utu\_mode is zero, it would mean no partitioning. Otherwise, the value of utu\_mode would indicate the partition structure. The UTU structure is only applied on intra CUs. Test results reportedly show gain of 0.65 % (Y), 1.26% (Cb) and 1.5% (Cr) in the case of UTU only performed on the luma component of I-slices. These results are from shortened (40-frame) test sequences.

The prediction process would operate on a TU basis (as in HEVC) rather than on the CU basis.

The encoding time is roughly doubled and the decoding time is increased about 10%. The contributor said the amount of decoder increase may be primarily a code optimization issue, estimating that the increase should really be about 4%.

It was commented that since this is splitting the tree deeper, it should be compared to using a deeper tree depth, which also provides gain.

Further study is requested.

[JVET-K0464](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3993) Crosscheck of JVET-K0145: Non-CE1:On Transform Unit Partition-Uniform Transform Unit Structure [P.-H. Lin, C.-H. Yao, S.-P. Wang, C.-C. Lin, C.-L. Lin (ITRI)] [late]

[JVET-K0220](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3730) Non-CE1: Proposal for a partitioning method by Fraunhofer HHI and Technicolor [J. Ma, A. Wieckowski, H. Schwarz, D. Marpe, T. Wiegand (HHI), F. Le Léannec, T. Poirer (Technicolor)]

This contribution was discussed Saturday 14 July 1445 (GJS).

This contribution proposes a partitioning scheme for VVC as a combination of different partitioning aspects tested in the Core Experiment 1: Partitioning (JVET-J1021). The proposed partitioner is configurable to reach different trade-off points.

The VTM encoder uses a maximum BTT depth of 3 (although the decoder also supports other depths and the encoder can be configured differently as well).

About 0.5%/0.7%/0.9% for AI/RA/LB gain was measured (configuration “C2”) without changing the partitioning structure or tree searching depth and with faster encoding, although changing the syntax and changing the boundary handling. It was estimated that about 0.3% for RA was from the boundary handling. It was commented that the encoder optimization and boundary handling rather than the other syntax difference seemed likely to be primarily responsible for the difference.

With the addition 4-way splits included, an additional 0.0%/0.2%/0.5% gain was reported.

It was noted that the particular variation of QT+BTT had been intended to be a "placeholder" with no presumptive status.

It was commented that the decoding time reported in the contribution had increased by 7% for AI.

This was further discussed Saturday 2000 (GJS). The proponent indicated that they had concluded there was no difference between the proposal without the additional 4-way splits and the current VTM QT+BTT and offered their encoder optimization for the current design. This was welcomed.

Decision: It was agreed that the QT+BTT as per draft 1 now *does* have presumptive status; it is not just a placeholder.

[JVET-K0434](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3957) Crosscheck of JVET-K0220 Non-CE1: Proposal for a partitioning method by Fraunhofer HHI and Technicolor [X. Li (Tencent)] [late]

[JVET-K0230](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3740) CE1-related: Separate tree partitioning at 64x64-luma/32x32-chroma unit level [T.-D. Chuang, C.-Y. Chen, Y.-W. Huang, S.-M. Lei (MediaTek)]

In a typical pipelined hardware decoder architecture, the data is pipelined with NxN luma blocks and MxM chroma blocks, where NxN and MxM are the same as the maximum luma transform block (TB) size and the maximum chroma TB size. In VTM-1.0, the maximum luma and chroma TB sizes are 64x64 and 32x32. In separate tree partitioning, for each coding tree unit (CTU), the luma coding tree block (CTB) is first signalled, then the chroma CTBs are signalled, which makes hardware decoder architecture to change from maximum TB pipelining to CTU pipelining (i.e., the data is pipelined with luma CTBs and chroma CTBs). In VTM-1.0, the luma CTB size is 128x128, and the chroma CTB size is 64x64. The number of samples for a decoder pipeline stage to support CTU-level separate tree partitioning is four times of that without separate tree partitioning, which will lead to significant increase of hardware areas for pipeline stages. In this contribution, it is proposed to start separate tree partitioning starting at 64x64-luma/32x32-chroma units instead of CTU. In the proposed method, each intra-slice CTU is first implicitly split into 64x64-luma/32x32-chroma units. Then the coding tree under each 64x64-luma/32x32-chroma unit is separated, and luma syntax is signalled before chroma syntax within each 64x64-luma/32x32-chroma unit. It is claimed that the proposed method can be easily supported by hardware decoder architecture of data pipelining with 64x64 luma blocks and 32x32 chroma blocks. Compared against the CTU-level separate tree partitioning, simulation results reportedly show negligible BD-rate differences for the 64x64-luma/32x32-chroma unit-level separate tree partitioning. It is also reported that the 64x64-luma/32x32-chroma unit-level separate tree partitioning with multiple intra chroma direct modes (multi-DMs) can achieve 0.99% and 0.81% Y BD-rates, 11.24% and 9.80% U BD-rates, 11.20% and 9.80% V BD-rates, 17% and 40% encoding time decreases for VTM-1.0-AI and BMS-1.0-AI, respectively, when both the anchor and the test include linear model (LM) chroma mode and the 65 intra angular modes. The impact on decoding time seems negligible.

It was commented that the current shared tree scheme also is not friendly to 64x64 pipeline architecture if there is a top-level split that is a ternary split.

It was suggested that we could just have a smaller maximum CU size for intra than inter. The contributor said this would likely have no impact on coding efficiency.

It was noted that there is an interaction with CCLM.

For separate tree operation, all luma would be sent before all chroma on a 64x64 basis or a CTU basis, whichever is smaller.

Decision: Adopt separate trees for intra slices (without multi-DMs) with an implicit split to 64x64 (into both VTM and BMS).

Decision: Prohibit ternary split of something bigger than 64 in width or height (and not send the bit to indicate ternary type at that level). See also later contribution K0556.

As a change of the software and CTC configuration, it was suggested to increase the chroma QP for intra when the trees are separate. The contributor had tested increasing the chroma QP offset by 1 for intra (with CCLM in both the anchor and test) and said this showed an increase to 3%/1.5%/0.4% over the VTM, that they would provide the test results in a revision of the contribution. Decision (SW & CTC): Agreed.

There was further discussion on Monday 1500 (GJS) about why separate trees are only planned for intra slices rather than also intra CTUs in inter slices. It was commented that the intra/inter switch is at the CU level rather than the CTU level, so in the current scheme the tree ends before the intra/inter decision is made. A proposed approach (JVET-K0354) that was studied in CE1 was to add a flag at the CU level that would continue the tree, and the two trees would separate from that point downward. It was remarked that the way current-picture referencing works means that if we only support separate trees for intra slices, we cannot combine separate tree usage with current picture referencing.

This was further discussed Tuesday 0945 (GJS). It was noted that supporting separate trees for intra CUs in inter slices shows little improvement in the CTC and has an encoder complexity impact, but it was reported that there was little decoder complexity impact, and enabling this would make intra more consistent between inter slices and intra slices (e.g., an encoder would not need to change the slice type in order to get access to the separate tree functionality, and some encoders may seldom use intra slices). There was some questioning of decoder complexity impact, but there was no clearly significant impact on decoder complexity).

In track A, it was initially agreed to enable separate trees for intra CUs in inter slices as described. There should a high-level (e.g., SPS) flag to enable or disable separate trees, and if enabled, there should be a flag to enable it or disable it in inter slices. The use in inter slices should be disabled in the CTC. Further study in a CE will be performed to confirm the decoder complexity impact.

This was further discussed in JVET plenary Tuesday 1330 (GJS & JRO). It was suggested that, in an inter slice, a flag at the whole-CTU level would indicate that the CTU is intra with separate trees. If the flag is zero, there would be the current scheme with a common tree. This and the other described scheme should be tested in a CE.

It was also suggested to consider a local switch at the CTU or CU level for whether a separate tree or a single tree or common tree is used. This should also be tested in the CE.

As a matter of design principle, it was agreed that having intra work similarly in an inter slice as in an intra slice is desirable (in the absence of some justification to do otherwise).

[JVET-K0402](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3922) Crosscheck of JVET-K0230: CE1-related: Separate tree partitioning at 64x64-luma/32x32-chroma unit level [X. Xu, J. Ye (Tencent)] [late]

[JVET-K0320](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3834) CE1-related: Zero-Unit for Picture Boundary Handling [K. Zhang, L. Zhang, H. Liu, Y. Wang, P. Zhao, D. Hong (Bytedance)]

Considered in BoG on picture boundary handling.

[**JVET-K0535**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4065) **Cross-check of JVET-K0320: CE1-related: Zero-Unit for Picture Boundary Handling [M. Xu (Tencent)] [late]**

[JVET-K0362](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3880) CE1-related: Context modeling for coding CU split decisions [S.-T. Hsiang, S.-M. Lei (MediaTek)]

This contribution proposes a modified method for entropy coding the coding unit (CU) split decisions. The proposed method reportedly reduces the total number of contexts by 1 and reduces the numbers spatial neighbouring CUs used context selection from four to two. The proposed method reportedly leads to 0.10%, 0.15%, and 0.14% luma BD-rate gains for the AI, RA, and LB settings, respectively, under the VTM-1.0 CTCs (with all BMS tools off). The proposed method reportedly leads to 0.10%, 0.12%, and 0.20% luma BD-rate gains for the AI, RA, and LB settings, respectively, under the BMS-1.0 CTCs (with all BMS tools on).

It was commented that although this looks logical, it seems like a very small refinement that is not necessary to consider at this time and changing it could interfere with other ongoing work. This should be kept in mind if it remains relevant as the project proceeds.

[JVET-K0414](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3935) CE1-related Cross-check of JVET-K0362 [S. Jeong (Samsung)] [late]

[JVET-K0366](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3884) CE1-related: Partial CU for picture boundary [M. Xu, X. Li, S. Liu (Tencent)]

Considered in BoG on picture handling.

[JVET-K0523](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4053) Cross-check of JVET-K0366: CE1-related: Partial CU for picture boundary [K. Zhang (Bytedance)] [late]

[JVET-K0497](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4027) CE1.4 related: Evidence of Split Unit Coding Order [Y. Piao, J. Chen, C. Kim (Samsung)] [late]

This contribution presents evidence of potential gain of split unit coding order (SUCO) described in CE1 subtest 1.4. Since the result of SUCO on VTM in JVET-K0133 is not consistent with results observed in other contexts, some evidence of potential gain from SUCO is presented in this contribution to justify further study of SUCO for VVC. SUCO on the HM reportedly provides 2.1% and 2.1% BD-rate gains in AI and RA, respectively. SUCO on JEM3.1 reportedly provides 2.9% gain in class A2 in RA configuration. In the IFVC software (JVET-J0072) of the Cfp response JVET-J0024, a maximum 1.7% gain was reported with less complexity in the IFVC software than in the HM and JEM because of encoder optimization.

Further study (although not currently in a CE, because it not yet clear exactly how to test it and there is an interaction with partitioning modifications) was encouraged.

[JVET-K0554](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4084) CE1-related: Joint proposal for picture boundary partitioning by Fraunhofer HHI and Huawei [A. Wieckowski, J. Ma, H. Schwarz, D. Marpe, T. Wiegand (HHI), H. Gao, S. Esenlik, J. Chen (Huawei)] [late]

Considered in BoG on picture handling.

[JVET-K0556](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4086) CE1-related: Constraint for binary and ternary partitions [C.-W. Hsu, T.-D. Chuang, C.-Y. Chen, Y.-W. Huang, S.-M. Lei (MediaTek)] [late]

Discussed Tuesday 1000 (GJS)

Virtual pipeline data units (VPDUs) are defined as non-overlapping MxM-luma(L)/NxN-chroma(C) units in a picture. In hardware decoders, successive VPDUs are processed by multiple pipeline stages at the same time; different stages process different VPDUs simultaneously. The VPDU size is roughly proportional to the buffer size in most pipeline stages, so it is said to be very important to keep the VPDU size small. In HEVC hardware decoders, the VPDU size is set to the maximum transform block (TB) size. Enlarging the maximum TB size from 32x32-L/16x16-C (as in HEVC) to 64x64-L/32x32-C (as in the current VVC) can bring coding gains, which results in 4X of VPDU size (64x64-L/32x32-C) expectedly in comparison with HEVC. However, in addition to quadtree (QT) coding unit (CU) partitioning, ternary tree (TT) and binary tree (BT) are adopted in VVC for achieving additional coding gains, and TT and BT splits can be applied to 128x128-L/64x64-C coding tree blocks (CTUs) recursively, which is said to lead to 16X of VPDU size (128x128-L/64x64-C) in comparison with HEVC. To reduce the VPDU size in VVC, a constraint for TT and BT is proposed, and the VPDU size is defined as 64x64-L/32x32-C for the following.

* Cond. 1: For each VPDU containing one or more CUs, the CUs are completely contained in the VPDU.
* Cond. 2: For each CU containing one or more VPDUs, the VPDUs are completely contained in the CU.

Proposed constraint: For each CTU, the above two conditions shall not be violated, and the processing order of CUs shall not leave a VPDU and re-visit it later.

For the current scheme, these constraints would be satisfied if three constraints are applied.

* Prohibit ternary split of edges longer than 64 (32 for chroma)
* Prohibit vertical split when width is 64 and height is 128 (half these for chroma)
* Prohibit horizontal split when width is 128 and height is 64 (half these for chroma)

It was reported that imposing these three constraints would have a significant coding efficiency impact. Another way to meet the constraint is to set MAX\_TT\_SIZE and MAX\_BT\_SIZE to 64, likely accompanied by increasing the BT/TT depth.

Further study in a CE is neeed to test some approaches and determine the coding efficiency impact.

## CE2 related – Loop filters (15)

Contributions in this category were discussed Sunday 15 July in Track B 0900–1220 (chaired by JRO).

[**JVET-K0042**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3543) **A study on the overlap in functionality between SAO and ALF [S. Sethuraman, Nijil K. (Ittiam)]**

This contribution provides a study report of disabling SAO in the presence of ALF to understand the coding loss, if any. The motivation is to understand the overlap in functionality between these two in-loop filtering stages with the aim to see whether SAO can be disabled when ALF is enabled so as to reduce the number of cascading in-loop filtering stages and thus reduce the internal memory needs. Since there have been some enhancements to SAO that have been proposed in CE2 experiments, three of the BDRATE improving enhancements have been included during the study. The study results show that the tool OFF BDRATE drop for this modified SAO averages 0.88% in luma under CTC, while the tool ON gain for SAO in the absence of ALF averages -2.0% in luma. The chroma BDRATE drops are higher (at ~3%). Since chroma ALF has only a single class, SAO seems to provide higher improvements in chroma than in luma. In informal visual evaluations that compared ALF-only against ALF+SAO based streams, the latter was seen to remove certain motion trail artifacts (due to the CTB level signaling), implying that SAO still provides perceivable visual quality improvements and additional efforts at improving ALF may be required before SAO can be dropped in the cascade of in-loop filtering stages. Also, other objective quality metrics need to be tried to see which one correlates better with the subjective visual quality in order to reduce the subjectivity in conclusions going forward.

Interesting study – rate gains of SAO and ALF are partially interdependent, however the impact on visual quality justifies that each has its own benefit.

[JVET-K0068](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3569) CE2 related: Hadamard Transform Domain Filter [V. Stepin, S. Ikonin, R. Chernyak, J. Chen (Huawei)]

This contribution proposes in-loop filter in 1D Hadamard transform domain which is applied on CU level after reconstruction and has multiplication free implementation. Proposed filter is applied for all CU blocks that meet the predefined condition and filter parameters are derived from the coded information. It is reported that for the random access configuration the proposed method provides 0.50% of luma BD-rate saving with 105% encoding time and 104% decoding time compared to VTM 1.0.

Filtering is done in the Hadamard domain, applying a weight to the Hadamard coefficients that depends on the coefficients and the quantization (somewhat similar as a Wiener filter frequency domain equation). Attenuation factor becomes lower for higher QP. Furthermore, threshold is applied.

Hadamard transform is only applied within the current transform block, not overlapped (same position in processing chain as bilateral filter)

Lookup table is required which is approx. 17 k bytes

Average gain in VTM is 0.5%, BMS 0.4%

Should be compared if it is interdependent with bilateral filter – include in the same sub CE as bilateral filter.

[JVET-K0201](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3710) Non-CE2: On SAO parameter signalling [G. Laroche, J. Taquet, C. Gisquet, P. Onno (Canon)]

This contribution presents a modification of the high level SAO parameters signaling. In addition to the SAO HEVC CTU level parameters derivation, the Frame, Line, Column, 2x2 CTU, 3x3 CTU, Temporal and Temporal 90° SAO parameters derivation are available at encoder side and signaling in the slice header. For the additional derivations, the traditional SAO Up and Left Merge flags are removed. In this contribution, only the parameters derivation is modified and the SAO filtering stays at CTU level. An average BDR YUV (14:1:1) gain is reported compared to VTM1.0 of -0.11%, -0.21%, -0.35%, -0.61% for respectively AI, RA, LDB and LDP configurations and an average BDR YUV (14:1:1) gain of -0.07%, -0.24%, -0.57%, -0.51% for respectively AI, RA, LDB and LDP configurations compared to BMS1.0.

Question is raised, as the proposal changes the granularity of SAO adaptation, does it have impact on the visual quality? Not investigated

The proposal also performs inheritance from the reference picture. This requires additional storage of SAO parameters which is undesirable. The fact that there is hardly any gain for AI could suggest that most of the gain comes from that.

The aspect of CTU grouping makes parameter optimization more complicated, kind of slice level optimization. Also conventional SAO could use a similar method with lookahead.

No action on this.

[JVET-K0453](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3978) Cross-check of JVET-K0201: Non-CE2: On SAO parameter signalling [F. Galpin, P. Bordes (Technicolor)] [late]

[JVET-K0202](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3711) Non-CE2: On SAO Edge Offset classification [G. Laroche, J. Taquet, C. Gisquet, P. Onno (Canon)]

This contribution presents a modification of the SAO Edge offset classification and offsets coding. The modification of the Edge offset classification is similar to those proposed in CE2-3.4 which consist in modifying the sign function used for the Edge offset category determination. In this contribution, when the same classification as CE2-3.4 is enabled, the peak and valley offsets are coded with an explicit sign signaling as Band offsets and the Luma offsets are predicted by a default value. Moreover, in a second modification, several sign functions are competing at encoder side and explicitly signaled in the bitstream. An average BDR YUV (14:1:1) gain is reported compared to VTM1.0 of -0.11%, -0.21%, -0.18%, -0.48% for respectively AI, RA, LDB and LDP configurations for the first modification and an average BDR YUV (14:1:1) gain of -0.12%, -0.26%, -0.37%, -0.95% for respectively AI, RA, LDB and LDP configurations for both modifications.

Almost no gain in BMS (likely due to interdependency with BMS) – as ALF is now in VTM, this would likely be the case for the next VTM as well.

Proponent is asked to investigate whether the sign in EO might have positive impact in visual quality.

[JVET-K0454](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3979) Cross-check of JVET-K0202: Non-CE2: On SAO Edge Offset classification [F. Galpin, P. Bordes (Technicolor)] [late]

[JVET-K0203](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3712) Non-CE2: Higher-precision modifications to VVC deblocking filters [C. Gisquet, J. Taquet, G. Laroche, P. Onno (Canon)]

This proposal describes three modifications to parameters of the deblocking filter. It first asserts that the tC parameter, linearly derived according to bitdepth from a table, suffers from a lack of precision and thus proposes a bitdepth-dependent scalar multiplication approach. Secondly, said parameter, used to clip the output of deblocking filters within a range of the filtered sample, is asserted to be too high compared to at least the distortion produced by the BMS. It then proposes to reduce the maximal value in the same fashion for all deblocking filters. Finally, it proposes a new condition on whether to filter a chroma edge, dependent on the flatness of the luma as measured by the luma deblocking filter on the corresponding luma edge. It reports achieving for Y -0.3%/-0.1%/-0.3% over the VTM anchor for respectively AI/RA/LDB, and -0.8%/-0.3%/-0.6% over the BMS anchor. For YUV, using 14:1:1 weights, these numbers are respectively -0.3%/-0.3%/-0.4% and -0.9%/-0.5%/-0.8%.

The better gain in BMS is likely due to the interdependency with ALF (similar as reported in CE2.4.1.4i). However in case of deblocking filter, design aspects should be studied based on subjective impact rather than visual gain.

Unclear if the normative modifications would be really necessary, and what the contribution of each of the three design aspects is. The crosschecker also reports that the condition on chroma edges would introduce a dependency between luma and chroma which would be undesirable.

It is also noted that for the aspect of clipping testing with class F (sharp edges) would be important

Further study in CE.

[JVET-K0524](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4054) Crosscheck for CE2-related: Higher-precision modifications to VVC deblocking filters (JVET-K0203) [B. Wang, A.M. Kotra (Huawei)]

[JVET-K0237](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3747) CE2-related: Bugfix for deblocking at maximum transform block boundaries [C.-M. Tsai, C.-W. Hsu, Y.-W. Huang, S.-M. Lei (MediaTek)]

In this contribution, a bugfix is proposed for deblocking at transform block (TB) boundaries. In VTM-1.0 and BMS-1.0, each coding blocks (CB) larger than maximum TB is inferred to be further partitioned into multiple TBs by tiling with the maximum TBs, while each CB smaller than or equal to maximum TB has only one TB. However, deblocking is only performed at CB boundaries and always skipped at maximum TB boundaries that do not coincide with any CB boundaries. It is proposed to apply deblocking to maximum TB boundaries that do not coincide with any CB boundaries.

This bugfix was adopted in VTM (see notes under CE2)

[**JVET-K0503**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4033) **Crosscheck for CE2-related: Bugfix for deblocking at maximum transform block boundaries (JVET-K0237) [K. Andersson, Z. Zhang (Ericsson)] [late]**

[JVET-K0238](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3748) CE2-related: Improvements of sample adaptive offset [C.-Y. Lai, C.-Y. Chen, C.-W. Hsu, Y.-W. Huang, S.-M. Lei (MediaTek)]

(include abstract from new version)

Two aspects:

* Grouping of CTUs (one row) for optimization of SAO parameters (encoder only) – gives similar gain as K0201 for the cases of AI and RA, but does not require normative change
* Modification of syntax – does not provide any benefit

Decision(SW): Adopt the non-normative encoder trick (not CTC)

[JVET-K0465](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3994) Crosscheck of JVET-K0238: CE2-related: Improvements of sample adaptive offset [C.-H. Yao, P.-H. Lin, C.-C. Lin, S.-P. Wang, C.-L. Lin (ITRI)] [late]

[JVET-K0489](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4019) Cross-check of JVET-K0238: CE2-related: Improvements of sample adaptive offset [T. Ikai (Sharp)] [late]

[JVET-K0239](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3749) CE2-related: Filter size reduction in CTB-based ALF [Y.-C. Su, C.-Y. Chen, Y.-W. Huang, S.-M. Lei (MediaTek)]

Based on CE2.4.2.2, a 9x7cross+3x3square filter shape is proposed for CTB-based adaptive loop filter (ALF) for reducing two line buffers in comparison with 9x9cross+3x3square. Compared with VTM-1.0, the 9x7cross+3x3square CTB-based ALF can achieve -2.63%, -4.78%, and -4.45% luma BD-rates with 30%, 39%, and 30% decoding time increases, for AI, RA, and LB, respectively. Compared with BMS-1.0 with ALF disabled, the 9x7cross+3x3square CTB-based ALF can achieve -2.12%, -4.52%, and -4.16% luma BD-rates with 23%, 25%, and 23% decoding time increases for AI, RA, and LB, respectively. Compared with 9x9cross+3x3square CTB-based ALF under VTM-1.0 configuration, 9x7cross+3x3square CTB-based ALF achieves 0.06%, 0.04%, and 0.08% luma BD-rates with 2%, 4%, and 4% decoding time decreases for AI, RA, and LB, respectively. Compared with 9x9cross+3x3square CTB-based ALF under BMS-1.0 configuration, 9x7cross+3x3square CTB-based ALF achieves 0.05%, 0.05%, and 0.08% luma BD-rates with 0%, 1%, and 2% decoding time decreases for AI, RA, and LB, respectively.

Was reviewed in BoG JVET-K0521

[JVET-K0467](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3996) Cross-check of JVET-K0239: Filter size reduction in CTB-based ALF [Q. Yu, Y. Lin (HiSilicon)] [late] [miss]

[**JVET-K0274**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3786) **CE2 related: Reduced complexity bilateral filter [J. Ström, P. Wennersten, J. Enhorn, D. Liu, K. Andersson, R. Sjöberg (Ericsson)]**

This contribution proposes a modified version of the bilateral filter from JVET-F0034, JVET-F0096 and JVET-J0021. The main modification is a reduction in the size of the look-up table (LUT) that is used to store the filter coefficients. The contribution claims to reduce the total size of the stored variables (including the LUT) from 2783 bytes in JVET-F0096 to 816 bytes, a reduction of 71%. The proposal states that this is achieved by approximating the 34 rows in the LUT (one row is used for every qp) by four rows plus shifting. The contribution further claims that the need for a division table is removed by using the approximation proposed in JVET-J0021. The non-local filtering for inter blocks proposed in JVET-J0021 is reportedly also used. The BD rate figures for an implementation in BMS 1.0 are reported to be -0.33% / -0.52% / -0.60% for AI/RA/LD respectively, and the VTM figures are reported to be ‑0.33% / -0.81% / -0.59% for AI/RA/LD respectively. The BMS decoder run times are reported to be 101% / 101% / 101% for AI/RA/LD and the VTM decoder run times are reported to be 104% / 103% / 103%.

The reduction in run time is due to re-using difference computations, at the same time increasing the difference computation window which gives small compression gain.

Additional results are also presented that demonstrate almost identical results (small loss for AI, small gain for LD) when the bilateral filter is not used for 4x4 blocks. Gain becomes larger when also 4x8 and 8x4 block are disabled.

Interesting LUT reduction and computation reduction. However, the problem remains that bilateral filter is at a critical path between inverse transform and intra prediction, which might introduce latency in pipelining

It is pointed out that in software implementation the LUT operations cannot be performed in parallel.

Further study (CE) of the aspect of block size restrictions, in terms of performance and whether this resolves the latency issue (e.g. when boundary samples needed for next prediction are filtered first after the inverse transform). How many additional cycles are needed between inverse transform and before the prediction can be started? An initial analysis was shown in track B Monday afternoon, where it was shown that the processed edge samples could be available within 10 cycles after the inverse transform is done. Further consideration necessary if that would be acceptable implementation wise. Further results are shown that the loss by further reducing the number of LUTs to 16 is marginal.

This additional should be provided in an update of the document.

[JVET-K0563](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4093) Cross-check of contribution JVET-K0274 [J. Rasch (HHI)] [late]

[JVET-K0318](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3832) CE2-2.1.1-related: HEVC luma filters and decisions for chroma deblocking [K. Andersson, Z. Zhang, R. Sjöberg (Ericsson)]

This contribution proposes to use HEVC luma filters and decisions for chroma deblocking with some minor adaptations in decisions. For BMS it also applies deblocking of implicit TU boundaries after deblocking of sub-block boundaries from motion prediction to ensure that implicit TU boundaries can be deblocked with more than 1 pixel even when it exist sub-block boundaries 4 samples from the implicit TU boundary. The modifications are implemented on top of CE2-2.1.1.

The proposed solution is claimed to improve subjective quality especially notable for Campfire at low bitrates. It also provides a luma,Cb,Cr BD rate impact of -0.19%,-0.18%,0.17% / -0.18%,-2.20%,-2.22% / -0.12%,-1.89%,-1.69% for AI/RA/LD compared to VTM-1.0 and -0.12%,-1.08%,-0.93% / -0.12%,-2.08%,-2.28% / -0.20%,-2.46%,-2.40% for AI/RA/LD compared to BMS-1.0. Decoding time 105%/106%/107% for AI/RA/LD compared to VTM-1.0 and 103%/104%/103% for AI/RA/LD compared to BMS-1.0.

These are additional changes beyond the bug fix of enabling DBF on large TU boundaries.

It is suggested to use the luma type of decisions for chroma as well.

Samples close to large TU boundaries might be sampled twice (once over subblock and again with long filter at large TU). This would require two passes of DBF, and somewhat inhibit parallelism.

Further study in CE.

[JVET-K0494](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4024) Crosscheck of JVET-K0318 (CE2-2.1.1-related: HEVC luma filters and decisions for chroma deblocking) [C.-M. Tsai (MediaTek)] [late]

[JVET-K0369](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3887) CE2-related: Longer Tap Deblocking Filter [A.M. Kotra, B. Wang, S. Esenlik, Z. Zhao, J. Chen (Huawei)]

A new longer tap deblocking filter for the luma component is proposed. The proposed filter mainly targets the filtering of blocking artifacts which arise due to the usage of larger size transform units and coding units. The “longer tap” filter introduces new filter condition checks which consider a wider range of spatial activity along the edges. Furthermore, new longer tap filter coefficients are proposed in order to effectively smooththeedges with blocking artifacts belonging to larger blocks. Our proposal also filters the implicit TU boundaries and allows for parallel deblocking of different CUs.

Moreover, to reduce the line buffer requirements for the “longer tap” filter: For the horizontal edges which overlap with the CTU boundaries, the maximum number of samples used in filter decision and the maximum number of samples used in filter modification from the top block are restricted to be the same as in HEVC deblocking filter. Compared to the deblocking filter used in BMS, the proposed method improves the subjective quality of sequences, especially for the ones which are encoded at lower bitrates. Objective results of the proposed longer tap deblocking filer are as follows:

Over VTM Anchor (AI, RA, LDB): Luma BD-Rate gain of -0.12%, -0.13%, -0.02% is achieved without any increase in EncT and DecT.

Over BMS Anchor (AI, RA, LDB): Luma BD-Rate gain of -0.08%, -0.01%, 0.04% is achieved without any increase in EncT and DecT.

Unlike some other proposals of CE2.2, only one condition is checked to decide for the longer filter. Further, it is claimed that line buffer requirements are reduced. Up to 7 samples are filtered at each side.

Include this in the next CE.

The CE shall also report about complexity of the different proposals such as additional line buffer requirements, number of operations due to additional rules, number of operations (worst case) for the filtering, etc.

[JVET-K0492](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4022) Cross-check of JVET-K0369: CE2-related: Longer Tap Deblocking Filter [C. Gisquet, J. Taquet] [late]

[JVET-K0372](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3890) CE2-related: Additional results for CE2.4.1.4 with chroma filter shape aligned with luma [N. Hu, V. Seregin, N. Shlyakhov, M. Karczewicz (Qualcomm)]

This contribution presents additional test results for CE2 test 4.1.4 where ALF filter shape for chroma is aligned with luma component. In the BMS ALF, luma filter can be switched between 5x5, 7x7, or 9x9 filter shapes, while chroma filter size is always 5x5. The same flexible filter structure for chroma is tested in this contribution. For maximum ALF filter shape size 9x9, test results reportedly show 3.26%, 5.34%, and 4.63% luma gain in AI, RA, and LB configurations respectively over VTM-1.0 anchor.

Slide deck is available in JVET-K0371.

Somewhat obsolete after the decision on inclusion of another ALF version in VTM.

[JVET-K0479](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4009) Cross-check of JVET-K0372: CE2-related: Additional results for CE2.4.1.4 with chroma filter shape aligned with luma [R. Vanam (InterDigital)] [late]

[JVET-K0373](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3891) CE2-related: Two-dimensional ALF classification [M. Karczewicz, N. Hu, V. Seregin (Qualcomm)]

This contribution proposes a modification to BMS ALF classification. In modified classification, two characteristics: Laplacian based activity and direction are used to form a joint classification. The categorization for each characteristic is signaled to the decoder side, and the joint classification is used instead of the ALF classification in BMS-1.0. Test results reportedly show 3.38%, 5.48%, and 4.82% luma gain in AI, RA, and LB configurations respectively over VTM-1.0 anchor.

Slide deck to be uploaded.

Current classification scheme combines activity and direction. Here, classification for act. and dir. is performed independently, and then certain combinations are mapped with filters. This provides 0.1% bit rate gain, but decoder runtime is highly increased. Not clear if it is conceptually more complex.

Somewhat obsolete after the decision on inclusion of another ALF version in VTM.

[JVET-K0534](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4064) Crosscheck of JVET-K0373: CE2-related: Two-dimensional ALF classification [M. Ikeda (Sony)] [late]

[JVET-K0382](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3901) CE2-related: CTU Based Adaptive Loop Filtering [M. Karczewicz, A. Gadde, N. Hu, V. Seregin (Qualcomm)]

In this contribution, additional mode for the adaptive loop filter in BMS is proposed. In this mode, selection of the set of filters is done for each CTU. Test results reportedly show 3.10%, 4.96%, 4.31% luma gain for AI, RA, and LB configurations respectively comparing to VTM-1.0 anchor. Additional results with low-delay ALF encoder, where the filters used for the current picture are derived from the previous coded pictures, show 3.05%, 4.72%, 4.16% luma gain for AI, RA, and LB configurations comparing to VTM-1.0 anchor.

Beneficial for low latency encoding. Further study in CE, compared to an equivalent low latency mode with the current ALF (determining coefficients from previous picture, and switching filter on/off at CTU level.)

[JVET-K0488](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4018) Cross-check of JVET-K0382: CE2-related: CTU Based Adaptive Loop Filtering [[T. Ikai (Sharp)](mailto:ikai.tomohiro@sharp.co.jp)] [late]

[JVET-K0388](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3907) CE2-related: Improvement on the implementation of adaptive loop filter [Y. Li, Z. Chen (Wuhan University), X. Li, S. Liu (Tencent)] [late]

This contribution presents some implementation problems about adaptive loop filter (ALF) in the current BMS1.1 software. To solve the problems, a numerical method based on LDLT decomposition is suggested, which can be used to replace the current Cholesky factorization method.

The case could happen when the covariance matrix is not positive definite, or if it does not have full rank, was detected e.g. for uniform pictures.There was a ticket #59 reported, and a patch to resolve this already existing. Therefore the problem is resolved for now, it anyway does not happen under CTC.

The proponents are encouraged to study if the problem of ill-conditioned covariance might also occur with different picture content, e.g. some type of screen content. It might also be the case that it would be better to completely disable ALF if it occurs.

[JVET-K0502](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4032) Crosscheck report of JVET-K0388 (Improvement on the implementation of adaptive loop filter) [Y. Zhao (Huawei)] [late]

[**JVET-K0540**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4070) **CE2-related: Reduced filter shape size for ALF without classification [V. Seregin, N. Hu, M. Karczewicz (Qualcomm)] [late]**

This contribution presents results of CE2 test 4.1.4 where the largest ALF filter shape is reduced from 9x9 to 7x7 and classification process is disabled. Test results provide 1.58%, 3.08% and 2.79% luma gain in AI, RA, and LB configurations respectively over VTM-1.0 anchor for 7x7 max filter size, and provide 1.78%, 3.30% and 3.10% luma gain in AI, RA, and LB configurations respectively over VTM-1.0 anchor for 9x9 max filter size.

The approach investigated here is using one filter for the whole picture, but it can locally be disabled at block level.

Just for information – no action.

## CE3 related – Intra prediction and mode coding (26)

Contributions in this category were discussed XXday XX July XXXX–XXXX (chaired by XXX).

[JVET-K0539](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4069) BoG report on intra prediction and mode coding (CE3-related) [G. Van der Auwera]

This BoG report was discussed Saturday 1815 (GJS).

The BoG reviewed related input contributions to Core Experiment 3 on intra prediction and mode coding, and formulated recommendations for consideration by the track. The meeting took place on Saturday July 14th from 9:00am until 1:20pm.

Regarding the division by the width + height used for the DC mode, among the proposed schemes, the second method in K0122 seemed preferable. Decision: Adopt the second method in K0122 (i.e., use only the longer side to compute the average for non-square blocks).

The BoG recommended testing K0500 in a CE. During the discussion of the BoG report, a proponent suggested adoption rather than a CE. The proposal is for “wide-angle intra prediction”, and substitutes the wide angles for other angles when the block shape is rectangular. The proposal is reported to add 6 or 10 modes to particular blocks. The total number of directional modes in the design is increased by 20.

The reported gain for wide-angle intra prediction is 0.3% for AI and 0.2% for RA. It was reported (in a revision of K0500) that there is no effect on encoder and decoder runtime. It was noted that the reported testing had been done with the 35 mode intra scheme, not the 67 mode intra scheme. The tables in the document for comparison to the BMS were empty at the time of this discussion.

The BoG recommended having CE of CCLM improvements and multi-reference-line prediction. CE tests of intra mode coding may also be desirable to consider.

Not presented in BoG: K0058 (presenter not available), K0518 (later reviewed Monday pm), K0529 (later reviewed Saturday pm), K0536 (later reviewed Tuesday am), and later K0542.

Presentation requested to be deferred: K0276, K0348 (and related K0377).

In further discussion Saturday 14 July at 2020 (GJS) a participant asked about having a CE on line-based intra coding. It was agreed that a CE study of this would be desirable, and should take into account the complexity concerns expressed about small transform units (e.g., using a minimum of 16 sample units rather than the 4-sample minimum unit size used in the prior CU).

Wide-angle filtering was further discussed Monday 16 July 1520 (chaired by GJS) to report the results of the ongoing experiment. On top of BMS (i.e., on top of the 67-mode scheme), the wide-angle filtering was reported to provide 0.3% for AI, 0.1% for RA, and another participant said 0.1% for LB.

Boundary smoothing, which had been used in some previous experiments, was not included in the test and was not being proposed.

A participant commented that the proposed wide angle scheme actually solves an inconsistency in the reference samples needed. The length of the reference sample array becomes twice the length of the longer edge instead of the width plus the height, which is said to be more consistent with what happens for square blocks.

Decision: Include the wide angles in the expansion of the number of angles (85 angles, 87 modes total).

[JVET-K0058](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3559) CE3-Related: Priority List Based Intra Mode Coding with 5 MPM [Y. Yu, K. Panusopone, S. Hong, L. Wang (Arris)]

Presenter not available.

[JVET-K0059](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3560) CE3-related: Reference sample processing for wide-angle intra-prediction [A. Filippov, V. Rufitskiy, J. Chen (Huawei)]

[JVET-K0064](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3565) CE3-related: On MDIS and intra interpolation filter switching [G. Van der Auwera, A.K. Ramasubramonian, V. Seregin, T. Hsieh, M. Karczewicz (Qualcomm)]

[JVET-K0498](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4028) Crosscheck for CE3-related: On MDIS and intra interpolation filter switching (JVET-K0064) [K. Sharman (Sony)] [late]

[JVET-K0074](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3576) CE3-related: Combination of LAM (Test 4.1.1) and MNLM (Tests 4.1.3, 4.1.4, and 4.1.5) [Y.-J. Chang, H.-Y. Jiang (Foxconn)]

[JVET-K0533](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4063) CE3-related: Cross-check of JVET-K0074 [S. Paluri (LGE)] [late]

[JVET-K0122](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3624) CE3-related: Alternative techniques for DC mode without division [A. Filippov, V. Rufitskiy, J. Chen (Huawei)]

[JVET-K0483](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4013) Cross-check of JVET-K0122: CE3-related: Alternative techniques for DC mode without division [X. Qi, L. Wang (Hikvision)] [late]

[JVET-K0169](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3677) CE3-related: Block Shape Adaptive Intra Prediction Directions [G. Rath, F. Urban, F. Racapé (Technicolor)]

[JVET-K0442](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3967) Crosscheck of JVET-K0169: CE3-related: Block Shape Adaptive Intra Prediction Directions [[L. Zhao (Tencent)](mailto:leolzhao@tencent.com)] [late]

[JVET-K0172](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3680) CE3-related: Gradient-Based Boundary Filtering in Intra Prediction [G. Rath, F. Urban, F. Racapé (Technicolor)]

[JVET-K0458](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3984) Cross-check of JVET-K0172: CE3-related: Gradient-Based Boundary Filtering in Intra Prediction [Y. He (InterDigital)] [late]

[JVET-K0175](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3683) CE3-related: Advanced MPM based on intra reference line selection scheme [H.-Y. Jiang, Y.-J. Chang (Foxconn)]

[JVET-K0441](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3966) Crosscheck of JVET-K0175: CE3-related: Advanced MPM based on intra reference line selection scheme [L. Zhao (Tencent)] [late]

[JVET-K0196](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3705) CE3-related: Non-linear weighted intra prediction (cross-check report in JVET-K0262) [P. Helle, T. Hinz, R. Rischke, J. Pfaff, P. Merkle, M. Schäfer, B. Stallenberger, V. George, H. Schwarz, D. Marpe, T. Wiegand (Fraunhofer HHI)]

[JVET-K0262](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3772) Cross-check report on JVET-K0196 CE3-related: Non-linear weighted intra prediction [S. Ikonin, J. Chen (Huawei)] [late]

[JVET-K0204](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3713) Non-CE3: On cross-component linear model simplification [G. Laroche, J. Taquet, C. Gisquet, P. Onno (Canon)]

[JVET-K0512](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4042) Cross-check of JVET-K0204: On cross-component linear model simplification [A. K. Ramasubramonian (Qualcomm)] [late]

[JVET-K0221](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3731) CE3 Related: Additional results of JVET-J1023 Core Experiments 5.2.3, 5.2.4 and 5.2.5 [S. Keating, K. Sharman (Sony)]

[JVET-K0242](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3752) CE3-related: Intra planar mode prediction [M. G. Sarwer, C.-W. Hsu, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-K0477](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4007) Cross-check of JVET-K0242: CE3-related: Intra planar mode prediction [Y. Kidani, K. Kawamura, S. Naito (KDDI)] [late]

[JVET-K0243](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3753) CE3-related: Intra mode coding [M. G. Sarwer, C.-W. Hsu, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-K0525](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4055) Crosscheck for CE3-related: Intra mode coding (JVET-K0243) [B. Wang, A.M. Kotra (Huawei)]

[JVET-K0276](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3788) CE3-related: Mode-dependent multiple reference line intra prediction with bi-prediction [B. Bross, P. Keydel, H. Schwarz, D. Marpe, T. Wiegand (HHI)]

This document proposes to add bi-prediction to mode-dependent multiple-reference line intra prediction as tested in sub-CE 3.5, test 5.4.4 and reported in JVET-K0051. An angular prediction that uses extend reference lines can be averaged with PLANAR prediction using the nearest reference lines. When the reference line index indicates that extended reference samples are used in angular intra prediction, an additional flag indicates whether bi-prediction with PLANAR mode applies or not. The results show that MRL bi-prediction achieves -1.17% (AI) and -0.54% (RA) BD-rate Y, 181% (AI) and 114% (RA) enc. time as well as 99% (AI) and 100% (RA) decoding time for the VTM configuration. It is proposed to study MRL bi-prediction in a new CE on MRL.

Presented Track B Tuesday 17th 1445

New element: planar prediction is averaged with the result from angular prediction of different reference lines, and compute average as final prediction.

Gain is on top of K0051

High increase of encoder run time in AI. (180%), which includes the additional effort of extended reference lines plus the additional bi-pred mode. The additional is approx is approx. 0.4% for AI, the relative run time increase is 0.35%

Would there also be gain without additional reference lines?

Test in CE3.

[**JVET-K0544**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4074) **Cross-check of JVET-K0276: CE3-related: Mode-dependent multiple reference line intra prediction with bi-prediction [M. Sychev (Huawei)] [late]**

[JVET-K0289](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3802) CE3-related: Wide angular intra prediction for non-square blocks [L. Zhao, S. Liu, X. Zhao, X. Li (Tencent)]

[JVET-K0515](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4045) Cross-check of JVET-K0289: CE3-related: Wide angular intra prediction for non-square blocks [G. Rath, F. Racape, F. Urban (Technicolor)] [late]

[JVET-K0293](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3806) CE3-related: Simplifications for chroma intra coding [L. Zhao, X. Zhao, X. Li, S. Liu (Tencent)]

[JVET-K0336](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3851) CE3-related: Improved multi-directional LM [X. Ma, H. Yang, J. Chen (Huawei)]

[**JVET-K0478**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4008) **Cross-check of JVET-K0336: CE3-related: Improved multi-directional LM [Y. Kidani, K. Kawamura, S. Naito (KDDI)] [late]**

[JVET-K0348](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3865) CE3-related: On the combination of multiple reference lines, bilateral reference line filtering, PDPC and 65 directional modes for intra prediction [G. Van der Auwera, A.K. Ramasubramonian, V. Seregin, M. Karczewicz (Qualcomm), B. Bross, P. Merkle, H. Schwarz, D. Marpe, T. Wiegand (HHI)]

This document proposes combining tools described in CE3-5.4.4 (multiple reference lines for intra prediction), CE3-2.7.1 (bilateral reference sample filter), CE3-2.4.1 (simplified PDPC), and 65 intra directional modes. It is asserted that the gains of the tools are additive with a BD-rate improvement of −3.16% (Y), −1.77% (U), −1.73% (V) with runtimes of 188% (Enc) and 107% (Dec) for AI configuration over BMS-VTM-1.0, and −2.03% (Y), −1.31% (U), −1.17% (V) with runtimes of 120% (Enc) and 100% (Dec) for AI configuration over BMS-1.0. For the random-access configuration, the BD-rate improvement is −1.57% (Y), −0.27% (U), −0.35% (V) with runtimes of 116% (Enc) and 100% (Dec) over BMS-VTM-1.0 and −1.1% (Y), −0.6% (U), − 0.6 % (V) with runtimes of 112% (Enc) and 101 % (Dec) over BMS-1.0. The source code of this tool combination is provided together with this proposal. In addition, informative results for combining the intra tools with more intra tools and proposed tools from transform and quantization showing that BD-rate savings resulting from these combinations are additive as well.

Presented track B Tuesday 17th 1500

Analysis of combination of tools.

Gains of intra tools add up

Gains of intra tools plus AMT are even more than additive.

[JVET-K0504](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4034) Cross-check of JVET-K0348 [W. Zhu, A. Segall(Sharp)] [late]

[JVET-K0377](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3896) CE3-related: Combining CE3-5.2.5 on using two rows of reference lines for prediction, CE3-2.4.1 on PDPC and 65 angular intra modes [S. Keating, K. Sharman (Sony), G. Van Der Auwera, A. Ramasubramonian, V. Seregin, A. Said, M. Karczewicz (Qualcomm)]

This document proposes combining tools CE3-2.4.1 (simplified PDPC), CE3-5.2.5 (using two rows of ref. samples for intra prediction) and 65 intra angular modes. The results are reported as being additive with a BD-rate improvement of -2.59% (Y), -1.75% (U), -1.62% (V) for AI over VTM1.0 with run-times of 134% (Enc) and 114% (Dec).

An additional proposal is to also include the merging of the intra reference sample smoothing filter with the intra interpolation filter as proposed in JVET-K0064. The results for this are reported to give a BD-rate improvement of -3.35% (Y), -2.03% (U), -1.92% (V) for AI with run-times of 149% (Enc) and 121% (Dec).

No need for presentation. Similar report of additive gain of intra tools as of K0348.

[JVET-K0400](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3920) CE3-related: DC mode with only shift operators based on sub-sampling [D. Kim, G. Ko, J. Son, J. Kwak (WILUS), J. Seok, Y. Lee (Humax)] [late]

[JVET-K0506](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4036) Cross-check of JVET-K0400: CE3-related: DC mode with only shift operators based on sub-sampling [B. Lee (Chosun Univ.)] [late]

[JVET-K0431](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3954) Cross-check of JVET-K0293: CE3-related: Simplifications for chroma intra coding [T.-H. Li, H.-Y. Jiang, Y.-J. Chang (Foxconn)] [late]

[JVET-K0469](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3999) CE3-related: Chroma intra prediction simplification [C.-H. Yao, P.-H. Lin, C.-C. Lin, S.-P. Wang, C.-L Lin (ITRI)] [late]

[JVET-K0561](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4091) Cross check of JVET-K0469: CE3-related: Chroma intra prediction simplification [X. Zhao (Tencent)] [late] [miss]

[JVET-K0482](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4012) CE3-related: MPM based multi-line intra prediction scheme [L. Zhao, X. Zhao, X. Li, S. Liu (Tencent), H.-Y. Jiang, Y.-J. Chang (Foxconn)] [late]

[JVET-K0500](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4030) CE3-related: Wide-angle intra prediction for non-square blocks [F. Racape, G. Rath, F. Urban (Technicolor), L. Zhao, S. Liu, X. Zhao, X. Li (Tencent), A. Filippov, V. Rufitskiy, J. Chen (Huawei)] [late]

[JVET-K0517](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4047) Cross-check of JVET-K0500 (CE3-related: Wide-angle intra prediction for non-square blocks) [J. Lainema (Nokia)] [late]

[JVET-K0518](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4048) CE3-related: Interpolation filtering for intra-prediction within rectangular blocks [A. Filippov, V. Rufitskiy, J. Chen (Huawei)] [late]

Discussed Mon 1600 (GJS).

In CE3, different intra-interpolation filters were considered to improve the coding efficiency of intra prediction for skew directional modes. This contribution proposes an alternative mechanism of selecting interpolation filters. On top of the VTM, the simulation results for this technique reportedly show 0.60%for the AI configuration with 105% encoding time and 100% decoding time on average.

This scheme would add several filters. It is a variation of a scheme tested in CE3. Further study was encouraged.

[JVET-K0529](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4059) CE3.3 related: Intra 67 modes coding with 3MPM [N. Choi, Y. Piao, K. Choi, C. Kim (Samsung)] [late]

This contribution was discussed Sat 1930 (chaired by GJS).

This contribution presents the performance and text of 67 intra mode coding when using 3MPM and a 6-bit fixed-length coding (FLC) for remaining modes which is similar to HEVC intra mode coding. 67 modes with 3MPM on BMS reportedly shows 0.6% loss in AI and 0.3% loss in RA. The AI encoding time was increased by 15%. (See below as K0545 regarding the encoding time.) Some of the tests had not been completed.

The 6 MPM scheme used for comparison does more RDO mode checking and has higher encoding complexity than the 3 MPM scheme as tested.

Additional testing was being done to check the performance of the 3 MPM scheme when the encoder search is checking more to be comparable to the encoder search used for the 6 MPM scheme.

This was further discussed Monday 1605 (GJS).

The tests described above had been completed. The estimated impact described above was confirmed. With an addition RD check in the encoder to equalize the search complexity, the coding loss was 0.5% in AI and 0.2% in RA relative to the 6 MPM scheme with truncated binarization. Decision: As recorded in section 12.2, the 3 MPM scheme will be used in VTM & draft 2.

[**JVET-K0545**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4075) **Cross check of JVET-K0529, Intra mode coding with 3 MPMs and results for 6 MPM with optimized encoder [A.M. Kotra, B. Wang, J. Chen (Huawei)] [late]**

Discussed Sat 2005 (GJS)

Results were presented for 6 MPM with truncated binarization. The gain reported was 1.3% for AI and 0.6% for RA. The AI encoding time was increased by 8%. There was some difference in the encoding optimization used in this comparison relative to what was used in the prior CE and in K0529, which accounts for the higher encoding time impact measured in that contribution.

[**JVET-K0536**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4066) **Non-CE3: Adaptive multiple cross-component linear model [S.-P. Wang, P.-H. Lin, C.-H. Yau, C.-L. Lin, C.-C. Lin (ITRI)] [late]**

In this contribution, an adaptive grouping mechanism is utilized to classify neighbouring samples into groups to derive a linear model for each group. Then, chroma samples are predicted based on the reconstructed luma samples by using a linear model in the nearest group. Test results reportedly show 1.54%, 9.65% and 10.06% BD-rate saving on luma and chroma, respectively, under AI test condition over VTM1.0. The contribution does not compare the method to the BMS CCLM method.

The decoder is proposed to perform a classification based on the decoded sample values and apply CCLM separately to the regions.

Another CCLM modification is MMLM which is being tested in a CE.

More reference lines are used. To be further studied in a CE.

[**JVET-K0542**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4072) **CE3-related: Combination of CE3-5.5.1, 5.4.4 and 5.5.2 [P.-H. Lin, C.-C Lin (ITRI), L. Zhao, X. Zhao, X. Li, S. Liu (Tencent), B. Bross, H. Schwarz, D. Marpe, T. Wiegand (HHI)] [late]**

This contribution lays out a combination of the methods of 3 CE tests, test 5.1.1, test 5.4.4 and test 5.5.2. The results show that 0.53% gain with 24% encoder run-time increase and 0.31% gain with 5% encoder run-time increase for AI/RA and VTM configuration respectively are observed. The results show that the combination yield a reduction of the relative encoding run-time by up to 80% to compared to the individual CE3 tests. It is proposed to further study this combination in a CE.

Combination of only using subset of directional modes for ref lines 1 and 3; restricting narrow blocks; fast encoder search.

The original proposals had gains around 1%, however, the run time was unacceptable high.

Further study in CE.

[**JVET-K0555**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4085) **Cross-check of JVET-K0059: CE3-related: Reference sample processing for wide-angle intra-prediction [Jonathan Pfaff (HHI)] [late]**

## CE4 related – Inter prediction and motion vector coding (41)

Contributions in this category were discussed in BoG K0546 (chaired by H. Yang), unless noted otherwise.

[JVET-K0052](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3553) Non-CE4: A study on the affine merge mode [M. Zhou (Broadcom)]

This contribution studied line buffer usage of the affine merge mode in BMS1.0 and advocated the following three changes to simplify the design, namely, 1) directly using 4-parameter affine motion model to derive the seed vectors for the current PU of the affine merge mode; 2) disabling the affine merge mode for PUs whose width is less than 8 to enable motion data line buffer sharing, and 3) having the affine (merge) mode and the regular merge/skip and AMVP mode share the same motion data line buffer. For 4K video, the proposed changes reduce the line buffer size of the affine merge mode from roughly 18,688 bytes to 320 bytes, without comprising compression efficiency. Relative to the BMS1.0 anchor, the overall BD-rate changes are 0.03% in RA, -0.07% in LD-B and -0.03% in LD-P, respectively. Compared to the BMS1.0 VTM configurations but with “Affine” and “HighPrecMv” on, the overall BD-rate changes are 0.00% in RA, 0.02% in LD-B and 0.06% in LD-P, respectively. It is recommended to study the simplification of the affine (merge) mode to make it more implementation friendly.

This contribution proposes using 4-parameter affine motion model to derive the seed vectors for the current PU of the affine merge mode. BMS affine firstly derive the bottom-left MV and then use the pseudo 6-param model to derive two CPMVs of the current block, which is a redundant operation.

Recommendation: when inheriting the 4-param affine model from neighboring blocks, remove redundant operation of deriving the bottom-left CPMV of the neighboring block. Use 4-parameter model to compute seed vectors for the current CU. Adopt this in BMS affine.

It is proposed to reuse the motion data line buffer as regular inter mode. It is noted that the potential increase of line buffer size for 4K is about 18 KB if we store affine CPMVs separately.

Recommendation: further study in CE.

It is shown in this contribution there is no performance drop when storing 1/4 precision MV instead of 1/16 precision MV in line buffer. Suggest further study.

It is commented that storing high-precision MV in motion buffer causes higher memory bandwidth access and larger line buffer size as well.

[JVET-K0508](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4038) Crosscheck of JVET-K0052 (Non-CE4: A study on the affine merge mode) [H. Chen, H. Yang, J. Chen (Huawei)] [late]

[JVET-K0056](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3557) Non-CE4: Merge mode modification [T. Solovyev, J. Chen, S. Ikonin (Huawei)] [late]

This contribution describes the enhanced merge mode based on VTM 1.0 software. The proposed method extends the merge candidates list by spatial and temporal candidates. Additional spatial candidates are located in current CTU line and bottom line of above CTU. The total amount of checking positions is limited by 22 and final number of candidates is limited to 10.

It is claimed that additional two TMVP candidates provide about 0.2% coding gain.

The order of candidates, regular spatial and temporal candidates, additionally 2 temporal candidates, at most three long distance spatial candidates.

Recommendation: study this in CE.

[JVET-K0549](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4079) Crosscheck of JVET-K0056: Non-CE4: Merge mode modification [J. Ye, X. Li (Tencent)] [late] [miss]

[JVET-K0065](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3566) CE4 related: Candidate list reordering [L. Xu, F. Chen, S. Ye, L. Wang (Hikvision)]

This contribution proposes a candidate list reorder method for Merge and AMVP mode. Instead of using fixed candidate list construction order, the proposed method adopts different approaches of candidate list construction order for blocks with different shapes. Compared with BMS1.0, the results reportedly show that the proposed swapping method achieves 0.35%, 0.32% and 0.16% BD rate reduction for, LB, LP, and RA configurations, respectively. Compared with VTM1.0, it also reportedly shows that the proposed method achieves 0.32%, 0.32% and 0.13 % BD rate reduction for, LB, LP, and RA configurations respectively, while brings negligible increase of encoding and decoding complexity.

When constructing the list either for merge mode or for AMVP, the checking order of neighboring blocks is dependent on CU shape.

Recommendation: study this in CE.

[JVET-K0417](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3938) Crosscheck of JVET-K0065: CE4 related: Candidate list reordering [J. Chen, K. Choi (Samsung)] [late]

[JVET-K0080](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3583) CE4-2.6 related: Simplified ATMVP with fixed sub-block size [H. Jang, J. Lim, J. Nam, S. Kim (LGE)]

Contribution discussed together with JVET-K0346.

[JVET-K0412](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3933) Crosscheck of JVET-K0080: Simplified ATMVP with fixed sub-block size [Y. Han, W.-J. Chien (Qualcomm)] [late]

[JVET-K0095](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3598) CE4-related: Harmonization of CE4.1.7 and CE4.1.3 [J. Lee, J. Nam, N. Park, H. Jang, J. Lim, S. Kim (LGE)]

It is about 1) slice level switching of 4/6-param affine model, 2) conditional signaling of 6-param model at block level, and 3) by-pass coding of the flag for 4/6-param model switching at block level.

Recommendation: further study in CE.

[JVET-K0413](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3934) Crosscheck of JVET-K0095: Harmonization of CE4.1.7 and CE4.1.3 [Y. Han, Y. Zhang, W.-J. Chien (Qualcomm)] [late]

[JVET-K0101](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3604) CE4-related: Affine MVD Coding [S. Paluri, M. Salehifar, S. Kim (LGE)]

In this proposal, a new motion vector difference coding (MVD) for Affine Motion Vector Difference (AMVD) is introduced. In affine mode, two control motion vectors are used to derive affine motion vectors for each sub-block. One is motion vector from left top and the other is motion vector from right top. Hence, two MVDs are coded. The proposed method jointly exploits the similarity of the MVD between the MVD from left and MVD from right. It is reported that 0.09% BD-rate saving is observed for Random Access configuration.

It is claimed that the joint coding of component from 2 CPMVs could also be applied to the case of 3 CPMVs.

It is claimed that this method could be combined with the MVD coding in JVET-K0337.

It is commented that affine MVD coding requires coding 2/3 MVs instead of 1 MV, which may need a coding method with higher efficiency.

Further study is encouraged.

[JVET-K0432](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3955) Cross-check of JVET-K0101: CE4-related: Affine MVD Coding [T.-H. Li, H.-J. Jhu, Y.-J. Chang (Foxconn)] [late]

[JVET-K0102](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3605) CE4-related: Interweaved Prediction for Affine Motion Compensation [K. Zhang, L. Zhang, H. Liu, Y. Wang, P. Zhao, D. Hong (Bytedance)]

With the affine motion compensation (AMC) in the benchmark set (BMS), a coding-block is divided into sub-blocks as small as 4×4, each of which is assigned with an individual motion vector derived by the affine model. In this contribution, an interweaved prediction approach is proposed for AMC. With the interweaved prediction, a coding block is divided into sub-blocks with two different dividing patterns. Then two auxiliary predictions are generated by AMC with the two dividing patterns. The final prediction is calculated as a weighted-sum of the two auxiliary predictions. The interweaved prediction is only applied on the luma-component for affine-coded blocks with uni-prediction. Simulation results reportedly show 0.29% BD-rate savings under BMS Random Access (RA) configurations.

Multi-hypothesis prediction with two hypothesis is proposed for affine MC. Sub-block motion vectors at two sets of different sampling positions are derived and two prediction hypothesis is obtained by applying MC using the two sets of motion vectors. The two predictions are then averaged to get the final prediction.

It is applied to luma component in case of uni-prediction of affine coded blocks. So the worst case of memory access bandwidth is not changed.

It is noted that weighting matrices are designed to for combining the two prediction hypothesis.

It is noted that for boundary and corner positions in the auxiliary block partitioning patent, motion compensation is performed on 2x2 and 2x4/4x2 blocks.

-0.3% BD-Rate is observed for BMS tool test.

Recommendation: further study in CE.

[JVET-K0455](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3981) Cross-check of JVET-K0102: CE4-related: Interweaved Prediction for Affine Motion Compensation [Y. He (InterDigital)] [late] [miss]

[JVET-K0103](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3606) CE4-related: Simplified Affine Prediction [K. Zhang, L. Zhang, H. Liu, Y. Wang, P. Zhao, D. Hong (Bytedance)]

In this contribution, affine prediction in the benchmark set (BMS) is modified in three aspects. First, the sub-block size is fixed to be 4×4, instead of being calculated by an equation with division operations.

This was already decided in track B discussion earlier.

Second point in the contribution, the constraints of block size for affine merge mode and affine AMVP mode are unified. Both affine merge mode and the affine AMVP mode are applicable when the width and height of a block is greater than or equal to 8.

The coding performance is slightly degraded. And the encoding time for VTM tools test is increased by 2%. In case of 6-param affine inter prediction for 8x8 CU, 3 MVs instead of 4 MVs are signaled. It is commented that this does not reduce parsing efforts in BMS affine. It is claimed that the proposed change could reduce the number of lines in BMS SW.

Third, the sub-block size for chroma components is expanded from 2×2 to 4×4. Simulation results reportedly show 0.03% BD-rate increase on the Y component for Random Access (RA) configuration in average compared with BMS-1.1.

This implies using different motion vectors for the MC of luma and chroma. Additional sub-block MV derivation for chroma component is required.

Recommendation: test the 4x4 block chroma sub-block in CE.

[JVET-K0507](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4037) Crosscheck of JVET-K0103 (CE4-related: Simplified Affine Prediction) [H. Chen, H. Yang, J. Chen (Huawei)] [late] [miss]

[JVET-K0104](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3607) CE4-related: History-based Motion Vector Prediction [L. Zhang, K. Zhang, H. Liu, Y. Wang, P. Zhao, D. Hong (Bytedance)]

This contribution presents a History-based Motion Vector Prediction (HMVP) method for inter coding. In HMVP, a table of HMVP candidates is maintained and updated on-the-fly. After decoding a non-affine inter-coded block, the table is updated by adding the associated motion information as a new HMVP candidate to the last entry of the table. A First-In-First-Out (FIFO) or constraint FIFO rule is applied to remove and add entries to the table. The HMVP candidates could be applied to either merge candidate list or AMVP candidate list. It is asserted that the line buffer size is kept unchanged compared to VTM. When the merge candidate list size is extended by 10, compared with VTM-1.0, simulation results reportedly show that HMVP with FIFO and 16 entries of the table achieves 1.00%, 0.51% and 1.04% BD rate reduction for RA Main10, LDB Main10, and LDP Main10 configurations respectively. Compared with BMS-1.0, simulation results reportedly show that HMVP achieves 0.81%, 0.42% and 0.44% BD rate reduction for RA Main10, LDB Main10, and LDP Main10 configurations respectively. In addition, when the merge candidate list size is kept unchanged, compared with VTM-1.0, 0.82% BD rate reduction for RA Main10 configurations are reported by applying HMVP with constraint FIFO and only 8 entries of the table.

Reviewed in track B (chaired by JRO) Sunday 1330

Very promising (as compared to other merge proposals from CE4), much less complexity

Method is applied both for merge and MV prediction, gain on MV prediction is said to be <0.2%

Only partial results for BMS available so far

Test 3 is most promising in terms of complexity vs. performance

Further study in CE. Also study the impact on merge and MV prediction separately.

Revisit: Results for test 3 configuration only for merge, RA in BMS should be provided (confirmed by crosscheck). Could be a candidate for BMS adoption, if it provides reasonable gain.

[JVET-K0456](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3982) Cross-check of JVET-K0104: CE4-related: History-based Motion Vector Prediction [Y. He (InterDigital)] [late] [miss]

[JVET-K0189](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3698) Non-CE4: ATMVP simplification [H. Chen, H. Yang, J. Chen (Huawei)]

Contribution discussed together with JVET-K0346.

[JVET-K0460](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3986) Cross-check of JVET-K0189: Non-CE4: ATMVP simplification [Y. He (InterDigital)] [late] [miss]

[JVET-K0193](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3702) CE4-related: On performance improvements of Enhanced Interpolation Filter (EIF) [M. Sychev, G. Zhulikov, T. Solovyev, H. Chen, J. Chen (Huawei)] [late]

It is shown that the coding artifacts along a sharp edge in case of rotation could be alleviated by the per-pixel affine MC using EIF interpolation filter.

Combined with other techniques in affine MC, at maximum 0.7% coding gain could be obtained for VTM tool test in RA configuration.

It is commented that EIF may have interdependencies with OBMC and ALF in terms of coding gain.

Further study is encouraged.

[JVET-K0448](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3973) Crosscheck of JVET-K0193: CE4-related: On performance improvements of Enhanced Interpolation Filter (EIF) [A. Henkel (Fraunhofer HHI)] [late]

[JVET-K0194](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3703) CE4-related: On further complexity reduction of Enhanced Interpolation Filter (EIF) [M. Sychev, G. Zhulikov, T. Solovyev, J. Chen (Huawei)] [late]

Proposed two aspects of design improvement and complexity reduction of Enhanced Interpolation Filter (EIF) for inter prediction. EIF was initially proposed in JVET-J0024 and further studied in JVET-K020. Currently studied version of EIF use 3-tap high-pass filter (EIF3) and proposed as a solution for complexity reduction and worst case memory bandwidth reduction raised in JVET-J0081 for sub-block size 4x4. There are two aspects included in the document:

1. Directional Bilinear Interpolation Filter (DBIF) with reduced number of operation by 1/3 in respect to bilinear in EIF.
2. 16-bit friendly design of EIF.

Further study is encouraged.

[JVET-K0449](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3974) Crosscheck of JVET-K0194: CE4-related: On further complexity reduction of Enhanced Interpolation Filter (EIF) [A. Henkel (Fraunhofer HHI)] [late]

[JVET-K0195](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3704) CE4-5 related: Inter/Intra Boundary Padding [J. Brandenburg, R. Skupin, H. Schwarz, D. Marpe, T. Schierl, T. Wiegand (Fraunhofer HHI)]

[JVET-K0246](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3756) CE4.2-related: MV buffer reduction for non-adjacent spatial merge candidates [Y.-L. Hsiao, T.-D. Chuang, C.-Y. Chen, C.-W. Hsu, Y.-W. Huang, S.-M. Lei (MediaTek)]

[JVET-K0473](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4003) Crosscheck of JVET-K0246: MV buffer reduction for non-adjacent spatial merge candidates [J. Ye, X. Li (Tencent)] [late]

[JVET-K0267](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3778) CE4-related: Virtual Temporal Affine [F. Galpin, T. Poirier, F. LeLeannec, A. Robert (Technicolor)]

This contribution reports the results of a modified temporal affine merge candidate using virtual affine model from collocated candidates. We report the results of implementation in JVET VTM-1.0 and BMS-1.0. For tool ON/tool OFF test, simulation results reportedly show an average luma BD-rate gain of -0.03%, with gain up to -0.40% on CTC sequences, for Random Access (RA) compared with VTM. Moreover, additional tests on affine class sequences show an average luma BD-rate gain in RA of -0.23% for VTM, with gain up to -0.78%. In BMS configuration on the same affine class, simulation results reportedly show an average luma BD-rate gain of -0.28%, with gains up to -1.00%.

It is proposed to use the motion info in the top-left and top-right, or top-left and bottom-left corner, in the temporal collocated block, scale the two motion vectors, and then use them as merge candidates. At maximum two candidates will be derived.

Recommendation: study in CE.

[JVET-K0268](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3779) CE4-related: Affine tools from J0022 [F. Galpin, A. Robert, T. Poirier, F. LeLeannec (Technicolor)]

This contribution reports the results of all affine mode enhancement described in JVET-J0022 [1], implemented within the JVET VTM-1.0 and BMS-1.0. Simulation results reportedly show an average luma BD-rate gain of -3.77% in Random Access (RA) configuration, compared with VTM. Moreover, additional tests on affine class sequences show an average luma BD-rate gain in RA of -12.21% for TM, resp. -11.53%, for BMS without affine.

[JVET-K0548](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4078) Cross-check of JVET-K0268: CE4-related: Affine tools from J0022 [Y. He (InterDigital)] [late] [miss]

[JVET-K0297](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3811) CE4-related: Reduce line buffer for additional merge candidates [J. Ye, X. Li, S. Liu (Tencent)]

[JVET-K0427](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3949) Cross-check of JVET-K0297: CE4-related: Reduce line buffer for additional merge candidates [B. Choi (Sharp)] [late]

[JVET-K0301](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3815) CE4-related: Extension of merge and AMVP candidates for inter prediction [G. Li, X. Xu, X. Li, S. Liu (Tencent)]

In this contribution, it is proposed to add two spatial neighboring positions, which are from the middle position at the left and top edge of the current block, into motion vector predictor candidate list. The new candidates are put at the beginning of the merge candidate list and the AMVP candidate list, if available. The number of allowed merge candidates or AMVP candidates is kept unchanged. The proposed method was tested on BMS-1.0. It is reported that with VTM configuration, the result has 0.14%, 0.12%, and 0.16% BD-rate improvement for Random Access, Low Delay B, and Low Delay P, respectively.

Recommendation: study this in CE.

[JVET-K0491](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4021) Cross-check of JVET-K0301: CE4-related: extension of merge and AMVP candidates for inter prediction [T. Zhou, T. Ikai (Sharp)] [late]

[JVET-K0302](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3816) CE4-related: Bilateral Motion Vector Prediction [B. Choi, F. Bossen, A. Segall (Sharp)]

[JVET-K0415](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3936) CE4-related Cross-check of JVET-K0302 [S. Jeong (Samsung)] [late]

[JVET-K0304](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3818) CE4-related: Ranking based spatial merge candidate list for inter prediction [G. Li, X. Xu, X. Li, S. Liu (Tencent)]

In this contribution, it is proposed to replace the original spatial merge candidate list construction method with a ranking based method and the positions of merge candidates are extended to include more potential candidates. The proposed spatial candidates are derived and put at the beginning of the merge candidate list. The number of maximum merge candidates is kept unchanged. Other non-spatial merge candidates are added afterwards until the merge list is full. The proposed method was tested on BMS-1.0 reference software. It is reported that with VTM configuration, the result has 0.23%, 0.38%, and 0.51% BD-rate improvement for Random Access, Low Delay B, and Low Delay P, respectively.

It is proposed to sort spatial candidates in descending order of occurrence. The spatial candidates can be fetched from an L-shaped area around the current block. The existing spatial candidates are replaced by the candidates derived in this way.

For VTM test, 0.2%/0.4%/0.5% for RA/LDB/LDP configurations.

Recommendation: study this in CE.

[JVET-K0493](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4023) Crosscheck of JVET-K0304 (CE4-related: ranking based spatial merge candidate list for inter prediction) [C.-M. Tsai (MediaTek)] [late]

[JVET-K0335](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3850) CE4-related: Shape dependent control point selection for affine mode [Y. He, X. Xiu, Y. Ye (InterDigital)]

Affine coding mode is a tool in Benchmark Set (BMS). It is a four-parameter motion model which can describe translation, zooming, and rotation. The model parameters are signaled by motion vectors at two control points: top-left and top-right corners of an affine coding block. A shape dependent control point selection method is proposed in this document, where the positions of two control points are determined based on the coding block shape. Two sets of control points are defined for different coding block shape: top-left and top-right corners are selected for the coding block with horizontal and square shape; top-left and bottom-right corners are selected for the coding block with vertical shape. Compared to the VTM-1.0 anchor, the proposed method achieves average Y BD-rate reduction of 3.22% for Random Access and 2.26% for Low Delay B. Compared to the VTM-1.0 with affine mode enabled, it achieves average Y BD-rate reduction of 0.24% for Random Access and 0.20% for Low Delay B. Compared to the BMS-1.0 anchors, Y BD-rate reduction of 0.19% for Random Access cases and 0.24% for Low Delay B cases is reported.

The second part of the proposal is about affine MVP list construction. The sorting of constructed affine motion predictors are removed in the proposal. Not applicable to JVET-K0337 affine MVP list construction which had just been adopted in BMS.

Shape dependent control point selection could be applied to sorting inherited affine model from neighboring blocks in the list construction of affine merge and affine MVP.

Recommendation: further study shape dependent control point selection in CE, for both affine merge and affine MVP.

[JVET-K0509](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4039) Crosscheck of JVET-K0335 (CE4-related: Shape dependent control point selection for affine mode) [H. Chen, H. Yang, J. Chen (Huawei)] [late]

[JVET-K0346](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3863) CE4-related: One simplified design of advanced temporal motion vector prediction (ATMVP) [X. Xiu, Y. He, Y. Ye (InterDigital) (modify author list according to v4)]

In this contribution, one simplified design of the advanced temporal motion vector prediction (ATMVP) tool in the BMS-1.0 is proposed to reduce both the average and worst-case complexity. The proposed method uses the aspects from CE4.2.6 Test 1 for collocated block derivation and CE4.2.5 Test 2 for adaptive ATMVP sub-block size. Compared to CE4.2.6 Test 1, the proposed method adds one early termination to avoid checking both prediction lists of all the spatial merge candidates. Moreover, for further complexity reduction, it is proposed to derive the collocated blocks for the ATMVP from the same constrained range as that used for the temporal motion vector prediction (TMVP) in HEVC.

The BD-rate results of the proposed ATMVP scheme using uncompressed motion field and 8x8 compressed motion field are provided. With uncompressed motion field, the proposed method reportedly achieves average luma BD-rate savings of 1.01% for RA, compared to the VTM-1.0 anchor. The average encoding and decoding time are 102% and 105% for RA. With compressed motion field, the corresponding BD-rate savings are 0.95% for RA with the encoding and decoding time of 102% and 105% for RA.

This contribution was initially discussed in BoG JVET-K0546 with the following recommendations:

Three modifications to BMS ATMVP are suggested,

1. One fixed collocated picture is used to derive temporal motion information.
2. Slice level adaptive sub-block switching, 8x8 or 4x4.
3. Constrain the region from where ATMVP motion is derived to the collocated CTU plus one 4x4 block column outside the collocated CTU at the right hand side, the same region for HEVC TMVP.

Note that the first one point is the same as CE4.2.6.a, and the second point is the same as CE4.2.5.b.

It is commented that the third point could reduce the worst case memory access bandwidth.

Recommendation: integrate the three modifications into BMS ATMVP.

It is noted that several experts support adopting the modified BMS ATMVP to VTM, some other experts request further discussion.

The BoG report was presented in Track B Sun 1230, and it was decided to adopt ATMVP to VTM, with the modifications suggested in K0346, and the version of 8x8 motion storage (which is denoted as “test 2” in section 4 of K0346v4)

Decision(VTM): Adopt ATMVP from BMS, with the modifications from JVET-K0346 as described above, and 8x8 MV storage (using always the motion information from the top left block in the 8x8 region if there are blocks smaller than 8x8)

[JVET-K0423](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3944) Cross-check of JVET-K0346: CE4-related: One simplified design of advanced temporal motion vector prediction (ATMVP) [F. Le Léannec] [late]

[JVET-K0553](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4083) Crosscheck of JVET-K0346: CE4-related: One simplified design of advanced temporal motion vector prediction (ATMVP) [?? (??)] [late]

[JVET-K0350](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3867) CE4-related: Improvement on Merge/Skip mode with line buffer restriction [Y. Han, Y.-H. Chao, W.-J. Chien, M. Karczewicz (Qualcomm)]

[JVET-K0352](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3870) CE4-related: Encoder optimization on top of CE 4.2.13 [J. Ye, X. Li, S. Liu (Tencent)]

This contribution proposes an encoder optimization algorithm on top of CE 4.2.13. The proposed method reports 0.30% and 0.29% luma coding gain compare to CE 4.2.13 Test A in RA and LDB configuration on VTM with 1-2% encoding time increase.

It is claimed that about 0.15% coding gain if it is applied to VTM.

Recommendation: consider using this encoder optimization in CE on merge related tests, and in the anchors as well.

Decision (SW/CTC both VTM and BMS): Adopt JVET-K0352 encoder optimization

Note that the same encoder optimization is used in K0198. Software will be provided by the K0352 contribution.

[JVET-K0520](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4050) Cross-check of JVET-K0352 (CE4-related: Encoder optimization on top of CE 4.2.13) [X. Chen (HiSilicon)] [late]

[JVET-K0364](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3882) Non-CE4: separate merge candidate list for sub-block modes [T. Fu, H. Chen, H. Yang, J. Chen (Huawei)]

Currently, several sub-block based motion vector prediction (MVP) methods has been proposed, e.g., ATMVP, STMVP, Affine Merge enhancement in JVET-J0024 and Planar MVP in JVET-J0061. All of these methods generate fine granularity motion fields for CUs. Because of the commonality among these methods, this contribution presents a separate merge candidate list for sub-block based MVPs. The candidate list is constructed by integrating two or more sub-block based MVPs into a separate merge candidate list.

Three tests in this contribution,

1. Three merge list, one for VTM merge list with ATMVP, one for planar MV, one for affine candidates.
2. Two merge list, one for VTM, one for affine candidates + planar MV + ATMVP.
3. Two merge list, one for VTM merge list with ATMVP, one for affine candidates + planar MV.

It is commented that part of the coding gain could be obtained by more encoder RD checks.

It is commented that separate list can reduce the complexity of list construction at the decoder side.

Recommendation: study this in CE.

[JVET-K0367](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3885) Non-CE4: BMS affine improvements [H. Chen, H. Yang, J. Chen (Huawei)]

In this contribution, affine prediction in BMS1.0 is improved by 4 aspects: 1) bit-exact SIMD implementation for affine ME; 2) optimize for adaptive sub-block in affine MC; 3) division free for affine merge candidate derivation; 4) modify the condition check for affine merge mode.

For affine SIMD optimization, experimental results reportedly show on average 9%/13% encoding time reduction in RA/LB configurations over VTM with affine on, and 3%/7% encoding time reduction in RA/LB configurations over BMS anchor. For other aspects, experimental results reportedly show on average 0.00%/0.01% luma BD-rate change with 1%/1% encoding time reduction and 2%/2% decoding time reduction in RA/LB configurations over VTM with affine on.

Recommendation: adopt the SIMD implementation into BMS affine, after code review from W.-J. Chien and Yuwen He.

The second point on adaptive sub-block in affine MC is not relevant since fixed 4x4 sub-block partition has been adopted into BMS affine.

The contribution further replace the division operation by shifting when inheriting the affine model from neighboring blocks. This could be applied to both affine merge candidate derivation and affine MVP candidate derivation. It is commented that the rounding of the motion vectors should be aligned with other rounding operations applied to motion vectors, e.g. to AMVR.

The contribution further proposes to enable affine merge to blocks w&h >= 8. Note that in current BMS, the block size restriction w\*h >=64 pixels is applied to affine merge, w&h >= 16 to affine inter. It is commented that this fixes a bug of CPMV storage in the current BMS affine merge.

Recommendation: adopt this restriction to BMS affine. Note that the same restriction on affine merge mode is also proposed in JVET-K0103 and JVET-K0052.

It is mentioned that the JVET-K0103 propose applying the same block size restriction, w&h >= 8, to both affine inter and affine merge mode.

[JVET-K0484](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4014) Cross-check of JVET-K0367: Non-CE4: BMS affine improvements [K. Zhang (Bytedance)] [late]

[JVET-K0161](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3669) Spatial-temporal merge mode (non subblock STMVP) [T. Zhou, T. Ikai (Sharp)]

Note that this contribution number is invalid on JVET document website. And the contribution is registered again as JVET-K0532.

This contribution presents a spatial-temporal merge mode which is a simplified version of STMVP mode in JEM. It is asserted that this simplified merge mode produces more coding gain by courtesy of its optimized reference position while its’ no subblock feature is beneficial for hardware design. It is reported the bd-rate gain is 0.80 % and 0.28 % in RA and LB condition respectively.

Recommendation: study in CE.

[JVET-K0511](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4041) Crosscheck of JVET-K0161: Spatial-temporal merge mode (non subblock STMVP) [G. Li (Tencent)] [late]

[JVET-K0519](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4049) Cross-check of JVET-K0161 (Spatial-temporal merge mode (non subblock STMVP)) [X. Chen (HiSilicon)] [late]

[JVET-K0481](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4011) Integrating affine-based motion model in HEVC encoder for future video coding [M. S. Sayed (Egypt-Japan Univ. Sci. & Tech.)] [late]

In this contribution, it is proposed to add deformable block matching as an additional tool to the HEVC encoder. It is proposed to use affine transformation in a simple way to be integrated in the HEVC encoder to represent more complicated motion types such as rotation, zooming and deformation. It is claimed that the proposed tool shows slight reduction in the BD-rate however it represents a start point for further improvements in the direction of deformable block matching. It is claimed that better performance is expected from the proposed tool with improvements in its motion estimation and compensation processes.

Triangle-shaped PU is proposed. The motion vector at three vertices are coded. Per-pixel motion compensation is performed. Scan line algorithm for 2-D affine transformations is employed for calculating motion vector of each pixel inside a triangle partition. Motion accuracy is 1/4. Motion vector coding is not changed.

The proposed method is implemented on top of HM. Results show only slight coding gain on affine sequences.

It is commented that 1/4 accuracy for affine motion compensation may be a reason why this is no coding gain.

[JVET-K0514](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4044) CE4-related: Encoder optimization on top of CE 4.2.3 [Y. Han, W.-J. Chien, M. Karczewicz (Qualcomm)]

[JVET-K0532](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4062) Spatial-temporal merge mode (non subblock STMVP) [T. Zhou, T. Ikai (Sharp)]

[JVET-K0558](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4088) Report on complexity analysis of affine MVP candidate list construction [H. Chen]

(add abstract)

Was presented – see notes under CE4.

## CE5 related – Arithmetic coding engine (5)

Contributions in this category were discussed Friday 13 July 1600–1700 (chaired by GJS).

A CE for the arithmetic coding engine will be done. Throughput issues should be understood and sufficient gain should be shown to justify the change of the engine.

[JVET-K0273](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3785) CE5-related: Implementation considerations for entropy coding engine [F. Bossen]

Modifications are proposed to the entropy coding core engine such as to enable a wider variety of implementations. These newly enabled implementations may be beneficial in both software and hardware. For example, it is claimed that software implementations with reduced per bin cycle counts are enabled. It is asserted that the proposed changes do not noticeably impact compression efficiency.

Two changes are proposed, relative to techniques studied in CE5.

* Modify the constant 2b to 2b−1 (e.g., 32768 to 32767) in the probability estimate update function
* Modify the subinterval range computation for the LPS symbol to ((r >> 5) \* (qLPS >> (b − 5)) >> 1) + 4. Note that, alternatively, this equation can be implemented using a 32×8×8 = 2048 bit lookup table.

Throughput improvements and cycle count reductions on the order of 10-20% were reported for these tricks.

The coding efficiency impact of the first technique was estimated at 0.00% and for the second was 0.04%.

[JVET-K0385](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3904) CE5-related: Context state memory reduction [A. Said, H. Egilmez, Y-H. Chao, M. Karczewicz, V. Seregin (Qualcomm)] [late]

This contribution proposes a method to reduce the amount of RAM and ROM needed for binary arithmetic coding tools tested in CE 5.1 and CE 5.1A (JVET-K381 and JVET-K380). This is done by using a single adaptation window per context. The proposed method reportedly reduces the RAM and ROM memory requirements by 15 bits and 1 bit per-context, respectively. The proposed method reportedly provides average BD-rate gains of 0.8%, 0.8% and 0.9% for AI, RA and LD over the BMS, respectively, and the gains over the VTM are 0.6% for AI, RA and LD coding.

The proposed single window solution leads to a BD-rate loss about 0.4% for AI, RA and LD coding as compared to the results in CE5.1 and CE5.1A.

[JVET-K0510](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4040) Cross-check of JVET-K0385: CE5-related: Context State Memory Reduction [V. Lorcy (bcom), P. Philippe (Orange)] [late]

[JVET-K0430](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3952) CE5-related: State-based probability estimator [H. Kirchhoffer, J. Stegemann, D. Marpe, H. Schwarz, T. Wiegand (HHI)] [late]

An extension of the state-based probability estimator of VTM-1.0 to two states per context model is proposed. The transition table size is reduced from 64 to 32 elements and the two states per context model require 8 and 12 bit, respectively. Experimental results for the VTM configuration reportedly show overall luma BD rate reductions of 0.67%, 0.45%, and 0.41% for AI, RA, and LB, respectively. The BMS configuration reportedly show overall luma BD rate reductions of 0.71%, 0.46%, and 0.44% for AI, RA, and LB, respectively.

This proposal consists of two core elements. First, a state-based probability estimator is presented. The derivation of the subinterval range is the second part.

It was commented that for software it may be preferable to use a multiply rather than a table-lookup and that the bit width of storage is not critical unless it affects multiples of 8, 16, or 32 bits.

With custom window sizes there would be somewhat more gain, but this contribution did not consider how to combine the concept of custom window sizes with this.

[JVET-K0495](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4025) Crosscheck of JVET-K0430 (CE5-related: State-based probability estimator) [C.-M. Tsai (MediaTek)]

## CE6 related – Transforms and transform signalling (19)

Contributions in this category were discussed Friday 13 July 1700–1840 (chaired by GJS).

### Primary transforms

[JVET-K0113](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3617) CE6-related: EMT signalling [C. Rosewarne, A. Dorrell (Canon)]

Not needed to review, since it is not relevant to the adopted AMT scheme.

[JVET-K0130](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3635) CE6-related: Type4 only AMT [K. Abe, T. Toma (Panasonic)]

This proposes a simplified AMT that uses a DCT4 instead of a DST7. A DCT4 is part of what is needed for a DCT2, which is said to make this easier to implement. About 0.2% loss is reported relative to using a DST7. K0265 and K0394 are said to be similar, and K0292 also has a similar spirit but a different approach.

These are to be further studied in a CE.

[JVET-K0476](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4006) Cross-check of JVET-K0130: CE6-related: Type4 only AMT [Y. Kidani, K. Kawamura, S. Naito (KDDI)] [late]

[JVET-K0265](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3775) CE6-related: Reduction of the number of core transforms in AMT [K. Naser, F. Le Léannec, E. François (Technicolor)] [late]

See notes for K0130.

[JVET-K0462](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3991) Cross-check of JVET-K0265: CE6-related: Reduction of the number of core transforms in AMT [S. Bandyopadhyay, Y. He, Y. Ye (InterDigital)] [late]

[JVET-K0394](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3914) CE6-related: AMT with only Type2/Type4 DCT/DST [T. Tsukuba, M. Ikeda, T. Suzuki (Sony)] [late]

See notes for K0130.

[JVET-K0426](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3947) Cross-check of JVET-K0394: CE6-related: AMT with only Type2/Type4 DCT/DST [X. Zhao (Tencent)] [late]

[JVET-K0292](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3805) CE6-related: Compound orthonormal transform [X. Zhao, Z. Zhang, X. Li, S. Liu (Tencent)]

See notes for K0130. This contribution keeps a DST7 for 4 point and 8 point transforms and embeds a DST7 into a DCT2 for larger sizes.

[JVET-K0290](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3803) CE6-related: On 8-bit primary transform core [X. Zhao, X. Li, S. Liu (Tencent)]

The current VTM has 10 bit coefficients for the 64-point transform, but has 8 bit coefficients for shorter transform. This proposes an 8 bit transform design that uses the shorter HEVC transforms as a component of the 64-point transform. The performance impact was said to be negligible (0.00%).

Low QP usage was suggested to be especially important. For very low QP with AMT, a penalty of about 0.1% for AI and 0.05% for RA was reported. Without AMT, the penalty is said to be 0.00%.

It was commented that the basis functions have repeated values that could potentially cause plateaus visually, and so suggested visual testing.

Another participant noted that for high bit-depth data we have historically used a high-precision forward transform that is matched – i.e., designed as a high-precision inverse of the lower-precision inverse transform, and using it in the encoder gives measurable gain. It was asked how this would interact with such a scheme.

Further study is needed to study this and other potential approaches to transform simplification.

[JVET-K0419](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3940) Cross-check of JVET-K0290: CE6-related: On 8-bit primary transform core [T. Tsukuba (Sony)] [late]

[JVET-K0291](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3804) CE6-related: Fast DST-7/DCT-8 with dual implementation support [Z. Zhang, X. Zhao, X. Li, S. Liu (Tencent)]

This reports on a partial butterfly implementation of a DST7 that is compatible with a matrix multiply approach if a couple of number are changed by 1. About a 7-8% reduction in encoder runtime is reported by using it (with no loss in coding efficiency) in an implementation that is written in ordinary C without SIMD optimization.

40% to 50% operation count reduction is reported.

Only 3 numbers are affected (by 1).

It was commented by some participants that there would not really be a benefit expected for this.

The proponent has both 8 bit and 10 bit variations available.

To be further studied with other potential ways of simplifying the transform.

[JVET-K0429](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3951) Cross-check of JVET-K0291: CE6-related: Fast DST-7/DCT-8 with dual implementation support [P. Philippe (Orange)] [late]

[JVET-K0420](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3941) Cross-check of JVET-K0291: CE6-related: Fast DST-7/DCT-8 with dual implementation support [T. Tsukuba (Sony)] [late]

[JVET-K0299](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3813) CE6-related: Further simplification for AMT complexity reduction (CE6.1.2) [P. Philippe (Orange), V. Lorcy (bcom)]

This is a proposed way of reducing the implementation complexity of the inverse transform process for AMT. This should be further studied along with other complexity reduction methods for the inverse transform.

[JVET-K0126](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3631) CE6-related: Simplified multiple-core transform for intra residual coding [Y. Lin, Q. Yu, J. Zheng (HiSilicon), X. Cao, C. Zhu (UESTC)]

This contribution was discussed Saturday 14 July 1215 (chaired by GJS).

This contribution presents two simplified versions of the adaptive multiple-core transform (AMT) in BMS. On the one hand, the number of transform cores for intra residual is reduced from 5 to 3, as a unified transform design of AMT for intra and inter residual coding. On the other hand, encoding is accelerated by reducing the number of signalled transform pairs. It is reported that the proposed transform versions achieve better trade-off between coding performance and encoding/decoding complexity.

This is similar to what is proposed in K0171. The proposal is to not support one of the 5 transform combinations used in AMT. The one it proposes to not include is having a DCT8 style transform in both dimensions. The coding efficiency impact of omitting this combination is reported to be negligible (0.06% for AI).

It was asked whether there is any significant impact on the decoder for whether this combination is supported or not.

It was not clear whether there is a benefit for prohibiting the combination. If the benefit is intended to be saving encoder complexity, the scheme should be tested relative to an encoder-only optimization (the most obvious being simply not checking this combination). Generally, when considering potential syntax restrictions, if there is no benefit for decoders, testing should consider a good encoder-only alternative. Further study was encouraged.

[JVET-K0499](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4029) Crosscheck of JVET-K0126 (Simplified multiple-core transform for intra residual coding) [M.-S. Chiang (MediaTek)] [late]

### Secondary transforms

This topic remained open after the discussions of Friday 13 July.

This was further discussed on Saturday 14 July 1000 (chaired by GJS).

A CE will be done to measure the available gain and complexity of methods of secondary transforms relative to the VTM (which will now include AMT).

[JVET-K0100](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3603) CE6-Related: Matrix multiplication based NSST with reduced memory map [M. Salehifar, M. Koo, J. Lim, S. Kim (LGE)]

A non-separable secondary transform (NSST) called reduced secondary transform (RST) was proposed and investigated in CE 6.2.6.

A direct matrix multiplication NSST for 4x4 NSST (16x16 direct matrix multiplications) and 8x8 NSST (16x64 direct matrix multiplication) is introduced and investigated in this contribution. relative to a full secondary transform, this reduces the multiplication and multilayer complexity. Also results with memory reduction also reported.

This uses 16 secondary transform kernels instead of ~100 as used in the CE test.

Ordinarily, implementing a secondary transform larger than 4x4 has high complexity. This proposal use a sparse matrix decomposition to simplify the computation. The number of transform kernels is also reduced.

[JVET-K0440](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3965) Cross-check of JVET-K0100: CE6-Related : Matrix multiplication based NSST with reduced memory map [X. Zhao (Tencent)] [late]

[JVET-K0306](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3820) CE6-related: “Set of Transforms” selection and signalling scheme tested with different types of secondary transforms sets [M. Siekmann, C. Bartnik, S. Matlage, H. Schwarz, D. Marpe, T. Wiegand (HHI)]

This proposal involves having a set of secondary transforms and selecting a candidate set of secondary transforms using a LUT based on the transform size and intra mode, then sending an index to select the transform to apply (e.g., among 5 candidates). The secondary transform sizes are 4x4 and 8x8.

[JVET-K0501](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4031) Crosscheck of Section 2.3 in JVET-K0306 (CE6 - related: “Set of Transforms” selection and signalling scheme tested with different types of secondary transforms sets) [M.-S. Chiang (MediaTek)] [late]

[JVET-K0405](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3925) CE6-related: Secondary Transforms Coupled with a Simplified Primary Transformation [H. Egilmez, A. Said, Y.-H. Chao, M. Karczewicz, V. Seregin (Qualcomm)] [late]

[JVET-K0110](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3614) CE6-related: NSST restriction [C. Rosewarne, A. Dorrell (Canon)]

This proposes prohibiting NSST when the block aspect ratio is greater than 2:1. However, this does have some coding efficiency penalty. No action was taken on this.

### Shrink transform

This topic was discussed Friday 13 July 1840 (chaired by GJS).

[JVET-K0399](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3919) CE6-related: Simplification of Shrink Transform (CE6.1.9) [K. Kawamura, Y. Kidani, Sei Naito (KDDI)] [late]

This is applied to transform length 64 (only). From the decoder perspective, the decoder does a length-32 inverse transform and then upscales the result to form the final residual. An encoder may be designed to perform a transform of input length 64 and keep the lower frequency coefficients or to downsample and perform a shorter transform.

In the CE the upscaling used an 8 tap filter. In this contribution it used value replication.

Text was not available.

The BMS uses a 64 point inverse transform with only the 32 lowest-frequency transform coefficients.

This processing treats this particular block length with a different processing in a way that did not seem clearly better and potentially inconsistent with the rest of the design.

[JVET-K0416](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3937) Cross-check of JVET-K0399: CE6-related: Simplification of Shrink Transform [K. Abe, T. Toma (Panasonic)] [late]

## CE7 related – Quantization and coefficient coding (7)

Contributions in this category were discussed Saturday 14 July 1045–XXXX (chaired by GJS).

[JVET-K0070](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3571) Non-CE7: Entropy Coding for Dependent Quantization [M. Coban, J. Dong, M. Karczewicz (Qualcomm)]

In this document an alternate dependent quantization entropy coding scheme that allows grouping of bypass coded bins at coefficient group level is described. The state machine driving the quantizer selection and coefficient coding algorithm is driven by the significance of a coefficient rather than the coefficient level parity allowing grouping of remaining level bypass coded bins for efficient parsing. The simulations results reportedly show 4.41% AI, 3.01% RA, 2.44% LB BD-Rate gains versus the VTM-1.0 anchor.

To be studied in a CE.

[JVET-K0424](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3945) Cross-check of JVET-K0070: Non-CE7: Entropy Coding for Dependent Quantization [H. Schwarz (Fraunhofer HHI)] [late]

[JVET-K0072](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3574) Non-CE7: Alternative entropy coding for dependent quantization [H. Schwarz, T. Nguyen, D. Marpe, T. Wiegand (Fraunhofer HHI)]

This contribution describes an alternative approach for entropy coding of transform coefficient levels for dependent quantization. It is asserted that the proposed entropy coding can be used with dependent scalar quantization and conventional scalar quantization, only a single 16-bit parameter (which represents the state transition table for dependent quantization) has to be set depending on the quantization approach used. The proposal uses exactly the same concept of dependent quantization as JVET-K0071 (tested in CE7-2.1). In comparison to the transform coefficient coding in JVET-K0071, the transform coefficient levels of a subblock are coded using multiple passes over the scan positions, where all bypass-coded bins of a subblock are transmitted after the regular-coded bins. The proponents assert that the bin-to-bin dependencies are reduced relative to the approach in JVET-K0071 and that the alternative approach is more suitable for high-throughput hardware implementations.

The following average results are reported relative to VTM-1:

Coefficient coding only:

* AI: −1.6%, −1.5%, −1.7% (Y, Cb, Cr) at 105% encoder and 102% decoder run time;
* RA: −1.0%, −1.0%, −1.0% (Y, Cb, Cr) at 104% encoder and 100% decoder run time;
* LB: −0.7%, −1.6%, −0.4% (Y, Cb, Cr) at 104% encoder and 101% decoder run time;
* LP: −0.8%, −1.7%, −0.3% (Y, Cb, Cr) at 104% encoder and 99% decoder run time.

Combination of coefficient coding and dependent quantization:

* AI: −4.8%, −1.6%, −1.6% (Y, Cb, Cr) at 119% encoder and 100% decoder run time;
* RA: −3.2%, −1.5%, −1.5% (Y, Cb, Cr) at 113% encoder and 99% decoder run time;
* LB: −2.6%, −2.8%, −0.9% (Y, Cb, Cr) at 116% encoder and 100% decoder run time;
* LP: −2.4%, −2.7%, −1.0% (Y, Cb, Cr) at 118% encoder and 98% decoder run time.

Encoder runtimes are increased about 15-20%.

It was discussed whether sign data hiding can be combined with the TCQ schemes. Further study of this is encouraged.

As a fallback mode, the state aspects can be disabled with ordinary quantization.

It was suggested to put the fallback mode switch at the picture level.

Decision: Adopt (to WD/VTM) with fallback switch at the picture level.

[JVET-K0428](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3950) Cross-check of JVET-K0072: Non-CE7: Alternative entropy coding for dependent quantization [M. Coban (Qualcomm)] [late]

[JVET-K0319](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3833) CE7-related: TCQ with high throughput coefficient coding [J. Dong, M. Coban, M. Karczewicz (Qualcomm)]

This is another proposed variation of TCQ. This was trying to combine the JEM coefficient coding with TCQ. Detailed presentation was not requested, although the proponent may wish to include it in the planned CE.

[JVET-K0513](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4043) Cross-check of JVET-K0319: CE7-related: TCQ with high throughput coefficient coding [L. Zhang (Bytedance)] [late]

[JVET-K0281](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3794) Inter-Component Context Modelling for Coded Block Flag [C. Rudat, P. Haase, H. Schwarz, D. Marpe, T. Wiegand (HHI)]

This is reporting the gain for a change in the context modeling for the coded block flag, making the second chroma CBF dependent on the first one, with a reported gain of about 0.2% for AI and 0.1% for RA. Detailed presentation was not requested.

[JVET-K0310](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3824) AHG13: Sign Data Hiding [P. Yin, S. McCarthy, F. Pu, T. Lu, W. Husak, T. Chen (Dolby), T. Tsukuba (Sony)]

It was said that sign data hiding can provide 0.9% on AI and 0.7% on RA for the VTM with 3-4% encoder increase (and no decoder complexity increase). See also the AHG13 report, which indicated that the previous non-adoption of SDH had a larger impact on coding efficiency than other removed features relative to HEVC (other having 0.4% impact or less, and in RA it was the only one with a significant impact), and had a good tradeoff of coding efficiency and complexity.

Decision: Adopt SDH into WD/VTM (can only be used when TCQ is disabled).

## CE8 related – Current picture referencing (1)

Contributions in this category were discussed Saturday 14 July 1145–XXXX (chaired by GJS).

[JVET-K0050](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3551) CE8 related: Intra Region-based Template Matching for luma and chroma [G. Venugopal, K. Müller, H. Schwarz, D. Marpe, T. Wiegand (HHI)]

This document proposes the region-based template matching method for both luma and chroma for intra prediction. The proposed method is an extension of the algorithm explained in the JVET document JVET-K0048 to the chroma components. The experimental results of the RTM method for luma and chroma for the all intra (AI) configuration against VTM-1.0 shows an overall BD-rate gain of 2.30% with 120% and 215% decoder and encoder run-time respectively. For the random access (RA) case, the overall coding gain is 0.76% with 103% decoder run-time and 117% encoder run-time, and for the low delay (LB) case it is 0.18% coding gain with 101% decoder run-time and 109% encoder run-time. While using BMS-1.0 as the anchor, the overall BD-rate gain is 1.32% for the all intra (AI) configuration with 104% decoder complexity and 113% encoder complexity. For the random access (RA) configuration, the overall coding gain is 0.43% with 101% decoder run-time and 105% encoder run-time, and for the low delay (LB) case it is 0.16% overall gain with 100% decoder run-time and 102% encoder run-time.

This extends a prior proposal to support chroma, using displacement vectors derived from luma.

This provides more gain on the CTC (2.3% for AI, 0.8% for RA) than other CPR proposals (1.3% for AI, 0.5% for RA), although it provides less gain for SCC than other CPR methods. It uses a decoder template matching search.

Further study is encouraged. Complexity issues need to be considered.

[JVET-K0401](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3921) Crosscheck of JVET-K0050: CE8 related: Intra Region-based Template Matching for luma and chroma [X. Xu, X. Li (Tencent)] [late]

## CE9 related – Decoder side motion vector derivation (16)

Contributions in this category were discussed Saturday 14 July in track B 1115–1310 and 1420-1645 (chaired by JRO).

[JVET-K0093](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3596) CE9-Related: Restricted template matching schemes to mitigate pipeline delay [N. Park, J. Nam, H. Jang, J. Lee, J. Lim, S. Kim (LGE)]

Conventional template matching method considers reconstructed samples of the previous coding block as template. It could cause serious pipeline delay when reconstructed pixel is not ready.

This contribution proposes a restricted template matching (RTM) scheme to reduce pipeline delay issue. The proposed RTM is used to reorder MERGE candidates. The following three tests were investigated to evaluate the proposed method;

- Test #1 : Restricted template region

- Test #2 : Restricted reordering of merge candidate list

- Test #3 : Harmonization of the above two tests

The experimental results for Test #1 reportedly show xx% and xx% bit rate savings compared to VTM anchor and xx% and xx% bit rate savings compared to BMS anchor in RA and LB configurations, respectively. The experimental results on Test #2 reportedly show xx% and xx% bit rate savings compared to VTM anchor and also xx% and xx% bit rate savings compared to BMS anchor in RA and LB configurations, respectively. Finally, the experimental results for Test #3 reportedly show xx% and xx% bit rate savings compared to VTM anchor and also show xx% and xx% bit rate savings compared to BMS anchor in RA and LB configurations, respectively.

The proposal avoid using the immediately preceding block for template matching (test 1)

The proposal reduces the memory bandwidth by applying template matching only for sub-group of merge candidate (2 sub-group, it is identified via merge index which one is used)

Combined test 3 provides 0.42% in VTM and 0.28% in BMS for RA conf., roughly 0.5% for both in LDB.

Pipelining solution is interesting

However, memory bandwidth would still be significantly increased. This seems to be a general problem when template matching is used between current picture and new areas from reference picture that have not been accessed before.

Further study would be welcome, to solve the memory bandwidth problem.

It is suggested that for example, template matching might only be performed for candidates in the merge list that are very similar and would access the same area from the reference picture.

[JVET-K0505](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4035) Cross-check of JVET-K0093 (CE9-Related : Restricted template matching schemes to mitigate pipeline delay) [A. Karabutov, S. Ikonin, R. Chernyak (Huawei)] [late]

[JVET-K0105](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3608) CE9-related: Simplification of Decoder Side Motion Vector Derivation [H. Liu, L.Zhang, K. Zhang, Y. Wang, P. Zhao, D. Hong (Bytedance)]

This contribution presents two modifications on Decoder Side Motion Vector Derivation (DMVR) in the benchmark set (BMS): 1) DMVR for 4x4 blocks is disabled; 2) Partial Sum of Absolute Differences (SAD) is calculated instead of full SAD. Simulation results reportedly show 0.05%/0.07% BD-rate increase for BMS-1.0/VTM-1.0 under BMS Random Access (RA) configurations.

Two aspects:

* Disallow DMVR for 4x4 blocks (having 4x8/8x4 as smallest block size decreases worst case memory bandwidth increase from 140% to 136%)
* Perform SAD calculation on every second row

Further study in CE, but also restriction to even larger block sizes should be considered.

It is noted that the results show an encoding time increase, but the measurement is inaccurate according to proponents.

[JVET-K0490](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4020) Cross-check of JVET-K0105: CE9-related: Simplification of Decoder Side Motion Vector Derivation [T. Ikai (Sharp)] [late]

[JVET-K0187](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3696) CE9-related: Low latency template based motion vector refinement [F. Chen, L. Wang (Hikvision)] [late]

This contribution describes a motion vector refinement (MVR) based on neighboring motion information. With simple motion information of neighboring blocks, a template is generated to refine an initial MV of the Merge mode candidate. Since the template only utilizes the basic motion information, rather than the reconstructed or intra predicted samples of the spatial neighbors, the proposed method brings in low latency. It is reported that, compared with BMS-1.0, 0.85 %, 0.96 % and 0.20% coding gains are earned for LDP, RA, and LDB configurations on average, respectively. And 2.80%, 0.47%, 0.28% coding gains are earned for LDP, RA and LDB configurations on average when compared with VTM-1.0.

The latency issue is solved by using the initial MV (not the refined) for the intialization of TM in current block.

It should from implementation aspects if the restriction of using only 1 line or 1 column is really necessary.

Further study in CE.

[JVET-K0446](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3971) Crosscheck of JVET-K0187: CE9-related: Low latency template based motion vector refinement [S. H. Wang, S. S. Wang (PKU)] [late]

[JVET-K0256](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3766) CE9-related: MV prediction modifications for decoder-side MV derivation tools [T.-D. Chuang, C.-Y. Chen, Y.-W. Huang, S.-M. Lei (MediaTek)]

Decoder-side motion vector derivation (DMVD) tools can cause serious degradation of processing throughput in hardware decoders. In DMVD, an initial motion vector (MV) is first obtained with simple processes without any need to access reference picture samples, and then the initial MV is refined within a search range centered at the initial MV using refinement processes that need to access reference picture samples to obtain the final MV. In this contribution, a MV prediction modification is proposed for DMVD as follows. When reconstructing MV(s) of a non-DMVD current block or initial MV(s) of a DMVD current block, “MVs from non-DMVD previous blocks” and “initial MVs from DMVD previous blocks” (instead of final MVs from DMVD previous blocks) can be referenced. With the proposed method, at parsing stage of a pipelined hardware decoder, all reference picture samples needed to perform DMVD can be pre-fetched using initial MVs of DMVD blocks. To reduce coding efficiency degradation, final MV(s) of a DMVD previous block can be referenced when the DMVD previous block is above the current coding tree unit (CTU) row or in a previously decoded picture. Final MV(s) of DMVD blocks are used in overlapped block motion compensation (OBMC) and deblocking processes. Simulation results show that decoder-side motion vector refinement (DMVR) and OBMC with the proposed method can achieve -2.59% luma BD-rate for VTM-1.0-RA, and 72% of DMVR and OBMC coding gain is preserved by applying the proposed method. It is also shown that DMVR, OBMC, and Core Experiment 9.2.1 (CE9.2.1) bilateral matching merge mode with the proposed method can achieve -5.81% luma BD-rate for VTM-1.0-RA, and 83% of DMVR, OBMC, and CE9.2.1 bilateral matching merge mode coding gain is preserved by applying the proposed method.

Results without OBMC: The method that resolves the latency problem gains 1.58% for VTM, and 0.95% for BMS. This seems to be slightly better than the method CE9.1.1.a, as the current has some more aspects to resolve the latency problem.

Further study in CE (without OBMC)

For the upcoming CE, complexity characteristics of different algorithms need to be investigated more systematically. For each proposal, data must be provided that document

* Latency characteristics
* Memory bandwidth increase
* Storage of vectors, samples, …
* Complexity in terms of operations

Side activity (M. Zhou, W.J. Chien, X. Li, X. Xiu, S. Sethuran, X. Chen) to work out a list of criteria.

[JVET-K0275](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3787) Non-CE9: DMVR without Intermediate Buffers and with Padding [S. Esenlik, I. Krasnov, Z. Zhao, J. Chen (Huawei)]

A decoder side motion vector refinement (DMVR) method using bilateral matching is proposed, where the intermediate interpolation operation before calculation of MRSAD cost function is removed. Moreover the final prediction is obtained using padding in order to eliminate the additional the memory bandwidth requirement.

According to the proposal when the maximum MRSAD computations are restricted to 13, -4.08% luma coding gain with 107%, encoding time and 126% decoding time over VTM Anchor (RA), and -0.91% luma coding gain with 101%, encoding time and 99% decoding time over BMS Anchor (RA) is achieved.

The proposal targets the memory bandwidth problem (which is particularly more severe for bilateral matching) by applying padding. Further, integer samples are used, which reduces the computational complexity as well.

No refined motion vectors are used from spatial neighbours. Refined MVs are used for deblocking and temporal prediction, which does not cause a latency problem.

Further study in CE.

[JVET-K0516](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4046) Cross-check of JVET-K0275: Non-CE9: DMVR without Intermediate Buffers and with Padding [N. Park, J. Lim (LGE)] [late]

[JVET-K0288](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3801) CE9-related: Memory bandwidth reduction for DMVR [M. Xu, X. Li, S. Wenger, S. Liu (Tencent)]

Decoder side motion vector refinement (DMVR) techniques were proposed to refine motion information at decoder side. With DMVR, an initial MV pair for a block is first identified at decoder. Then the MV pair is refined within a search range. As the final MV may be at any position within the search range, the memory bandwidth requirement of DMVR mode would be higher than regular inter mode at decoder side. In this contribution, a memory bandwidth reduction method is proposed for DMVR. It is reported the memory bandwidth of DMVR can be kept the same as regular inter mode with the proposed method at the cost of 0.06% average luma BD-rate increase.

The method is based on using a shorter interpolation filter for DMVR than later for interpolation.

Investigate in CE. Other filter lengths than 6 should also be tested.

[JVET-K0421](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3942) Cross-check of JVET-K0288: CE9-related: Memory bandwidth reduction for DMVR [Y.-W. Chen, X. Wang (Kwai Inc.)] [late]

[JVET-K0295](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3808) CE9-related: Constrained Decoder Side Motion Vector Derivation [X. Li, M. Xu, S. Liu (Tencent)]

Decoder side motion vector derivation (DMVD) techniques were proposed to derive motion information at decoder side after parsing stage. As motion vector (MV) of an inter block may depend on the results of its spatial neighbor’s MV derivation, parsing/determining the initial MV of the current block and prefetching reference block cannot be started until the MV derivation process is finished for the previous block, which leads to serious latency in hardware pipeline. It is reported that the issue can be solved with the method proposed in this contribution at the cost of about 0.6% luma BD-rate increase for VTM-1 + DMVR. It is also reported that the method can be combined with other DMVD methods with similar luma BD-rate increase even when other DMVD method provides higher coding gain.

The approach is different from e.g. CE9.1.1.a, where the MV of a neighboring block is always used (when available), but if it has been refined, the starting point of the refinement is used. In K0295, the vector is marked as unavailable for MVP or merge, when it may have been refined (as per several conditions). This may save some storage, and according to the results presented, the loss is less than in CE9.1.1.a.

Further study in CE.

It is mentioned that just disallowing the dependency from the immediate neighbour may not fully solve the latency problem.

[JVET-K0407](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3927) Cross-check of JVET-K0295: CE9-related: Constrained Decoder Side Motion Vector Derivation [J. Ma (HHI)]

[JVET-K0347](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3864) CE9-related: Addressing the decoding latency issue for decoder-side motion vector refinement (DMVR) [X. Xiu, Y. He, Y. Ye (InterDigital)]

Decoder-side motion vector refinement (DMVR) is included in the benchmark set (BMS)-1.0. The DMVR needs to generate L0 and L1 prediction signals to derive the refined motion vectors (MVs) of one bi-predicted merge candidate. In the current DMVR design, the refined MVs are used to predict the neighboring MVs of the current coding unit (CU). Such design is reported to have decoding latency issue due to the interdependency among the decoding of spatial neighboring CUs.

In this contribution, three solutions are provided to address DMVR’s decoding latency issue. In solution one, instead of using the refined MVs, it is proposed to use the original MVs (i.e., unrefined MVs) to predict the MVs of a DMVR CU’s spatial neighbors. Solution one removes the decoding latency of the DMVR. In solution two, it is proposed to use the unrefined MVs to predict the MVs of a DMVR CU’s neighboring CUs that are in the same coding tree unit (CTU). But, for the neighboring CUs that are not in the same CTU, the refined MVs are used as the predictor. Solution two allows independent decoding of multiple inter CUs inside one CTU, but not across different CTUs. In solution three, it is proposed to use the unrefined MVs of one DMVR CU to predict the MVs of its neighboring CU that is in the same CTU row. But, for the neighboring CUs that is from one different CTU row, the refined MVs are used for spatial MV prediction. The third solution allows parallel decoding of multiple inter CUs within the same CTU row while the decoding of the CUs on the top boundary of one CTU is still dependent on that of the top neighboring CTU. Additionally, for both solutions, the unrefined MVs are used for deblocking, and the refined MVs are temporal motion vector prediction.

Experimental results show that compared to VTM-1.0 anchor, solution one provides average {Y, U, V} BD-rate savings of {1.41%, 1.45%, 1.47%} for the RA configuration with the encoding and decoding time of 112% and 130%. For solution two, the corresponding {Y, U, V} BD-rate savings are {1.71%, 1.67%, 1.73%} with the encoding and decoding time of 111% and 129%. For solution three, the corresponding {Y, U, V} BD-rate savings are {1.58%, 1.56%, 1.60%} with the encoding and decoding of 111% and 132%.

Difference of solution 1 versus CE9.1.1.a is the TMVP part (refined MV used for TMVP here). Results are in similar range.

Solutions 2 and 3 release the constraint of not using the refined neighbour at CTU boundary. This reportedly provides some compression benefit. More study on implications of imposing special rules would be necessary.

Further study in CE.

[JVET-K0425](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3946) Crosscheck of JVET-K0347: CE9-related: Addressing the decoding latency issue for decoder-side motion vector refinement (DMVR) [C.-H. Hung, W.-J. Chien, M. Karczewicz (Qualcomm)] [late]

[JVET-K0360](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3878) CE9-related: Bilateral Matching with Constrained Motion Vector Storage [C.-C. Chen, W.-J. Chien, M. Karczewicz (Qualcomm)]

This contribution introduced a constraint on the Template-free DMVR (CE9.2.5/9.2.6, JVET-K0359) to prohibit using refined motion vectors (MVs) for MV prediction and boundary strength (BS) calculation in deblocking filter. Specifically, refined MVs are taken only for motion compensation, but not for storage. None of them is kept in the motion field. Thus, only un-refined motion field is available for MV prediction and BS calculation. Experiments were conducted on top of the BMS (BMS-1.0 w/ BMS cfg.) and VTM anchors (BMS-1.0 w/ VTM cfg.) with Random Access configuration. Detailed performance results are summarized as follows.

Modified CE9.2.5 vs. BMS: (Y) -0.6%, (U) -0.7%, (V) -0.6%, (Enc.) 103%, (Dec.) 113%.

Modified CE9.2.6 vs. BMS: (Y) -0.3%, (U) -0.3%, (V) -0.3%, (Enc.) 100%, (Dec.) 103%.

Modified CE9.2.5 vs. VTM: (Y) -3.6%, (U) -3.4%, (V) -3.5%, (Enc.) 118%, (Dec.) 165%.

Modified CE9.2.6 vs. VTM: (Y) -3.0%, (U) -2.7%, (V) -2.8%, (Enc.) 109%, (Dec.) 142%.

The contribution proposes a low-latency version of 9.2.5 and 9.2.6 which per have a better performance than “BMS-DMVR”, but use different approaches e.g. larger search range. The loss compared to the original proposals is around 1.5% (which is larger than what was reported for BMS-DMVR before). The method completely gives up the usage of refined MVs except for MC of the current block. Therefore no extra buffer is required.

Same solution for solving the latency problem as in CE9.1.1.a.

Further study of 9.2.6 in CE, with the solution of latency problem proposed here.

(Note that 9.2.5 has better compression performance, but is unacceptable in terms of increased memory BW).

[JVET-K0463](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3992) Cross-check of JVET-K0360: Bilateral Matching with Constrained Motion Vector Storage [I. Krasnow (Huawei)] [late]

[JVET-K0361](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3879) CE9-related: Harmonization between CE9.2.6 (DMVR with Template-free Bilateral Matching, JVET-K0359) and CE9.2.9 (DMVR with Bilateral Matching, JVET-K0217) [C.-C. Chen, W.-J. Chien, M. Karczewicz (Qualcomm), S. Esenlik, I. Krasnov, Z. Zhao, M. Xiang, H. Yang, J. Chen (Huawei)]

This document proposes a harmonization method of CE9.2.6 (DMVR with Template-free Bilateral Matching, JVET-K0359) and CE9.2.9 (DMVR with Bilateral Matching, JVET-K0217). The followings are applied to the achieve the harmonized design: 1.) bilateral matching based on MVD mirroring without scaling; 2.) adaptive search pattern; 3.) MRSAD as the cost function; 4.) refined MVs not for spatial MV prediction; 5.) 2-tap bilinear filter for interpolating search range; 6.) early-skip condition based on MVD between merge candidates. When compared with the VTM anchor (i.e. BMS-1.0 with VTM cfg.) of CE9, the proposed method achieves an average Y BD-rate saving of 2.8%, with 5% and 25% increase in encoding and decoding time, respectively. When compared with the BMS anchor (i.e. BMS-1.0 with BMS cfg.), it shows an average Y BD-rate saving of 0.3%, without negative impact on runtime, respectively.

Benefit compared to CE9.2.9l (the solution which also resolves the latency problem) is not too obvious. 0.04% rate reduction, small reduction of enc./dec. run time. Would require more detailed analysis of complexity impact.

[JVET-K0406](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3926) Crosscheck of JVET-K0361: CE9-related: Harmonization between CE9.2.6 and CE9.2.9 [J. Li, C. Lim (Panasonic)] [late]

[JVET-K0041](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3542) Decoder Side MV Refinement/Derivation with CTB-level concurrency and other normative complexity reduction techniques [S. Sethuraman, J. Raj, S. Kotecha (Ittiam)]

This contribution presents a method for determining the availability of the refined motion vectors (MVs) of spatially neighboring coding units for use in AMVP process or as a starting point for DMVR/PMMVD process of a current coding unit. The proposed method enables concurrent reference data pre-fetch to be possible for all CUs within a CTB during one stage of a processing pipeline followed by concurrent decoder-side MV refinements, if required, and motion compensated prediction in a following stage of the pipeline. A configurable lag between consecutive CTB rows is proposed to allow more top CTB-row refined MVs to be available. When compared to considering all refined MVs within the current access unit to be unavailable, the proposed method offers BDRATE gains of up to 0.6% . In addition, two normative complexity reduction methods are proposed. The first is a modified search pattern based integer distance refinement procedure that attempts to strike a balance between the reduction of unconditional cost evaluations and the number of dependent stages during refinement In the second method, a parametric error surface based sub-pixel displacement estimation procedure approach is proposed to reduce internal memory and computational requirements. VTM tool ON results are provided (and BMS tool OFF results will be provided in an updated version of this contribution).

The results indicate that using bilinear filter in the search is not necessarily worse than DCTIF, in particular if larger number of iterations is used (i.e. enlarged search range).

There is also a variant which uses bilateral matching with symmetric vectors.

There is also a solution for the latency problem by a pipeline assumption.

This proposal has various aspects that are worthwhile studying in upcoming CEs, such as search strategies, usage of bilinear filters, etc.

[JVET-K0480](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4010) Non-CE9: A computational complexity analysis for DMVR [M. Zhou, B. Heng (Broadcom)] [late]

This contribution provides a computational complexity analysis for the DMVR relative to the bi-directional 8x8 motion compensation (MC). Based on the analysis results it is recommended to 1) disable the DMVR for PU sizes smaller than 8x8, 2) make the refinement search range PU size dependent for the DMVR, and 3) study the impact of using short tap filters for the DMVR.

Interesting contribution, to be further considered in the context of setting up complexity evaluation method for the CE.

For example, with shorter tap filters, search range could be increased with same memory bandwidth and same complexity.

It is noted in the discussion that it would be interesting in the upcoming CE to compare different approaches of DMVR with comparable amount of memory bandwidth usage, and comparable amount of computational complexity.

If we would be able to set up rules imposing some complexity limitations, it would be possible to better compare different algorithms (or parameter variations of an algorithm) for their RD performance at a certain worst-case complexity point.

S. Esenlik is mandated to coordinate setup of the next CE9.

[**JVET-K0485**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4015) **CE9-related: A simplified bi-directional optical flow (BIO) design based on the combination of CE9.5.2 test 1 and CE9.5.3 [X. Xiu, Y. He, Y. Ye (InterDigital), C.-Y. Chen, C.-Y. Lai, Y.-W. Huang, S.-M. Lei (MediaTek)] [late]**

This contribution proposes one combined bi-directional optical flow (BIO) method based on CE9.5.2 test 1 with a simpler gradient filter {-1, 0, 1} and CE9.5.3 with two-stage BIO early termination. Simulation results show that compared to VTM-1.0, the proposed scheme provides on average {Y, U, V} BD-rate savings of {2.80%, 0.94%, 0.68%} for RA with average encoding and decoding time of 108% and 123%.

This combination has slightly better performance than CE9.5.3 and slightly worse than CE9.5.2. However, the encoder/decoder run times (as reported so far relative to VTM) seem to be even faster than for CE9.5.3. Further, the worst case complexity is largely reduced (e.g. from >100 mul/sample to 13 mul/sample, and the compression benefit is large.

It was later reported that the method provides 1.2% BR reduction in BMS.

Decision (BMS): Adopt JVET-K0485.

[**JVET-K0538**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4068) **Cross-check of JVET-K0485 [Fabrice Le Léannec (Technicolor)] [late]**

[**JVET-K0550**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4080) **Cross-check of JVET-K0485: CE9-related: A simplified bi-directional optical flow (BIO) design based on the combination of CE9.5.2 test 1 and CE9.5.3 [Y.-W. Chen (Kwai Inc.)] [late]**

## CE10 related – Combined and multi-hypothesis prediction (5)

Contributions in this category were discussed Saturday 14 July 2040–2145 (chaired by JRO).

[JVET-K0148](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3655) CE10 related: Combining multi-hypothesis prediction with triangular prediction unit mode [R.-L. Liao, C. S. Lim (Panasonic)]

This contribution provides test results for combining the multi-hypothesis prediction with the triangular prediction unit mode. Two tests of multi-hypothesis prediction, the CE10.1.4 proposed by MediaTek and the CE10.1.8 proposed by Fraunhofer HHI, are used to combine with the CE10.3.2 triangular prediction unit mode. Two different combination are tested and their coding results are reported as follows:

* CE10.1.4 plus CE10.3.2:
  + (VTM configuration) (RA) -1.88% BD-rate with 139% encoding time and 107% decoding time

(LB) -2.04% BD-rate with 149% encoding time and 106% decoding time

* CE10.1.8 plus CE10.3.2:
  + (VTM configuration) (RA) -2.55% BD-rate with 150% encoding time and 106% decoding time

(LB) -3.75% BD-rate with 169% encoding time and 106% decoding time

Already discussed in context of CE – to be investigated in next round of CE10.

Revisit: Report results on BMS, if possible before end of meeting

[JVET-K0258](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3768) CE10-related: OBMC complexity reduction and parallel blending [C.-C. Chen, C.-W. Hsu, Y.-W. Huang, S.-M. Lei (MediaTek)]

Two aspects about overlapped block motion compensation (OBMC) are proposed. The first is to perform data reuse to achieve lossless runtime reduction for original JEM-based OBMC. The second is to replace sequential computing with parallel computing for original JEM-based OBMC sample blending process. For the first aspect, when OBMC is enabled additionally, the reductions of encoder runtime are 9% for RA VTM-1.0, 10% for LB VTM-1.0, 2% for RA BMS-1.0, and 3% for LB BMS-1.0; the savings of decoder runtime are 12% for RA VTM-1.0, 12% for LB VTM-1.0, 5% for RA BMS-1.0, and 5% for LB BMS-1.0. The proposed parallel blending has no noticeable BD-rate change or run time change.

When the proposed lossless runtime reduction techniques and parallel blending are applied, OBMC achieves -1.04%, -1.41%, -1.26%, and -1.93% luma BD-rates with 5%, 7%, 3%, and 5% encoding time increases and 11%, 13%, 25%, and 33% decoding time increases for RA VTM-1.0, LB VTM-1.0, RA BMS-1.0, and LB BMS-1.0, respectively. Chroma BD-rate savings are about 1% higher than luma BD-rate savings.

When the proposed lossless runtime reduction techniques and parallel blending are applied, CU-boundary-only OBMC (i.e., no sub-block OBMC) achieves -1.04%, -1.41%, -0.89%, and -1.38% luma BD-rates with 5%, 7%, 1%, and 3% encoding time increases and 11%, 13%, 6%, and 7% decoding time increases for RA VTM-1.0, LB VTM-1.0, RA BMS-1.0, and LB BMS-1.0, respectively. Chroma BD-rate savings are about 1% higher than luma BD-rate savings.

The first aspects are non-normative, but obviously help to reduce both encoder and decoder runtime of OBMC. The third aspect changes the blending procedure, to enable parallel processing

The worst case memory bandwidth increase of OBMC is still approx. 2.5X

[JVET-K0474](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4004) Crosscheck of JVET-K0258: OBMC complexity reduction and parallel blending [J. Ye, X. Xu (Tencent)] [late]

[JVET-K0422](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3943) Cross-check of JVET-K0258: CE10-related: OBMC complexity reduction and parallel blending [R.-L. Liao, C. S. Lim (Panasonic)] [late]

[JVET-K0259](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3769) CE10-related: OBMC bandwidth reduction and line buffer reduction [Z.-Y. Lin, T.-D. Chuang, C.-Y. Chen, Y.-W. Huang, S.-M. Lei (MediaTek)]

This contribution presents proposed methods to reduce overlapped block motion compensation (OBMC) memory bandwidth and line buffer. In order to remove the additional memory bandwidth required by OBMC, padding is applied to additional reference picture samples at the right-most w’ columns and bottom-most h’ rows. They are not fetched and replaced with generated samples by replicating the right-most column and the bottom-most row of the original fetched reference picture samples for fractional sample accuracy motion compensation (MC), where w’ and h’ are the width and height of OBMC region, respectively. For blocks coded with sub-block mode, the same padding method can be applied to reduce the required bandwidth. As for line buffer usage, the number of blended lines for above block boundary is reduced from 4 to 2, if current block is at coding tree unit (CTU) row boundary. Applying padding for OBMC at coding unit (CU) boundary suffers 0.03%, 0.06%, 0.01%, and 0.03% BD-rates for VTM-1.0-RA, VTM-1.0-LB, BMS-1.0-RA, and BMS-1.0-LB, respectively. Applying padding for OBMC at CU boundary and reducing the number of blended lines at CTU row boundary induce 0.05%, 0.07%, 0.03%, and 0.02% BD-rates for VTM-1.0-RA, VTM-1.0-LB, BMS-1.0-RA, and BMS-1.0-LB, respectively. Applying padding for OBMC at CU boundary and subblock boundary and reducing the number of blended lines at CTU row boundary introduce 0.05%, 0.07%, 0.04%, and 0.06% BD-rates for VTM-1.0-RA, VTM-1.0-LB, BMS-1.0-RA, and BMS-1.0-LB, respectively.

Due to padding, worst case memory bandwidth is reduced to approx. 2x. Loss compared to “normal” OBMC is <0.1%.

Investigate in CE together with K0258.

[**JVET-K0537**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4067) **Cross check of CE10-related: OBMC bandwidth reduction and line buffer reduction (K0259) [M. Siekmann (HHI)] [late] [miss]**

[JVET-K0270](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3781) CE10-related: Diagonal motion partitions on top of MTT block structure [Y. Ahn, D. Sim (Digital Insights)]

[JVET-K0526](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4056) Cross-check of JVET-K0270 (CE10-related: Diagonal motion partitions on top of MTT block structure) [T. Na, J. Lim (SK telecom), J. Shin (PIXTREE)] [late] [miss]

In JVET-H0087, diagonal motion partitions (DMPs) were proposed on top of quadtree plus binary tree (QTBT) block structure. In this contribution, the same method for inter prediction is proposed on top of multi-type tree (MTT) block structure. In the proposed partitioning method, a coding unit (CU) is split into two diagonal motion partitions. The proposed method includes only two diagonal directions, but it can represent various arbitrary partitions on top of MTT block structure. The proposed DMPs can achieve 1.28% and 1.75% BD-rate reduction over VTM-1.1 for random access and low-delay B configurations, respectively.

The slide deck showed additional information not in the word document – should be uploaded.

Similar or better performance than CE10 contributions on diagonal partitioning, however also significant increase in encoder runtime – that should be decreased.

In the presentation, preliminary results were presented that usage with uni prediction still preserves major part of the gain.

Blending of the two diagonal partitions was used, but that does not increase memory bandwidth.

Study in CE together with other geom. part. approaches.

[JVET-K0485](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4015) CE9-related: A simplified bi-directional optical flow (BIO) design based on the combination of CE9.5.2 test 1 and CE9.5.3 [X. Xiu, Y. He, Y. Ye (InterDigital), C.-Y. Chen, C.-Y. Lai, Y.-W. Huang, S.-M. Lei (MediaTek)] [late]

[JVET-K0531](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4061) CE10-related: Combined test of CE10.1.4 and CE10.1.8 [M.-S. Chiang, C.-W. Hsu, Y.-W. Huang, S.-M. Lei (MediaTek)] [late]

In this contribution, the combined test regarding multi-hypothesis prediction in JVET-J0018 and JVET-J0014 are proposed. In JVET-J0018, multi-hypothesis prediction is applied to advanced motion vector prediction (AMVP) mode, skip or merge mode, and intra mode, which are tested in CE10.1.1, CE10.1.2, CE10.1.3, respectively. CE10.1.4 is the combined test of CE10.1.1, CE10.1.2, and CE10.1.3. In JVET-J0014, multi-hypothesis prediction is applied to merge mode, which is tested in CE10.1.5 to CE10.1.8 with different parameter settings. In this contribution, combined results of CE10.1.4 and CE10.1.8 is proposed. It is reported that, compared to VTM-1.0, this proposal achieves -2.45% and xxx% luma BD-rates for RA and LB, respectively, with 39% and xx% encoding time increases and xx% and xx% decoding time increases. Compared to BMS-1.0, this proposal achieves xxx% and xxx% luma BD-rates for RA and LB, respectively, with xx% and xx% encoding time increases and xx% and xx% decoding time increases.

Revisit: Report results on BMS, if possible before end of meeting

Worst case complexity is the same as the individual tools, as max. number of hypotheses stays the same. Also worst case memory bandwidth is not increased compared to the individual tools.

Further study in CE.

[**JVET-K0543**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4073) **Cross-check of JVET-K0531: CE10-related: Combined tests of CE10.1.4 and CE10.1.8 [X. Xiu (InterDigital)] [late] [miss]**

## CE11 related – Composite reference pictures (3)

Contributions in this category were discussed Friday 13 July 1940–2010 (Track B chaired by JRO).

[JVET-K0157](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3665) CE11: HEVC-like encoder only solution for composite reference picture [W. Li, X. Zheng (DJI)]

This contribution provide a HEVC-like encoder only solution for composite reference that is evaluated in CE11. Implementation details and test results followed by common test condition are provided in the document. Simulations show that the proposed technique can achieve -2.46% and -1.46% coding gain over VTM1.0 and BMS1.0 at Lowdelay B Main10 configuration (LDB) with around 20% encoding time increase.

Decision(SW): Adopted – see further notes under CE11.

[JVET-K0159](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3667) CE11-related: Hardware friendly composite reference picture solution [W. Li, X. Zheng (DJI)] [late]

This contribution proposes a composite reference design. It can allow composite reference update at block level when a CTU’s coding is finished, which is said to show benefits on hardware design. It is said that the proposed method can achieve almost same coding performance of composite reference and get lower encoding runtime increase.

Relative to the method in CE11, the encoding run time is not decreased, and results are almost identical. It is claimed to be beneficial for encoder hardware implementation, but there is no need to investigate this in CE11, which should target investigating improved compression benefit of composite reference pictures.

[JVET-K0447](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3972) CE11 Related Work: Long-term Reference Simulated Implementation [C. Ma, D. Liu, Y. Li, F. Wu (USTC)] [late]

This contribution reports simulation results of long-term reference. Compared with BMS1.0, simulation results show that this tool achieves 2.95% BD rate reduction in LDB configuration. Compared with VTM1.0, simulation results show that this tool achieves 2.43% BD rate reduction in LDB configuration.

The approach is to define the first I picture as long term reference picture.

The gain comes again from sequences with static background.

The approach also provides gain for RA test cases. The IDR picture of each IDR period is defined as long term reference in this case. However, the contribution does not show results for the entire test set in the IDR case, only selected sequences which provide coding gain.

For LDB, the results are worse than those of CE11.

No superior methods compared to CE11, neither here nor in CE11 itself. Discontinue CE11– potentially take up again when evidence of benefit is shown for sequences with non-static background, which might require (normative) decoder-side tools.

## CE12 related – Mapping for HDR content (1)

Contributions in this category were discussed XXday XX July XXXX–XXXX (chaired by XXX).

[JVET-K0309](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3823) In-loop Reshaping for SDR Video [F. Pu, T. Lu, P. Yin, W. Husak, S. McCarthy, T. Chen (Dolby)]

(following notes by JRO when the document was discussed in context of BoG review Tue 17 track B)

Performs coding in a re-shaped domain, requires operating reshaper and inverse reshaper in the loop. Provides gain of 2% bit rate reduction, also for luma.

Reshaping function is computed for the IDR picture and then used over the whole IDR period.

Should be investigated if it has impact on visual quality, and what is the interrelationship with other tools of VTM and BMS (e.g. CCLM)

To be investigated in CE12 (rename as “mapping functions”

[JVET-K0468](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3998) CE12-related: In-loop chroma refinement [[E. François](mailto:edouard.francois@technicolor.com), C. Chevance, F. Hiron (Technicolor)] [late]

Discussed in BoG – applied for SDR and HDR as well. Test in CE12 along with K0309.

## CE13 related – Projection formats (1)

Contributions in this category were discussed XXday XX July XXXX–XXXX (chaired by XXX).

[JVET-K0332](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3846) CE13-related: Adaptive frame packing on top of CMP, MCP, and PAU [P. Hanhart, Y. He, Y. Ye (InterDigital)]

[JVET-K0522](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4052) Crosscheck of JVET-K0332: CE13-related: Adaptive frame packing on top of CMP and MCP [P. Wang (MediaTek)] [late]

## NN technology related (5)

Note: JVET-K0266 also relates to NN technology.

Contributions in this category were discussed Monday 16 July track B 1540–XXXX (chaired by JRO).

[JVET-K0158](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3666) AHG9: Separable Convolutional Neural Network Filter with Squeeze-and-Excitation block [T. Hashimoto, E. Sasaki, T. Ikai (Sharp)]

This contribution presents a Separable Convolutional Neural Network Filter with Squeeze-and-Excitation block (SESCNNF), which has fewer parameters than the network structure proposed in JVET-I0022. The current BMS software has multiple filters such as deblocking filter (DF), sample adaptive offset (SAO) and adaptive loop filter (ALF). In this contribution, we replace these three filters with SESCNNF, and SESCNNF shows 3.08%, 4.62%, and 5.73% gain on Y, Cb, and Cr average in BMS IO configuration.

Results are with AI

Comparison is against normal BMS anchor. Test is with other loop filters turned off, and CNNLF always enabled. Decoding time is >500x larger than anchor. Various configurations are shown, e.g. the decoding time is reduced to about 140-150x against anchor, the luma gain goes down to 2%.

Network is trained for different QP values (from CTC). Network does not know the current QP.

Sometimes the CNN generates artifacts. Might be better to disable the filter when it does not work

C++ implementation, floating point.

[JVET-K0443](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3968) Crosscheck of JVET-K0158: AHG9: Separable Convolutional Neural Network Filter with Squeeze-and-Excitation block [X. Song, L. Wang (Hikvision)] [late]

[JVET-K0222](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3732) AHG9: Convolution neural network loop filter [Y.-L. Hsiao, T.-D. Chuang, C.-Y. Chen, C.-W. Hsu, Y.-W. Huang, S.-M. Lei (MediaTek)]

This document presents three modifications of convolution neural network loop filter (CNNLF) introduced in JVET-J0018. The first modification is to train CNNLF parameters for each random access segment (RAS, roughly 1 sec in the conducted tests) of the sequence and signal CNN parameters at I-slice of each RAS, instead of training CNNLF parameters for the entire sequence and signaling CNNLF parameters in picture parameter set (PPS). The second modification is to simplify the CNNLF network from “eight layers with reconstructed samples, prediction samples, and residual samples as the input signals” to “four layers with reconstructed sample as the input signal” only. The third modification is that only those pictures with temporal ID equal to 0 or 1 are used in the training process to derive CNNLF parameters. Compared with VTM-1.0-RA, the proposed CNNLF achieves -2.57%, -18.52%, and -18.29% BD-rates for Y, U, and V, respectively, with 89% decoding time increase. Compared to BMS-1.0-RA, the proposed CNNLF achieves -0.88%, -13.76%, and -13.19% BD-rates for Y, U, and V, respectively, with 29% decoding time increase. Since the chroma BD-rate savings are much higher than the luma BD-rate savings, increasing chroma QP offset by 1 for both Cb and Cr is tested, and the BD-rate results are reported as follows: -3.67%, -10.10%, and -9.72% for Y, U, and V, respectively, for VTM-1.0-RA; -1.96%, -4.00%, and -3.48% for Y, U, and V, respectively, for BMS-1.0-RA. Further research on complexity reduction and training enhancement for improving coding efficiency is suggested. It is claimed that CNN is a promising research direction for further study.

Network is trained for each random access segment of a sequence, network parameters are sent in parameter set. One CNN for Y, another for Cb/Cr. Filter sizes in the 4 layers are 1x1/3x3/1x1/3x3; represented by 6 bit integer, 16/16/8/4 filters in the layers, ReLU after each layer. Output of network is added to original decoder output.

Time for training not included in encoder time (would anyway not be doable in realtime applications)

Chroma QP offset is increased by 1 relative to CTC.

Higher gain for higher resolution (likely due to the fact that the number of parameters is relatively less for high res)

Operated as last loop filter with block-level on/off flag.

Subjective quality? Looks better for Camp Fire (due to higher gain for chroma), sometimes for other sequences.

Theoretically, it could be done for low delay as well, when either using pre-trained networks, or training with previous segment; would however difficult with current technology in real-time.

Several experts expressed opinion this is very interesting. First time that decoding time on CPU comes t a realistic value.

[JVET-K0391](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3910) AHG9: Dense Residual Convolutional Neural Network based In-Loop Filter [Y. Wang, Z. Chen, Y. Li (Wuhan University), L. Zhao, S. Liu, X. Li (Tencent)] [late]

This contribution provides a dense residual convolutional network based in-loop filter (DRNLF) for VVC. In-loop filters, such as DF (deblocking filter), sample adaptive offset (SAO), are employed in VTM for suppressing compression artifacts, which contributes to coding performance improvement. In this contribution, the proposed DRNLF is introduced as an additional filter before SAO. Simulation results report -5.75%, -17.56%, -18.74% BD-rate savings for luma, and both chroma components compared with VTM1.1 under AI configuration, and -6.11%, -15.85%, -14.36% for RA configuration, and -6.05%, -11.53%, -12.30% for LDB configuration, and -7.14%, -12.16%, -12.72% for LDP configuration.

Operated between DBF and SAO. The decoding times reported are 17x for AI, 38x for RA, 37x for LDB (run on GPU). When run on CPU, AI decoding time increases to >800x.

Higher gain for low resolution video

Switchable on block level. Would be interesting to see how it performs when ALF is still used

Subjective quality? Not known.

Generally, the bit rate savings reported are higher now than reported in previous meeting. However, further study (AHG) necessary before we could define a CE. In the next meeting, also subjective viewing should be done to better understand the impact of CNN in comparison to other loop filters.

[JVET-K0444](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3969) Crosscheck of JVET-K0391: AHG9: Dense Residual Convolutional Neural Network based In-Loop Filter [X. Song, L. Wang (Hikvision)] [late]

## 360° Video related (5)

Contributions in this category were discussed XXday XX July XXXX–XXXX (chaired by XXX).

[JVET-K0141](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3647) AHG8: 360°-based inter/intra prediction for cubemap projection [C.-H. Shih, J.-L. Lin, H.-C. Lin, S.-K. Chang, C.-C. Ju (MediaTek)]

[JVET-K0142](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3648) AHG8: 360°-based in-loop filters for cubemap projection [S.-Y. Lin, L. Liu, C.-H. Shih, J.-L. Lin, H.-C. Lin, S.-K. Chang, C.-C. Ju (MediaTek)]

[JVET-K0466](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3995) Cross-check of JVET-K0142: AHG8: 360°-based in-loop filters for cubemap projection [P. Hanhart (InterDigital)] [late]

[JVET-K0183](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3692) AHG8: Face boundary filtering for 360° video [Xuchang Huangfu, Yule Sun, Bin Wang, Lu Yu (Zhejiang Univ.)] [late]

[JVET-K0333](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3847) AHG8: Horizontal geometry padding for PERP [P. Hanhart, Y. He, Y. Ye (InterDigital)]

[JVET-K0404](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3924) AHG8: Selective In-loop filtering for 360 Video Compression [C. Pujara, S.N. Akula, A. Singh, R. Narayana, W. Choi (Samsung)] [late]

## Extended colour volume related (0)

Contributions in this category were discussed XXday XX July XXXX–XXXX (chaired by XXX).

## HL syntax (6)

Contributions in this category were discussed Sunday 15 July 1700–1800 (chaired by GJS and JRO).

[JVET-K0155](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3663) AHG12: Flexible Tile Partitioning [Y. Yasugi, T. Ikai (Sharp)]

This contribution was discussed Sunday 1710 (chaired by GJS and JRO).

This contribution proposes a tile functionality that allows to split pictures into flexible partitioning tile, where the width or height of the unit of tile can be any multiplies of 4 (the minimum CU size), 8, 16, 32 and 64, i.e. smaller than a CTU.

In the proposed tile design, pictures would be split into constant-size CTUs as the conventional tile while the size of the right most and bottom most CTUs in tile boundary can be different from the constant CTU size. This flexible feature is asserted to provide better load balancing since all tile can be almost the same size in uniform spacing mode. It is also asserted that this feature is useful for 360 video sequences or frame packing sequences since the corresponding tile can fit the arbitrary face size.

The experimental results with this feature (unit=32) reportedly show that luma BD-rate coding losses on average are 0.66 % for All Intra and 0.99 % for Random Access under the common test condition (CTC) for SDR sequences.

The experimental results without this feature (i.e. HEVC like tile) reportedly show that luma BD-rate coding losses on average are 0.64 % for All Intra and 0.97 % for Random Access under the common test condition (CTC) for SDR sequences. In some cases, there was even some gain observed.

It was commented that the coding efficiency impact would be affected by the boundary handling in the design.

The contributor suggested that this is relevant to AHG12 (on parallelism).

2x2 composited images were shown as an example, where the tile boundaries could be set to align to the region boundaries.

Further study was encouraged.

[JVET-K0408](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3928) Cross-check of JVET-K0155: AHG12: Flexible Tile Partitioning [A. Wieckowski (HHI)] [late]

[JVET-K0260](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3770) Flexible Tiles [R. Sjöberg, M. Damghanian, M. Pettersson, J. Enhorn (Ericsson)]

This contribution was discussed Sunday 1725 (chaired by GJS and JRO).

This contribution proposes to include tiles into VVC. The tile sizes are proposed to be signaled individually, either by copying the tile size from the previous tile size in decoding order or by one tile width and one tile height code word. This is reported to enable tile partition structures that are not restricted by the HEVC rule that tile boundaries must span across the entire picture.

The following HEVC tile properties are included in the proposal:

* That CABAC is initialized for each tile, flushed after each tile and that the tiles are byte aligned
* That tiles break predictions as in HEVC
* That the tile structure is specified in the PPS
* That tile pointers are mandatory for all tiles except the first in a slice
* That either all CTUs in a tile belongs to the same slice or all CTUs in a slice belong to the same tile
* That the initial QP of a tile is set as in HEVC
* That tiles are CTU aligned
* That tiles can only have rectangular shape

The following tile properties different to HEVC tiles are also proposed:

* That tile boundaries are not required to span across the entire picture
* That the number of tiles in a picture is signaled in the PPS
* That the tile sizes are specified in subtile units to reduce the signaling bit cost
* That tiles are specified individually in subtile raster scan order by their individual tile size either by copying the tile size from the previous tile size in decoding order or by one tile width and one tile height UVLC code word each

The contribution presents experimental results for an OMAF 360 video partitioning example using HM 16.18. Compared to a realization using HEVC tiles and slices, using tiles with the proposed method, an average BDR of -1.1% was reported under RA configuration using the three 8-bit VVC CfP 360 video sequences.

The proponent said that both tiles and slices should be supported.

A particular syntax is proposed, with a prediction of tile sizes and a signalling of a granularity of the tile boundaries.

It is proposed that there be an established maximum number of tiles per picture.

In HEVC, tiles have a minimum width of 256, and there may also be a limit on minimum height.

It was commented that the ability for the tile structure to change from picture to picture may cause problems for parallel decoders.

It was commented that partitioning the picture for processor load allocation may be different from partitioning for other purposes.

It was commented that if we really want to think about tiles as a parallelization tool, this could get into profile/level constraints discussions that may be premature.

[JVET-K0300](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3814) Design goals for tiles [M. M. Hannuksela, A. Zare, M. Homayouni, R. Ghaznavi-Youvalari, A. Aminlou (Nokia)]

This contribution proposes that the VVC tile design should enable

* Encoding of motion-constrained tile sets (MCTSs) that are more efficient than HEVC MCTSs in terms of rate-distortion penalty;
* Avoiding visible MCTS boundaries with as small processing cost as possible;
* Intra block copy across tiles for enabling prediction from one constituent frame to another for frame-packed stereoscopic video, provided that intra block copy is adopted as a tool in VVC;
* Extracting VCL NAL units of a subset of MCTSs from one VVC bitstream and reposition them to another VVC bitstream without VCL NAL unit modifications.

The proposed design goals are asserted to make VVC tiles suited for viewport-dependent 360° streaming.

These design goals are proposed to be used in evaluating merits of technical contributions and to be included as mandates of an appropriate JVET ad-hoc group.

It was commented that a more fair comparison for some of the illustrated cases would use tiles that are not motion-constrained tiles.

The proposal also suggested supporting MCTS reordering and rewriting functionality.

This was further discussed Tuesday 1250 (GJS & JRO). It was agreed to establish an AHG on segmentation of a picture into coded regions, to investigate tiles, slices, etc., and what they can be used for and what additional functionalities for such regions may be beneficial beyond what has been done in past standards.

[JVET-K0325](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3839) On High Level Syntax Starting Point [S. Deshpande, B. Choi (Sharp)]

This was reviewed in JVET plenary Tuesday 1200 (GJS & JRO).

A high level syntax starting point is proposed for VVC that employs a NAL unit structure, a sequence parameter set, a picture parameter set, and a slice header. Initial NAL unit header and NAL unit types are also proposed.

It was commented that this appears to be an appropriate basic “trimmed down” approach. It includes the concepts of IRAP-vs-nonIRAP, NAL units, SPS, PPS, SEI, end of sequence, end of bitstream, and “slice”. Specification of access unit delimiters and filler data are not included.

It was commented that the traditional concept of a slice may not be needed, although it is likely that we would have some collection of CTUs that is not a whole picture (e.g., tiles). Using the term “slice” does not necessarily imply a traditional slice.

Decision: Adopted.

[JVET-K0403](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3923) A Flexible Syntax Framework for VVC [S. Deshpande, F. Bossen, A. Segall (Sharp)] [late]

This was reviewed in JVET plenary Tuesday 1230 (GJS & JRO).

This document proposes a framework for supporting a syntax structure in VVC. The core motivation for the proposal is to create syntax groups that may be skipped without parsing. This is asserted to allow decoders to extract syntax of interest without having to parse other syntax element groups. Here, the emphasis is on the signalling of high-level syntax in parameter sets, though it is asserted that the framework could be applicable to other signalling uses in the VVC standard.

The scheme is equivalent to the current method of carrying multiple SEI messages in one NAL unit.

It was commented that such a scheme could be only partially used - e.g., only for extension parameters.

It was commented that such schemes had been previously considered at some point - e.g., for groups of extension data.

A contributor said this would allow encoders to send the syntax they care about first, without needing to send other parameters earlier.

It was commented that if the order of the syntax is allowed to be changed, dependencies between syntax elements would need to be carefully considered, as “race conditions” would be possible.

It was asked whether this is needed at this stage, and suggested to just keep it in mind for further development.

This adds some overhead for type codes and length values.

It was commented that the lack of such a scheme has not really been a problem thus far.

Other than VUI, the syntax in the parameters sets is (at least generally) essential to enabling decoding of the bitstream (or for capability identification, in the case of the profile/tier/level syntax).

It was commented that the granularity of the synax has an effect on the usefulness - e.g., it doesn’t seem appropriate to wrap a single flag in this.

Several participants expressed interest in the scheme, and it should be considered in further work, but seems not necessary to use at this stage.

## Resilient intra refresh (1)

Contributions in this category were discussed Tuesday 17 July in Track B 0930–XXXX (chaired by JRO).

[JVET-K0212](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3722) Improved Cyclic Intra Refresh [J.-M. Thiesse, D. Nicholson] [late]

This contribution proposes a technique to improve the use of Cyclic Intra Refresh, a technique that is commonly used for real-time implementation of HEVC video encoders. Intra Refresh provides several desirable features, such as bit-rate smoothing, maintaining Intra data bit-rate under control, low delay and natural error resilience limiting temporal error propagation along frames. For video encoder’s manufacturers this feature is considered as mandatory for real-life applications.

In VVC Requirements, as contained in VCEG-BD03 and WG11/ N17074, low latency and error resilience requirements are present but have not been addressed by the responses to the CfP. One big factor of latency apart picture ordering and related dependencies, is the importance of Intra picture data amongst other, with an increasing Intra picture data size versus Inter picture data size ratio at each video codec generation. Cyclic Intra Refresh can be used for fulfilling these requirements.

But this feature, as applied by encoder manufacturer in previous standards, implies some encoding restrictions, limiting the coding efficiency, as well as Random Access capability. The proposed technique through appropriate signaling enables to leverage these limitations, allowing less coding restrictions. In this document results provided show an improvement toward classic use of Cyclic Intra Refresh together with usual LD configuration as a reference. Authors recommend considering the proposed technique for a potential inclusion in VVC future standard with further study to be conducted in an appropriate Core Experiment.

Some support is is expressed for studying the aspect of periodic intra refresh, which is commonly used in practice but there are more approaches for resolving it.

Benefit of vertical refresh regions not that obvious (except for ultra low delay, as it has less bit rate fluctuation over the frame). Horizontal refresh could be done with slices. However, independent slices cause some bit rate overhead.

Non-normative solution (just enforcing intra locally) causes up to 25% BR loss (maximum for class E). By introducing normative signalling of the vertical refresh area 2.5% is recovered. For other classes, the loss of non-normative solution is <10%, of which approx. 3.5% is recovered.

These results however reflect the case where the whole sequence is decoded from the beginning. As periodic intra refresh is used for random access, it should better be tested what happens if te decoding of a bitstreams starts at some random picture position. Also SPS?/PPS should be transmitted with every picture and counted in the bit rate, as well as the propagation errors should be counted in terms of quality.

Further study (AHG on low latency coding, J.-M. Thiesse, A. Duenas, A. Tourapis)

[JVET-K0560](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4090) Cross-check of JVET-K0212: Improved Cyclic Intra Refresh [E. Mora (Ateme)] [late]

X-check

## Palette mode (2)

Contributions in this category were discussed Tuesday 17 July in Track B 0900–0930 (chaired by JRO).

[JVET-K0411](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3931) AHG11: Palette mode [Y.-C. Sun, J. An, J. Lou (Alibaba)] [late]

This document proposes a simplified palette mode on top of VTM1.1. For Class F sequences, the results show -2.6%/-1.8%/-0.9% BD-rate luma gain in AI/RA/LDB configuration; for 4:2:0 TGM sequences in CE8, the results show -13.7%/-8.9%/-7.0% BD-rate luma gain in AI/RA/LDB configuration.

Palette mode is not as in HEVC, but rather similar to an earlier version (SCM2)

Applicable to 4:4:4 and 4:2:0

Enabled at CU level (which is now rectangular whereas it was always square in case of HEVC

Small losses for classes A through E

Investigate in CE15:

* Interrelationship with CPR
* Also investigate the HEVC palette mode
* Study the complexity impact of the two palette variants and IBC (V. Seregin offers to help in porting the SCC palette mode to VTM).

[JVET-K0475](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4005) Crosscheck of JVET-K0411: AHG11: Palette mode [J. Ye, S. Liu (Tencent)] [late]

X-check

# Complexity analysis and reduction (8)

Contributions in this category were discussed XXday XX July XXXX–XXXX (chaired by XXX).

[JVET-K0057](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3558) Reduced Memory Storage for Collocated Picture [Y. Yu, S. Hong, K. Panusopone, L. Wang (Arris)]

The contribution is available for study but the presenter was not available.

[JVET-K0086](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3589) Non-CE: Temporal motion vector data storage reduction [H. Jang, J. Lim, J. Nam, S. Kim (LGE)]

This contribution reports coding performance impact with respect to the compression unit size of temporal motion vector storage base on spatial candidates located in compression unit. It is observed that the proposed compression method shows 0.1% BD-rate coding gain from 16x16 compression unit on VTM Random access configuration. The proposed method achieved that 0.1% in partial result and 0.28% BD-rate loss with a factor of 4 and 16 memory size reduction respectively in BMS Random access configuration.

Was reviewed in track B Monday afternoon

Selection of the vector for temporal storage is done based on spatial candidate.

Results are provided for the cases of 16x16 (as HEVC), 32x32 and 64x64 memory compression.

(update table from new version)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Randomaccess main10 | | | | | | | over VTM-1.0 | | | | | | | | OverBMS-1.0 | | | | | | |
| Y | | U | | V | | EncT | | | DecT | | Y | | U | | | V | | EncT | | DecT | |
| 16x16 | -0.09% | | -0.09% | | -0.08% | | | 100% | | 105% | | 0.28% | | 0.25% | | | 0.20% | | 101% | | 102% |
| 32x32 | 0.11% | | 0.08% | | 0.13% | | | 101% | | 105% | | 0.59% | | 0.48% | | | 0.51% | | 101% | | 101% |
| 64X64 | 0.56% | | 0.45% | | 0.53% | | | 102% | | 105% | | 0.98% | | 0.86% | | | 0.90% | | 102% | | 100% |

The complexity in number of operations is almost negligible. The gain in VTM (approx. 0.1%) shown for the case of 16x16 indicates that the method has benefit.

Further study (CE4) in combination with the new 8x8 compression scheme of VTM.

[JVET-K0530](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4060) Crosscheck of JVET-K0086: Temporal motion vector data storage reduction [Y. Han, W.-J. Chien (Qualcomm)] [late]

[JVET-K0106](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3610) Energy Assessment for Video Decoding [C. Herglotz (FAU)] [late]

This was discussed Sunday 15 July 1645 (GJS & JRO).

This proposal describes a contribution on the assessment of processing energy for video decoding. It is stated that the processing energy is a relevant metric especially for portable devices like smartphones or tablet PCs, because small energy requirements help to increase the operating time until the battery is empty. It is proposed to use energy measurements to assess the energy efficiency of the decoding process.

In this proposal, a general method for energy measurements is presented that can reportedly be used for various video coding platforms. The results can be used to evaluate a decoder in terms of processing energy. Finally, it is proposed to use the Bjøntegaard-Delta Metric to assess the energy efficiency of the complete decoding process.

Study and experimentation with the method was encouraged.

Slide deck to be provided.

[JVET-K0107](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3611) Energy Modeling for Video Decoding [C. Herglotz, M. Kränzler, A. Kaup (FAU)] [late]

This was discussed Sunday 15 July 1655 (chaired by GJS & JRO).

In this contribution, a recently published energy model is presented which can be used for video decoding energy estimations. It is reported that the model is able to estimate the energy of both the processor and the memory access. Furthermore, it is stated that using sophisticated training methods, the trained model values can be interpreted as coding tool specific energies, and that the model can be used for decoder energy optimization.

The model is based on the amount of usage of a technical element and the energy used when that element is exercised. Further study was encouraged.

Slide deck to be provided.

[JVET-K0108](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3612) Decoding-Energy-Rate-Distortion Optimization [C. Herglotz, M. Kränzler, A. Kaup (FAU)] [late]

This was discussed Sunday 15 July 1505 (chaired by GJS & JRO).

This contribution proposes to include “decoding-energy-rate-distortion optimization” (DERDO) into the encoder reference software. It is reported that DERDO is an extension of RDO which can be used to control and minimize the decoding energy on the receiver side which is interesting for portable, battery driven devices. As the complexity of the next coding standard is expected to increase significantly, this tool may help in decreasing the complexity requirements of decoders. In this work, the potential savings are analyzed for intra only coding and reach 5.74% of energy savings at the expense of a bitrate increase of 7.57% or 1.1% energy savings at 0.13% bitrate increase. The authors claim that due to the high amount of new coding tools proposed for VVC, energy savings of a higher magnitude at lower rate increases can be expected for inter prediction coding tools.

It was asked whether the energy parameters that were used were provided. The presenter said they would provide these in a revision of the contribution. Further study was encouraged.

It was noted that the “green MPEG” initiative seems related.

Slide deck to be provided.

[JVET-K0451](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3976) AHG5: How to use the software to evaluate memory bandwidth [R. Hashimoto, S. Mochizuki] [late]

This contribution was provided for information on how to use memory bandwidth analysis tool. It is available for study. Detailed presentation was not requested.

[JVET-K0452](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3977) AHG5: Proposal of template for comparing memory bandwidth [R. Hashimoto, S. Mochizuki] [late]

This was discussed Sunday 15 July 1630 (chaired by GJS & JRO).

This contribution provides a template to compare memory bandwidth measured with the VTM and BMS software. This template can be used to check and compare memory bandwidth in the test model decoders.

See the prior document JVET-J0090. The software is being integrated into the VTM and BMS software packages.

It was commented that it could be desirable to request the use of the tool in AHG13 for tool on/off testing.

# Encoder optimization (3)

Contributions in this category were discussed XXday XX July XXXX–XXXX (chaired by XXX).

Allocated to BoG (coord by F. Bossen)

[JVET-K0390](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3909) Rate Control for VVC [Y. Li, Z. Chen (Wuhan University), X. Li, S. Liu (Tencent)] [late]

[JVET-K0472](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4002) Crosscheck of JVET-K0390: Rate Control for VVC [Q. Yu, J. Zheng (HiSilicon)] [late]

[JVET-K0206](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3715) AHG10: Improved perceptually optimized QP adaptation and associated distortion measure [C. Helmrich, H. Schwarz, D. Marpe, T. Wiegand (HHI)]

# Metrics and evaluation criteria (0)

Contributions in this category were discussed XXday XX July XXXX–XXXX (chaired by XXX).

# Withdrawn (8)

JVET-K0077 Withdrawn

JVET-K0089 Withdrawn

JVET-K0091 Withdrawn

JVET-K0114 Withdrawn

JVET-K0128 Withdrawn

JVET-K0278 Withdrawn

JVET-K0296 Withdrawn

JVET-K0313 Withdrawn

JVET-K0322 Withdrawn

JVET-K0470 Withdrawn

JVET-K0471 Withdrawn

JVET-K0551 Withdrawn

JVET-K0557 Withdrawn

# Plenary meetings, joint Meetings, BoG Reports, and Summary of Actions Taken

## Plenary meeting Friday 13 July 1400

* Role of the BMS
  + Holding extra potential features we aren’t so sure about yet
  + Needs to have significant gain over the VTM
  + Sometimes unveils whether gains are independent
  + Can be a common basis for CE tests of modified versions of features
  + Features that don’t yet have a clear or properly worked out design
  + Keep it runable
* Separate tree for (intra) chroma under consideration; see K0230, CE 1.5.1.1, CE1.5.2.5? Not in BMS yet. Put K0230 in BMS?
* Spec text

WD / VTM

* Increasing the upper bound on QP by 12 (no effect on CTC results) (K0251)
* PDPC (from K0063). AI 1.0%, RA 0.5%
* Intra 67 modes with 6 MPM and truncated binarization of non-MPM modes; otherwise per 3.2.3 (K0368), pending confirmation of mode coding after some experiment result (LGE / Huawei / Qualcomm were to test). AI 1.3%, RA 0.6% Remark: Consider non-normative speed-up
* CCLM 1.2%/9.0%/8.0% for Y/Cb/Cr in AI, 0.8%/10%/9.2% for Y/Cb/Cr RA K0190
* AMT, both intra and inter, each controlled by an SPS flag, AI 3.3%, RA 2.0%, LB 1.3%). It was suggested to disable inter AMT for CTC (penalty 0.5%, only a non-normative issue - see section 6.6).

6.5% in AI for luma, 3.5% for RA for luma, significantly more for chroma

[Consider increasing QPs of CTC or adding more QPs or spacing them 7 apart]

BMS

* CPR as per HEVC SCC [Consider adding more SCC into CTC Class F and making mandatory, see also K0294 - JB to coordinate]
  + CTC: 1.3% for AI, 0.5% for RA
  + Class F: 21% for AI, 16% for RA
  + SCC TGM 1080p: 54% for AI, 39% for RA

Roughly what that is likely to provide, relative to the prior VTM: 7.8% for AI, 4.0% for RA

Other

* BoG on partitioning
* BoG on boundary handling
* BoG on 360° video
* Revisit division by block size in CE3 for DC intra prediction
* Revisit trellis quantization in CE7
* Revisit whether to keep secondary transforms in the BMS and which ones in CE6

Track B

* Moved AMVR from BMS into VTM ~1.8% for RA
* Removing some partitioning restrictions
* BoG on ALF
* Deblocking viewing

Roughly what that is likely to provide, together with the Track A actions, relative to the prior VTM: 7.8% for AI, 5.8% for RA

Planned activities:

* BoG on boundary handling in foyer left
* BoG on intra prediction and intra mode coding Geert
* BoG on inter prediction and motion vector coding Haitao
* Viewing on deblocking Saturday
* 360° viewing Friday pm
* BoG on 360° video

## Plenary meeting Sunday 15 July 1430

General:

* QPs for CTC (add another or space apart by 7)? - open
* CPR in CTC - yes (as part of the BMS).
* Text is to be reviewed during the meeting
* BMS status

Track A:

* Separate tree for luma and chroma in intra slices Coding eff benefit
* Implicit split to 64x64 for intra slice
* Prohibit ternary split of edges longer than 64
  + Intra and inter
  + What about 32x128? - See later contribution K0556 - to be further studied.
* Dependent quantization Coding eff benefit
* Sign data hiding when dependent quantization not being used
* Modification of DC prediction mode for rectangular blocks to avoid division
* QT+BTT status
* Chroma QP increase when separate tree (non-normative CTC or software)
* Intra prediction with 67 modes (see MPM note below)
* Open at the time of this plenary discussion:
  + BoG on picture boundary handling
  + Wide angles
  + 3 MPM vs. 6 MPM est. ~0.2% difference for BMS RA, about 0.5% for VTM AI when encoding search is equalized (there is also a difference due to using context modeling in the 6 MPM scheme). Decision: Use 3 MPM for now.
* Track B:
  + VTM
    - GALF (4x4 class. based) without filter pred. (temporal/from default filters), 7x7 luma, 5x5 chroma (fixed), CTU level switching
    - ATMVP with 8x8 MV storage & simplifications, slice switching 4x4/8x8 subblock
    - AMVR
    - Deblocking “bug fix“ at large TU & 8x8 grid
    - Affine: New prediction & difference coding, fixed 4x4 sub-block, switchable 4/6 parameter model, bug fix affine merge (affine agreed to be moved to VTM in JVET plenary Sunday afternoon)
  + BMS
    - BIO with limited WC complexity
    - DMVR modifications for latency problem solving
    - Generalized Bipred
  + Software:
    - Some speedups e.g. affine
    - long-term reference mechanism, in combination with pic\_output\_flag = 0 (non CTC)
    - Cross-CTU opt. of SAO (non CTC)

Planning of remaining reviews (Tracks A/B) was performed.

## Joint meetings

## BoGs (5)

[JVET-K0521](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4051) BoG Report on ALF [L. Zhang]

Was presented Sat. 14th 1900 Track B (see notes under CE2.4)

[JVET-K0527](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4057) BoG report on 360° video [J. Boyce]

The BoG met July 13 in two sessions to review CE13 and other 360 Video related contributions identified in the AHG8 report. Informal subjective viewing was performed, and the observations from the viewing discussed in the BoG.

The BoG recommends the following:

* Add the MCP projection format to 360Lib. Should add configuration options for flexible padding, including number of padding pixels, and which locations to apply padding (just between face rows vs. around the face rows including the picture edges.) (Note: the acronym MCP should be modified, better hybrid of something)
* Use the MCP format with padding of 4 samples around face row with blending (PMCP) as in Test 5.2 for coding tool experiments in the CE and the CTC. Replace the ERP and CMP anchors with MCP. Will still ask AHG6 to provide PERP and CMP results for VTM and BMS.
* Define a 360 video Core Experiment be defined to study removing subjective artifacts at face boundaries, and comparing normative decoding tools with non-normative pre- and post-processing (such as padding, cropping, blending, post-filtering). CE coordinators: P. Hanhart and J.-L. Lin.
  + Include in CE: inter and intra prediction using spherical neighbors, deblocking filter disabling and using spherical neighbors, post-filtering, combinations with padding and with other coding tools. Different padding widths.

The BoG encourages study of the following:

* Signaling of projection format parameters, coding tools parameters, and post-filtering hints.
* GPU complexity impact of rendering (including possible post-processing) of projection formats.

The BoG asks the track for guidance on normative tool adoption for 360 video. What complexity impact could be acceptable to remove subjective artifacts that pre- and post-processing alone are unable to resolve?

For the CE on tools, it is necessary to specify a single projection format (“MCP” was agreed).

A reference method that prevents visibility of face boundaries (which is the main target of the CE) with non-normative elements such as geometry-corrected padding, blending, post-filtering needs to be defined. Such a method was also agreed in the BoG. It must however be further studied if a non-normative method can be further improved. (Note: In the informal viewing performed at this meeting it seemed that non-normative methods available so far could not fully resolve the boundary visibility problem).

Guidance is seeked which amount of normative tools is acceptable:

* Disabling tools at face boundaries might be simple to do, but concepts about how to specify that have not been proposed yet. (this applies to frame packed neighbours that are not spherical neighbours)
* Tools requiring access from somewhere else (this applies to spherical neighbors that are not frame packed neighboured) in the same picture or a reference pictures appear to be more complicated, and no real concepts of how to specify this have been shown yet; assessment of additional operations, memory accesses or additional buffering should be done for these cases.

Only when knowing the compression benefit compared to the non-normative case, as well the complexity impact, and the simplicity level of specifying it, further decisions can be made.

[JVET-K0528](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4058) BoG report on partitioning structures (CE1 SubCE1) [B. Bross]

See section 7.1

[**JVET-K0539**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4069) **BoG report on intra prediction and mode coding (CE3-related) [G. Van der Auwera] [late]**

[**JVET-K0541**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4071) **BoG report on common test conditions [J. Boyce]**

TBP

[**JVET-K0546**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4076) **BoG report on CE4 related contributions [(H. Yang)]**

Was presented Sun 15th 1220.

The BoG reviewed documents from the following categories:

* Affine motion compensation (12)
  + BMS affine bugfix
  + Motion compensation
  + Merge & AMVP
  + MVD coding
  + 4/6-param model switching
  + Triangle partition based affine MC
* Merge mode enhancement (14)
  + ATMVP modifications
  + Modified merge candidates
  + List construction
  + Buffer reduction
* Motion vector coding (1)
* Reference picture boundary padding (1)

For specific notes on documents reviewed in the BoG, see under CE4 related section.

The following aspects were recommended (see for adoptions/decisions in CE4 related section):

* Recommendations on BMS
  + BMS affine bugfix on inheriting 4-param affine model
  + BMS affine bugfix on CU size restriction for affine merge mode (w&h >= 8)
  + SIMD implementation into BMS affine
  + Simplified ATMVP
    - One fixed collocated picture is used to derive temporal motion information.
    - Slice level adaptive sub-block switching, 8x8 or 4x4.
    - Constrain the region from where ATMVP motion is derived to the collocated CTU plus one 4x4 block column outside the collocated CTU at the right hand side, the same region for HEVC TMVP.
* Recommended tests in next round of CE4
  + Affine MC: 6
  + Merge enhancement: 9
  + Motion vector coding: 1
  + Reference picture boundary padding: 1

[**JVET-K0547**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4077) **BoG report on complexity analysis of long distance merge candidates and combined merge candidates [X. Li]**

This contribution is the BoG report on complexity analysis of long distance merge candidates and combined merge candidates. The BoG meeting was held Saturday 14 July from 8:30pm to 10:20pm.

**Complexity Analysis of long distance merge candidates**

The analysis on line buffer is based on the following conditions

|  |
| --- |
| 1. MV information stored in the current CTU is regarded as local and is not counted in line buffer calculation |
| 1. The last MV row (4x4 luma pixel level) above the current CTU is considered as in the current architecture, and is not counted in line buffer calculation |
| 1. Left CTU is not counted in line buffer calculation |

The BoG had consensus on condition 1 & 2 while no consensus on 3.

It was agreed in the discussion in track B that this memory should be counted, but not as part of “line buffer”, it is definitely additional buffer that is needed.

The complexity analysis of long distance merge candidates is summarized in the attachment data file.

The data are included in the attached Excel sheet.

The analysis unveils that the proposals that perform more operations and have more additional buffer requirements provide better results.

For the next round of CE, a limit should be imposed on

* Max additional line buffer (may be 0?)
* Max number of potential additional candidates
* Max number of operations/comparisons/conditions in the pruning
* Max merge candidates in the list (list size shall be constant)

The same basically applies to the other category below, which already has much more limited complexity; in the next CE, all proposals suggesting modifications on merge should be compared against each other and with data on the complexity increase that they impose. H. Yang should take over the overall coordination with help by others for sub-CEs.

**Complexity Analysis of combined merge candidates**

There are two proposals studied in this category. The analysis is summarized in the following table.

|  |  |  |
| --- | --- | --- |
|  | size of merge list | Newly introduced complexity in worst case |
| [JVET-K0198](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3707) | 5 | 6 scaling, 6 x 2 comparisons on refIdx |
| JVET-K0245 (CE 4.2.8.c) | 10 | 6 x (4 comparisons on refIdx, 2 MV scaling, 4 average)  4+5+6+7+8+9=39 comparisons on MV candidates |

[**JVET-K0552**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4082) **BoG report on HDR-related contributions [A. Segall]**

Reviewed in track B Tues 17 July 1020 (chaired by JRO)

The BoG met July 15 to review CE12 and other HDR video related contributions and topics identified in the AHG-7 report.

The BoG recommends the following:

* reconducting the CE on HDR mapping, to further investigate the out-of-loop and in-loop mapping and the luma/chroma rate allocation and balance
* creating a CE for investigating mapping for SDR content (based on JVET-K0309, JVET-K0468)
* updating the HDR CTCs as follows:
  + For tests not requiring visual assessment (fixed QP case), consider the first 300 frames of the following HLG content: DayStreet, PeopleInShoppingCenter, SunsetBeach, the initial version of FlyingBirds sequence provided by NHK
  + For tests requiring visual assessment (fixed bitrate case), consider the full 600 frames of the following HLG content: DayStreet, PeopleInShoppingCenter, SunsetBeach
  + Mandating usage of HDRTools to report HDR metrics (including wPSNR)
  + Macro WCG\_EXT should be enabled
* reporting performance comparisons of VTM and BMS against the HM for HDR content, as done for the SDR case

The recommendations above were approved in track B, however the investigations of benefits of in-loop reshaping and chroma refinement for SDR shall be included in CE12 (not separate CE), to be renamed as “mapping function”

It is noted that for the next meeting it should planned to have appropriate viewing equipment both for HDR and SDR.

[**JVET-K0559**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4089) **Report of BoG on Picture Boundary Partitioning [K. Misra]**

[**JVET-K0562**](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=4092) **Report of BoG on Software development and CTC [F. Bossen]**

Reviewed in track B Tues 17 July 1020 (chaired by JRO)

The BoG on Software development and CTC met on Monday July 16 2018 from 9:15 to 13:15 in room Povodni Moz 2. It reviewed input contributions related to software development and CTC, and encoder optimizations. Recommendations from the BoG include:

* JVET-K0054: adopt unified PSNR calculation with runtime configurability (default off). Usage should be described in CTC document
* JVET-K0149: enhance dtrace to be able to produce additional proposed statistics. Add reference to external viewer in SW manual (with proper disclaimer).
* JVET-K0154: enable FULL\_NBIT macro by default, mainly effective to increase chroma gain for HDR (pending SIMD code review)
* JVET-K0261: adopt proposed SW cleanup changes (change macro values if needed and remove disabled code) – JEM compatibility no longer practical
* JVET-K0312: add tool usage metrics to SW. SW coordinators to determine best implementation (e.g., with or without dtrace). Encourage CU-level tool proponents to add and report such metrics (proportion of samples affected by tool) – this would be highly beneficial to further assess compression benefit and complexity impact of a tool, as well as its usage in terms of sequence characteristics.
* JVET-K0349: migrate from svn to git (GitLab) (this was previously recommended in AHG3 report)
* JVET-K0410: add fields in Excel reporting template to capture information such as processor, SIMD settings, compiler, compiler options, decoder options
* JVET-K0461: adopt guidelines for VVC SW development, add paragraphs on “const” variables and preference for shorter functions
* JVET-K0390: adopt proposed rate control method and bug fixes (pending code review)
* JVET-K0206: adopt changes into VTM (remains disabled by default). SW AHG to decide whether to enable corresponding macro (command line parameter would still disable the feature by default)

All recommendations of the BoG were approved in track B

Revisit: Item of accessibility of CE software to be discussed in JVET plenary (this issue was mentioned in the context of discussing the BoG report)

Copy the verbose dispositions about individual proposals from the 360, HDR and software BoG reports to the respective sections.

## List of actions taken affecting Draft 2 of VVC, VTM 2, BTM and 360Lib

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.

### Encoder only or CTC/software changes

Implement VTM/BMS SW as the original HEVC deblocking, filtering on an 8x8 grid as minimum size

(no spec text on deblocking).

JVET-K0185: Fast encoder for affine from 4.1.5c in combination with normative adoption 4.1.3c

JVET-K0157: “long-term reference + no-output picture”

Increase chroma QP by 1 when trees are separate

JVET-K0238: Encoder side grouping of CTUs for SAO parameter derivation

JVET-K0352: Encoder optimization for merge

### Syntax/semantics/decoding process changes VTM/WD

General: QT/BT/TT no longer “placeholder”

JVET-K0351 (test c): Partitioning: Keep only the TT restriction (preventing binary split with same orientation in center partition of the ternary split) in VTM at both encoder and decoder

JVET-K0554: Partitioning: Implicit splitting at picture boundaries

JVET-K0307/JVET-K0237/JVET-K0369/JVET-K0232/JVET-K0315: Deblocking bugfix: Perform deblocking at boundaries of TUs with any size >=64

Implement VTM/BMS SW as the original HEVC deblocking, filtering on an 8x8 grid as minimum size (no spec text on deblocking).

JVET-K0371: ALF (based on subtest 2.4.1.4c, 4x4 classification based on Laplacian for luma only, 7x7 luma, 5x5 chroma filters); disable prediction of adaptive filters from fixed filter set; disable temporal prediction; put filter parameters into slice header; Enabling flag at CTU level.

JVET-K0063: PDPC

JVET-K0190: 4.1.8 CCLM only

JVET-K0184: Affine MC (CE4.1.1a 4x4 fixed subblock size).

JVET-K0337: Affine MC coding and models (4.1.3a, affine MVP list construction, and 4.1.3b, MV difference coding, and 4.1.3c, 4/6 parameter model, no slice level switch).

JVET-K0357: AMVR (4.3.3a) from BMS to VTM

Transform: Adopt AMT (both intra and inter, each controlled by an SPS flag) as follows (approx 3.3%, RA 2.0%, LB 1.3%):

* No 64-length DST7 and DCT8 (no AMT syntax sent when either dimension is larger then 32)
* No 128-length DCT2
* Only DCT2, DST7 and DCT8
* All transforms are to have 10 bit coefficients
* Uses the syntax that has been in the BMS. AMT is applied only for luma. There are separate enabling flags for intra and inter at SPS level. When AMT is enabled, then
  + If CBF=1, then
    - A flag for DCT2 in both directions; if not then
      * If (intra and the number of nonzero coefficients is greater than two) or inter (regardless of the number of nonzero coefficients)
        + Flag for horizontal is DST7 vs. DCT8
        + Flag for vertical is DST7 vs. DCT8
      * Otherwise (intra block with only 1 or 2 nonzero coefficients), DST7 is used both horizontally and vertically
* Rather than AMT, suggested name is multiple transform selection (MTS)

JVET-K0251: Increase max QP from 51 to 63.

JVET-K0230/K0556: Separate trees for intra slices (without multi-DMs) with an implicit split to 64x64; Prohibit ternary split of something bigger than 64 in width or height (and not send the bit to indicate ternary type at that level).

JVET-K0122: DC prediction second method (use only the longer side to compute the average for non-square blocks)

JVET-K0368 with modifications tested in JVET-K0529: TU binarization, 67 intra modes from BMS, 3 MPM

JVET-K0539: Include the wide angles in the expansion of the number of angles (85 angles, 87 modes total)

ATMVP from BMS to VTM, with modifications from JVET-K0346:

* One fixed collocated picture is used to derive temporal motion information.
* Slice level adaptive sub-block switching, 8x8 or 4x4.
* Constrain the region from where ATMVP motion is derived to the collocated CTU plus one 4x4 block column outside the collocated CTU at the right hand side, the same region for HEVC TMVP.

JVET-K0072: Alternative entropy coding for dependent quantization

JVET-K0310: Sign data hiding

JVET-K0325: Simple high-level syntax

### BMS

All modifications from VTM

JVET-K0248: Generalized B prediction to BMS

JVET-K0076: 8.2.2 Current picture referencing

JVET-K0217 (variant 9.2.9l) bi-directional template matching, along with the method from JVET-K0199 (9.1.1.a) that refined vectors are not used to predict MVs from neighbour blocks, and not for deblocking

JVET-K0485: BIO (simplified variant)

### Changes in 360Lib

JVET-K0131: Modified cubemap projection

# Project planning

## Core experiment planning (update)

The following CEs were initially planned (Wed 18th 1630) It was emphasized that this was an initial list, and it was still to be decided after a presentation of an initial CE description if the respective CE will be finally established:

1. Partitioning (J. Ma (primary), M. W. Park, [Thu: Add per document])
2. In-loop filters (L. Zhang, K. Andersson, [Thu: added Y. Tung])
3. Intra prediction and mode coding (G. Auwera, J. Heo)
4. Inter prediction and MV coding (H. Yang, S. Liu)
5. Arithmetic coding engine (T. Nguyen, A. Said)
6. Transforms and transform signalling (A. Said, X. Zhao)
7. Quantization and coefficient coding (M. Coban, H. Schwarz)
8. Current picture referencing (X. Xu, K. Müller)
9. Decoder side MV derivation (S. Esenlik, Y.W. Chen)
10. Combined and multi-hypothesis prediction (C.W. Hsu, M. Winken)
11. Composite reference pictures (X. Zheng)

CE draft developers shall present initial versions of CE proposals Thu. afternoon, containing

* list of sub-experiments, origin of the technology to be investigated (e.g., CfP response document number), expected results, method of investigation
* Participating parties and cross-checkers
* Expected interdependency with other CEs

Interested parties were asked to get in contact with CE draft developers as listed above.

Initial descriptions of CEs 1 and 2 were orally reviewed Thursday 19 April 1600–1630.

For CE1: transform coefficient coding should be used from test (or with minor alignments when necessary by the partitioning); estimated number of configurations that will be tested to be reported on Friday. JVET-J1021

For CE2: It was noted that deblocking in the BMS is already parallelizable. It was suggested to include HDR test sequences in deblocking tests.

Regarding the general rule applying to CE plans established at this meeting, it was confirmed on Friday 20 April (1200, GJS and JRO) that each CE is planned based on technology provided in responses to the CfP, there may be subtests within each CE that are based on other contributions (or hypothetical combinations, etc.), provided there is agreement to include such testing.

It was discussed on 1230 Friday 20 whether the adaptive-resolution CNN technology should be in the intra prediction CE. This seemed to be different from mere intra prediction, as the resolution reduction is also applied to the residual in that scheme. It seemed too late in the meeting to try to define another CE. It was commented that the proposed technology is certainly interesting and should be studied in the AHG 9.

It was furthermore agreed in the Friday plenary that each CE should have a maximum of 3 coordinators. The role of CE coordinators is again clarified. It is not necessary that each sub-CE has an own coordinator. People in sub-CEs should communicate with each other about how to compare if each other and agree on a compiled version of their part before sending it to the overall coordinator.

## JEM description drafting 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
* JEM 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 Thursday 11 Jan. 2018.

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

Move to appropriate place in notes: It was agreed that proponents should not publish specific claims or precise measurements about the subjective performance of their proposal in the CfP test.

This section was reviewed Thursday 19 April afternoon.

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 main branch of JEM by the next meeting.
* 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. The software coordinator will coordinate the creation of these branches. All JVET members can obtain read access to the CE software branches. The access method will be announced on the JVET reflector within two weeks after the meeting.
* During the experiment, further improvements of the planned experiment can be made
* 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-J1010.

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 company proponent perspective – e.g. 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 referenced documents that are also available in the JVET document archive.

Any technology must have at least one cross-check partner to establish an CE – a single proponent is not enough. It is highly desirable have more than just one proponent and one cross-checker.

Some agreements relating to CE activities were established as follows:

* Only qualified JVET members can participate in an CE.
* Participation in an 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.
* All substantial communications about a CE, other than logistics arrangements, exchange of data, minor refinement of the test plans, and preparation of documents shall be conducted on the main JVET reflector. In the case that large amounts of data are to be distributed is recommended to send an announcement to the JVET reflector without attaching the materials, and send the materials to those who have requested it directly, or provide a link to it, or upload the data as an input contribution to the next meeting.

General timeline

T1= 3 weeks after the JVET meeting: To revise EE description and refine questions to be answered. Questions should be discussed and agreed on JVET reflector.

T2 = Test model SW release + 2 weeks: Integration of all tools into separate EE branch of JEM 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 encouraged to study and make contributions to the next meeting with proposed changes

T3: 3 weeks before the next JVET meeting: Any changes to the exploration branches 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 and announced on the JVET reflector. The name of the cross-checkers and list of specific tests for each tool under study in the EE will be announced in JVET reflector by this time. Full test results must be provided at this time (at least for proposals targeting to be promoted to JEM at the next meeting).

New branches may be created which combine two or more tools included in the EE document or the JEM. Requests for new branches should be made to the software coordinators.

Don’t need to formally name cross-checkers in the EE 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 favorable outcome of CE does not indicate a need for adoption of the technology.

## Software development and anchor generation

The planned timeline for software releases was established as follows:

* VTM1.0 will be released by 2018-05-04. This version will also include the implementation of BMS configuration in a separate branch.
* JEM7.2 will be released by 2018-05-04.
* Further versions of VTM may be released for additional bug fixing, as appropriate.
* Tools that are not in TM or BMS will be kept in another separate branch for the current meeting cycle (e.g. may be used in some CEs), by default disabled by macros.

Timeline of 360lib6.0: 1 week after the release of VTM1.0 (2018-05-11). Further versions may be released as appropriate for bug fixing.

# 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)).

|  |  |  |
| --- | --- | --- |
| **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 CE and AHG studies. * Report on project status to JVET reflector. * Provide a report to next meeting on project coordination status. | J.-R. Ohm, G. Sullivan | N |
| **Draft text and test model algorithm description editing (AHG2)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Produce and finalize JVET-K1001 VVC text specification Working Draft 2. * Produce and finalize JVET-K1002 VVC Test Model 2 (VTM 2) Algorithm and Encoder Description. * Gather and address comments for refinement of these documents. * Coordinate with Test model software development AhG to address issues relating to mismatches between software and text. | B. Bross, J. Chen (co-chairs), J. Boyce, S. Kim, S. Liu, 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) and benchmark set (BMS) software based on the NextSoftware package and release software packages with associated configuration files (available by X, repository to be announced via reflector). * Produce documentation of software usage for distribution with the software. * Discuss and make recommendations on the software development process. * Propose improvements to the guideline document for developments of the test model 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) | N |
| **Test material and visual assessment (AHG4)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Maintain the video sequence test material database for development of the VVC standard. * Identify and recommend appropriate test materials for use in the development of the VVC standard. * Identify missing types of video material, solicit contributions, collect, and make available a variety of video sequence test material. * Evaluate new test sequences, and prepare for the visual assessment and availability of viewing equipment in the next meeting. | V. Baroncini, R. Chernyak, P. Hanhart, A. Norkin, T. Suzuki, J. Ye (co-chairs) | N |
| **Memory bandwidth consumption of coding tools (AHG5)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Develop improved software tools for measuring both average and worst case of memory bandwidth, and provide information for usage of these tools. * Study cache configurations for measuring decoder memory bandwidth consumption. * Identify coding tools in CEs, VTM, and BMS with significant memory bandwidth impact. * Study the impact of memory bandwidth on specific application cases. | R. Hashimoto (chair), Y. He, T. Ikai, X. Li, H. Yang, M. Zhou (vice-chairs) | N |
| **360° video conversion software development (AHG6)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Prepare and deliver the 360Lib-7.0 software version and common test condition configuration files according to JVET-K1012. * Generate CTC VTM and BMS anchors according to JVET-K1012, and finalize the reporting template for the common test conditions. * Produce documentation of software usage for distribution with the software. | Y. He and K. Choi, (co-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 of the CfP responses. * Compare the performance of the VTM, BMS, and HM for HDR/WCG content. * Prepare for expert viewing of HDR content at the 12th JVET meeting. * Coordinate implementation of HDR anchor aspects in the test model software with AHG3. * Study additional aspects of coding HDR/WCG content. | A. Segall (chair), E. François, W. Husak, D. Rusanovskyy (vice-chairs) | N |
| **360° video coding tools and test conditions (AHG8)**  ([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. * Discuss refinements of common test conditions, test sequences, and evaluation criteria. * Solicit additional test sequences, and evaluate suitability of test sequences on head-mounted displays and normal 2D displays. * Study coding tools dedicated to 360° video, their impact on compression, and implications to the core codec design. * Study the effect of viewport resolution, field of view, and viewport speed/direction on visual comfort. * Study complexity of GPU rendering of projection formats * Study syntax for signaling of projection formats | J. Boyce (chair), K. Choi, P. Hanhart, J.-L. Lin (vice chairs) | N |
| **Neural networks in video coding (AHG9)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Investigate the benefit of using neural networks in video compression such as CNN loop filter, intra prediction, re-sampling in adaptive resolution coding, and encoder side partition mode decisions. * Investigate the complexity impact of using neural networks in video compression. * Investigate the complexity measurement of neural network coding tools. * Investigate the impact of training materials on the performance of neural network coding tools. * Investigate the impact of the training process on performance and complexity. | S. Liu (chair), B. Choi, K. Kawamura, Y. Li, L. Wang, P. Wu, H. Yang (vice-chairs) | N |
| **Encoding algorithm optimizations (AHG10)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study the impact of using techniques such as GOP structures and perceptually optimized adaptive quantization for encoder optimization. * Study the impact of adaptive quantization on individual tools in the test model. * Study the quantization adaptation tool in the test model. * Investigate the feasibility of adding a CTC test category in which adaptive quantization is turned on. * Study quality metrics for measuring subjective quality using e.g. the CfP response MOS scores. * Investigate other methods of improving objective and/or subjective quality, including adaptive coding structures, adaptive quantization without signalling, and multi-pass encoding. * Study methods of rate control and their impact on performance, subjective and objective quality. | A. Duenas and A. Tourapis (co-chairs), C. Helmrich, S. Ikonin, A. Norkin, R. Sjöberg (vice-chairs) | N |
| **Screen content coding (AHG11)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Investigate coding tools targeted at screen content in terms of compression benefit and implementation complexity. * Identify test materials and discuss testing conditions for screen content coding. | S. Liu (chair), J. Boyce, A. Filippov, Y.-C. Sun, M. Zhou (vice-chairs) | N |
| **High-level parallelism and coded picture regions (AHG12)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study high-level parallelism techniques. * Study concepts and proposed methods of representation of coded picture regions such as tiles and slices. * Study usage and additional functionalities for coded regions that may be beneficial beyond what has been done in existing standards * Prepare software and configurations for the test model to facilitate parallel processing tests. * Study the coding efficiency impact of parallel processing and coded picture regions. | T. Ikai (chair), M. Coban, M. M. Hannuksela, H. M. Jang, R. Sjöberg, R. Skupin, Y.-K. Wang |  |
| **Tool reporting procedure (AHG13)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Prepare output document JVET-K1005, which describes the methodology of tool-on/tool-off testing, provides a reporting template, and a list of tools to be tested by identified testers. * Provide configurations files, bitstreams, and results of the tool-on/tool-off testing. * 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), R. Chernyak, K. Choi, R. Hashimoto, Y. He, Y.**-**W. Huang, S. Liu (vice-chairs) |  |
| **Low-latency random access (AHG14)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Define relevant test conditions to evaluate low-latency encoding with progressive intra refresh for random access without intra frames. * Study non-normative ways to produce progressive intra refresh with minimum losses in coding efficiency. * Propose software modifications for integrating encoder-only intra refresh in the VTM and BMS model. * Characterize progressive intra refresh performance objectively and subjectively. * Study normative solutions to improve intra refresh performance against encoder-only intra refresh. | J.-M. Thiesse (chair), A. Duenas, K. Kazui, A. Tourapis (vice-chairs) |  |
| **Bitstream decoding properties signalling (AHG15)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study syntax alternatives for interoperability point signalling * Study selection of constraint flags to be included in the VTM and their impact on syntax, semantics, and decoding process | J. Boyce (chair), J. Chen, M. Karczewicz, A. Tourapis, Y.-K. Wang, S. Wenger (vice-chairs) | Tel. (TBA approx monthly) |
| **Implementation (AHG16)**  ([jvet@lists.rwth-aachen.de](mailto:jvet@lists.rwth-aachen.de))   * Study draft and proposed coding tools to identify implementation issues relating to decoder pipelines, decoder throughput, and other aspects of implementation difficulty. * Solicit hardware analysis of complex tools. * Provide feedback on potential solutions to address identified issues. | M. Zhou (chair), E. Chai, K. Choi, S. Ethuraman, T. Hsieh, D. Symes, X. Xiu (vice-chairs) |  |

# 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-K1000 Meeting Report of the 11th JVET Meeting [G. J. Sullivan, J.-R. Ohm] (2018-07-03, near next meeting)

(Initial versions of the meeting notes (d0 … d8) were made available on a daily basis during the meeting.)

JVET-K1001 Versatile Video Coding (Draft 2) [B. Bross, J. Chen, S. Liu] [WG11 N17732] (2018-08-31)

(Initial version planned to be made available by 2018-05-07.)

JVET-K1002 Algorithm description for Versatile Video Coding and Test Model 1 (VTM 1) [J. Chen, Y. Ye, S. Kim] [WG11 N17733] (2018-08-31)

(Initial version planned to be made available by 2018-05-07.)

JVET-K1003 Guidelines for VVC reference software development [K. Sühring] (2018-07-31)

JVET-K1004 Algorithm descriptions of projection format conversion and video quality metrics in 360Lib Version 7 [Y. Ye, J. Boyce]

Remains valid - not reissued [JVET-J1005](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=3541) Methodology and reporting template for tool testing [W.-J. Chien, J. Boyce, E. Alshina, J. Chen, E. François, Y. He, Y. W. Huang]

JVET-K1010 JVET common test conditions and software reference configurations [F. Bossen, J. Boyce, X. Li, V. Seregin, K. Sühring] (2018-07-31)

(Initial version planned to be made available by 2018-04-20.)

JVET-K1011 JVET common test conditions and evaluation procedures for HDR/WCG video [A. Segall, E. François, D. Rusanovskyy] (2018-07-31)

(Initial version planned to be made available by 2018-04-27.)

JVET-K1012 JVET common test conditions and evaluation procedures for 360° video [P. Hanhart, J. Boyce, K. Choi] (2018-07-31)

(Initial version planned to be made available by 2018-04-27.)

JVET-K1021 Description of Core Experiment 1 (CE 1): Partitioning [J. Ma, F. Le Léannec, M. W. Park]

(Initial version presented XX.)

Discussion Monday 1830 (GJS & JRO)

* Boundary handling
* Implementation-friendly modifications (e.g., 64x64 pipeline friendly)
* Separate tree for intra regions in inter slices? []

JVET-K1022 Description of Core Experiment 2 (CE2): Adaptive Loop Filter [V. Seregin, C.-Y. Chen ]

(Initial version presented XX.)

Discussion Monday 1840 (GJS & JRO)

* ALF (filter shapes, CTU-based, filter parameter coding, classification, low-latency aspects) []

JVET-K1023 Description of Core Experiment 3 (CE3): Intra Prediction and Mode Coding [G. Van der Auwera, J. Heo, A. Filippov]

(Initial version presented XX.)

Discussion Monday 1850 (GJS & JRO)

* Multiple reference lines
* Interpolation
* Line-based prediction
* Nonlinear weighted intra prediction
* Modified cross-component prediction
* Intra mode coding (e.g., 6 MPM)
* Bidirectional prediction

JVET-K1024 Description of Core Experiment 4 (CE4): Inter prediction and motion vector coding [H. Yang, S. Liu, K. Zhang]

(Initial version presented XX.)

Discussion Monday 1900 (GJS & JRO)

* Merging (affine & non-affine)
* Other affine aspects?
* Padding
* MVD coding
* Illumination compensation
* Motion field compression

JVET-K1025 Description of Core Experiment 5 (CE5): Arithmetic Coding Engine [H. Kirchhoffer, A. Said]

(Initial version presented XX.)

Discussion Monday 1910 (GJS & JRO)

* Table-based probability estimation, single & double window, custom window size

JVET-K1026 Description of Core Experiment 6 (CE6): Transforms and transform signalling [A. Said, X. Zhao]

(Initial version presented XX.)

Discussion Monday 1920 (GJS & JRO)

* Primary transform (factorization, precision, selection of the transform, spatial coverage of transform, additional or alternative transform types, handling of chroma)
* Secondary transform

JVET-K1027 Description of Core Experiment 7 (CE 7): Quantization and coefficient coding [H. Schwarz, M. Coban, C. Auyeung]

(Initial version presented XX.)

Discussion Monday 1930 (GJS & JRO)

* Context selection
* Reduced number of context models
* Reduced number of context-coded bins
* Alternative state machine dependent quantization
* Scanning order
* Modified residual sign prediction
* Spatial-domain residual scaling

JVET-K1028 Description of Core Experiment 8 (CE8): Current Picture Referencing [X. Xu, K. Müller, L. Wang]

(Initial version presented XX.)

Discussion Monday 1940 (GJS & JRO)

* Constraints
* Template matching

JVET-K1029 Description of Core Experiment 9 (CE9): Decoder-Side Motion Vector Derivation [S. Esenlik, Y. W. Chen, F. Chen, X. Chen (4?)]

(Initial version presented XX.)

Discussion Monday 1945 (GJS & JRO)

* DMVR interpolation filters, padding, search range, partial usage of refined MVs
* Matching method

JVET-K1030 Description of Core Experiment 10 (CE10): Combined and multi-hypothesis prediction [C.-W. Hsu, M. Winken, X. Xiu]

(Initial version presented XX.)

Discussion Monday 1955 (GJS & JRO)

* OBMC, non-rectangular partitions, diffusion filtering, prediction with more than two hypotheses, other blending of multiple predictors

JVET-K1031 Description of Core Experiment 11 (CE11): Deblocking [A. Norkin, A. M. Kotra]

(Initial version presented XX.)

* longer filters,
* 4x4 deblocking, …

JVET-K1032 Description of Core Experiment 12 (CE12): Mapping functions [E. François, D. Rusanovskyy, P. Yin]

(Initial version presented XX.) [TBD]

JVET-K1033 Description of Core Experiment 13 (CE13): Coding tools for 360° omnidirectional video [P. Hanhart, J.-L. Lin]

(Initial version presented XX.) [TBD]

Discussion Monday 2010 (GJS & JRO)

* Intra prediction, inter prediction, in-loop filters, padding, post-filtering, blending

JVET-K1034 Description of Core Experiment 14 (CE14): Post-reconstruction filtering [L. Zhang, S. Ikonin]

(Initial version presented XX.) [TBD]

* Bilateral
* Hadamard-based

JVET-K1035 Description of Core Experiment 15 (CE15): Palette mode [Y.-C. Sun, Y.H. Chao, X. Xu]

(Initial version presented Friday morning.)

Discussion Tuesday morning track B (JRO)

* Investigate the palette variant proposed in K0411 and HEVC-SCC palette mode
* Investigate interrelationship with CPR
* Study the complexity impact of the two palette variants and CPR (in coordination with CE8)

# 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 (starting meetings on the Tuesday or Wednesday of the first week and closing it on the Tuesday or Wednesday of the second week of the SG 16 meeting – a total of 6–7.5 meeting days), and
* Otherwise meeting under ISO/IEC JTC 1/SC 29/WG 11 auspices when it meets (starting meetings on the Wednesday or Thursday prior to such meetings and closing it on the last day of the WG 11 meeting – a total of 8.5 meeting days).

In cases where 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. 3 – Fri. 12 Oct. 2018, 12th meeting under WG 11 auspices in Macao, CN.
* Wed. 9 – Fri. 18 January 2019, 13th meeting under WG11 auspices in Marrakesh, MA.
* Tue. 19 – Wed. 27 March 2019, 14th meeting under ITU-T auspices in Geneva, CH.
* Thu. 4 – Fri. 12 July 2019, 15th meeting under WG11 auspices in Gothenburg, SE.

The agreed document deadline for the 12th JVET meeting is XXday XX Sept. 2018. Plans for scheduling of agenda items within that meeting remain TBA.

Kenzler Conference Management and Silke Kenzler in person were thanked for the excellent hosting and organization of the 11th meeting of the JVET.

XXX were thanked for providing viewing equipment used during the 11th JVET meeting.

XXX were thanked for providing new test material for usage in standardization efforts.

The 11th JVET meeting was closed at approximately XXXX hours on Wednesday 18 July 2018.

# Annex A to JVET report: List of documents

# Annex B to JVET report: List of meeting participants

The participants of the tenth meeting of the JVET, according to a sign-in sheet circulated during the meeting sessions (approximately XXX people in total), were as follows:

1. …