|  |  |
| --- | --- |
| **Joint Collaborative Team on 3D Video Coding Extension Development**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  2nd Meeting: Shanghai, CN, 13–19 Oct. 2012 | Document: JCT3V-B0046 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **AHG5: MV-HEVC software for HTM** | | |
| *Status:* | Input Document | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Li Zhang Ying Chen Jewon Kang  5775 Morehouse Dr San Diego, CA 92121 USA | Tel: Email: | 1-858-651-6660 [lizhang@qti.qualcomm.com](mailto:lizhang@qti.qualcomm.com)  1-858-845-6589 [cheny@qti.qualcomm.com](mailto:cheny@qti.qualcomm.com)  1-858-651-8457 [jewonk@qti.qualcomm.com](mailto:jewonk@qti.qualcomm.com) |
| *Source:* | Qualcomm Incorporated | | |

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

# Abstract

The current HTM software doesn’t contain an MV-HEVC software branch/configuration. The current HTM software only supports 3D-HEVC, although there are macros in the software that could be used to turn off advanced coding tools, such as inter-view motion/residual prediction, to hypothetically simulate MV-HEVC for the purpose of e.g., roughly evaluate the performance of MV-HEVC. The MV-HEVC software as a branch of HTM is helpful for the evaluation of the MV-HEVC proposals and facilitates the MV-HEVC standardization procedure, including the conformance test at a later stage. Therefore, a software branch is provided to align the MV-HEVC working draft with minor encoder/decoder changes. This software branch can be easily configured as part of the HTM software.

# Introduction

To align with MV-HEVC working draft (WD), there are two major aspects that need to be taken care of.

* Alignment with the latest HEVC specification for base view and non-base views.
* Alignment with MV-HEVC WD, in terms of e.g., 3DV motion hook in HEVC and reference picture marking in MV-HEVC.

The software package mainly aims at the second aspect, while the first aspect belongs to a potential migration of the HTM software to the latest HEVC specification and HM software.

Although this proposal only focuses on MV-HEVC, some modified aspects, e.g., NAL unit header, reference picture marking, are applicable to both MV-HEVC and 3D-HEVC.

The following were modified to align the MV-HEVC.

* Parameter sets: VPS, VPS extension, SPS, including removal of camera parameters.
* NAL unit header syntax elements, including nal unit type, layer\_id and temporal\_id.
* Decoded picture marking for inter-view reference pictures.
* Slice header, including removal of camera parameters and reference picture list construction.
* Motion prediction to be aligned with HEVC changes for merge and AMVP, including motion hooks.
* Other code cleanups.

The following items focusing on high level syntax, reference pictures may be further modified as part of the future work:

* nal\_ref\_flag should be removed from NAL unit header.
* The new nal unit types introduced in HEVC WD 8 should be included.
* Motion compression should be done for a view component, which is to be used as an inter-view reference picture, especially if it is to be used as a co-located picture of the current picture.
* Removal of combined list.
* Removal of Adaptation Parameter Set (APS) and Adaptive Loop Filter (ALF).
* Other various aspects to reflect the difference between HM 8 and HM 6.
* The extractor needs to be modified to align with encoder/decoder changes.

Some of the software modifications are applicable to the 3D-HEVC software branch, since they are aligned to 3D-HEVC as well.

# Description of the major software changes

## Parameter Sets

* In video parameter sets (VPS), the view dependencies are signalled, besides, operation points are also signalled. The view\_id of each view is also signalled in the VPS.
* Camera parameters are not needed in both 3D-HEVC and MV-HEVC, so they are removed from sequence parameter set (SPS). Encoder supports sharing one SPS for multiple views.

We may consider transmitting camera parameter set in additional SEI messages.

## NAL unit header

* layer\_id is signalled with 5-bit by reusing the reserved\_zero\_5bits. View order index of each NAL unit is derived to be equal to layer\_id.
* nuh\_temporal\_id\_plus1 is signalled instead of temporal\_id.
* NAL unit type related to random access point (RAP), i.e., NAL\_UNIT\_CODED\_SLICE\_IDV is removed. MV-HEVC assumes that a random access point uses the same NAL unit type as that of the base view RAP view component, thus no new nal unit type for those RAP view components is needed.

Current software supports only two types of RAPs, IDR and CRA, more types need to be integrated from HM together with the removal of nal\_ref\_flag.

## Decoded picture marking

* At the encoder, temporal reference pictures, as in RefPicSetStCurrBefore and RefPicSetStCurrAfter are marked as “short-term”, pictures in inter-view reference picture set are marked as “long-term”.

## Slice header

* Camera parameters are removed from the slice header.
  + Camera parameters were signalled in the slice header when they are not constant within the whole sequence. Since they are not needed in MV-HEVC (actually also in 3D-HEVC), the transmitted scale and offset values used for warping of depth samples are excluded from the slice header.
* Reference picture list construction
  + In HTM, temporalID is always set to 0. It is changed to be aligned with HM, so that not all pictures (only those with lower or equal temporalID values) can be added into the reference picture list of the current picture.
  + To avoid unnecessary reference picture list modification, a new syntax element is introduced in slice header to specify the initial position of the inter-view reference picture (s) in the reference picture list.

## Motion prediction

In MV-HEVC, all the inter-view reference pictures are used as long-term reference pictures. Therefore, in various places, identifying an inter-view reference based comparing of POC value or view order index are replaced by checking whether the reference picture is/was long-term or not, including the following aspects.

* AMVP
  + Unless a neighboring (spatial or temporal) block and the current block both refer to the same type of reference pictures, i.e., either both short-term or both long-term, the candidate from the neighboring block is considered as unavailable.
* Merge
  + The target reference index of temporal merging candidate is always set to 0 to be aligned with the latest HM. When a co-located block in the co-located picture refers to a reference picture which has a different type (short-term or long-term) than identified by the target reference index (0), the candidate from the co-located block is considered as unavailable.

# Simulation results

To evaluate the performance of the proposed software, 3D-HTM 4.0 is used as the anchor and common test conditions [2] for coding texture views are utilized except that low-level coding tools such as inter-view motion/residual prediction are disabled. Note the target reference index of temporal merging candidate is set to 0 in 3D-HTM4.0 to be consistent of HEVC.

The performance of the MV-HEVC software compared with 3D-HTM4.0 is illustrated in Table 1, where temporal scalability is not supported by the encoder (fixed temporalID, i.e., 0 for all frames).

Results are also provided when temporalID setting is aligned with HM, as shown in Table 2. Even with some minor loss in the base view, some minor overall bitrate saving (0.03%) is observed. The loss in the base view is relatively small and is aligned with the loss that can be observed when similar changes (by setting tempoalID values in a hierarchical fashion at the encoder) were made in HM.

Table 1: MV-HEVC (without temporal scalability) VS HTM4.0 anchor

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Video 0 | Video 1 | Video 2 | Video only |
| Balloons | 0.0% | -0.1% | -0.2% | 0.0% |
| Kendo | 0.1% | -0.1% | -0.2% | -0.1% |
| Newspapercc | 0.0% | -0.2% | -0.2% | -0.1% |
| GhostTownFly | 0.0% | -0.8% | -1.4% | -0.3% |
| PoznanHall2 | 0.1% | -0.2% | -0.4% | -0.1% |
| PoznanStreet | 0.0% | -0.2% | -0.1% | 0.0% |
| UndoDancer | 0.0% | -0.1% | 0.0% | 0.0% |
| 1024x768 | 0.0% | -0.2% | -0.2% | 0.0% |
| 1920x1088 | 0.0% | -0.3% | -0.5% | -0.1% |
| average | **0.04%** | **-0.24%** | **-0.36%** | **-0.08%** |

Table 2: MV-HEVC (with temporal scalability) VS HTM4.0 anchor

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Video 0 | Video 1 | Video 2 | Video only |
| Balloons | 0.1% | -0.2% | -0.1% | -0.1% |
| Kendo | 0.1% | -0.1% | -0.2% | 0.0% |
| Newspapercc | 0.1% | -0.1% | -0.1% | 0.0% |
| GhostTownFly | 0.1% | -0.4% | -1.2% | -0.1% |
| PoznanHall2 | 0.1% | -0.2% | -0.9% | -0.2% |
| PoznanStreet | 0.1% | 0.1% | 0.1% | 0.1% |
| UndoDancer | 0.2% | 0.0% | 0.0% | 0.1% |
| 1024x768 | 0.1% | -0.2% | -0.1% | 0.0% |
| 1920x1088 | 0.1% | -0.1% | -0.5% | 0.0% |
| average | **0.10%** | **-0.13%** | **-0.34%** | **-0.03%** |

To evaluate the performance of the proposed reference picture list construction method, the coding gain of the proposed method with respect to the original method in HTM (the proposed method is disabled) is listed in the following table. It can be observed that there is 0.1% and 0.05% bit rate reduction for dependent views and overall texture views, respectively. Detail results could be found in the attached file with the title as ‘JCT3V-B0046\_RPLI on vs off.xls’.

Table 3: Coding gain of proposed reference picture list modification (without temporal scalability)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Video 0 | Video 1 | Video 2 | Video only |
| Balloons | 0.0% | -0.1% | -0.1% | -0.1% |
| Kendo | 0.0% | -0.1% | -0.1% | -0.1% |
| Newspapercc | 0.0% | -0.1% | -0.1% | -0.1% |
| GhostTownFly | 0.0% | -0.1% | -0.1% | 0.0% |
| PoznanHall2 | 0.0% | -0.2% | -0.2% | -0.1% |
| PoznanStreet | 0.0% | -0.1% | -0.1% | 0.0% |
| UndoDancer | 0.0% | -0.1% | -0.1% | 0.0% |
| 1024x768 | 0.0% | -0.1% | -0.1% | -0.1% |
| 1920x1088 | 0.0% | -0.1% | -0.1% | 0.0% |
| average | **0.00%** | **-0.12%** | **-0.11%** | **-0.05%** |

# Proposal

* To adopt the proposed MV-HEVC software as part of the HTM software package.
* To start the migration of the HTM software to be aligned to HM 8 or 9.
* To align the 3D-HEVC functions with MV-HEVC if applicable.

# References

1. G. Tech, K. Wegner, Y. Chen, M. M. Hannuksela, "MV-HEVC Working Draft 1," ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JCT3V-A1004, 1st Meeting: Stockholm, SE, 16–20 July 2012.
2. D. Rusanovskyy, K. Müller, A. Vetro, "Common Test Conditions of 3DV Core Experiments," ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JCT3V-A1100, 1st Meeting: Stockholm, SE, 16–20 July 2012.

# Patent rights declaration(s)

**Qualcomm Incorporated may have current or pending patent rights relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).**