|  |  |
| --- | --- |
| **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-B0088 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **3D-CE5.h related: Unified temporal collocated picture in DV derivation** | | |
| *Status:* | Input Document | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Yi-Wen Chen1, Jian-Liang Lin1, Yu-Wen Huang1, Shawmin Lei1, Na Zhang2,3, Jicheng An2, Kai Zhang2, Siwei Ma4, Debin Zhao3, Wen Gao4  1MediaTek Hsinchu, No. 1, Dusing Rd. 1, Hsinchu Science Park, Hsinchu, Taiwan 30078  2MediaTek Beijing, North Building 10F, Raycom Infotech Park Tower C, No. 2 Kexueyuan South Rd., Haidian District, Beijing, China 100190  3 Department of Computer Science and Technology, Harbin Institute of Technology, Harbin, China 150001  4School of EECS, Peking University, Institute of Digital Media, Room 2641, Science Building #2, No.5, Yiheyuan Rd., Haidian District, Beijing, China, 100871 | Tel: Email: | Shawmin Lei +886-3-5670766 ext. 25555  {yiwen.chen, jl.lin, yuwen.huang, na.zhang, jicheng.an, kai.zhang, shawmin.lei}@mediatek.com  {swma, dbzhao, wgao}@jdl.ac.cn |
| *Source:* | MediaTek Inc. | | |

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

# Abstract

In the HEVC-based 3D video coding, HTM 4.0.1, a mechanism that allows the collocated picture to be changed at prediction unit (PU) level is utilized to search temporal neighboring blocks to derive the disparity vector. This contribution proposes a constraint to unify the temporal collocated picture for all PUs within the same picture to ease the memory access burden caused by the PU-level adaptive collocated picture selection. In this contribution, two schemes are proposed to provide different signaling methods of the temporal collocated picture for disparity vector (DV) derivation. In scheme 1, an additional syntax is used to signal the temporal collocated picture for DV derivation in the same manner as the signaling of temporal collocated picture for temporal motion vector predictor (TMVP). In scheme 2, the signaling of the temporal collocated picture for DV derivation and the one for TMVP uses the same syntax. Besides, in scheme 2, one additional temporal neighboring block is checked to increase the probability of having an available DV. Experimental results show that both schemes have no overall coding performance changes compared to HTM-4.0.1 while the burden of memory access is reduced.

# Introduction

In the HEVC-based 3D video coding, HTM-4.0.1 [1], a disparity vector (DV) is derived to specify the location of a corresponding block in an inter-view reference picture for the inter-view motion parameter prediction and inter-view residual prediction. The DV is derived from the spatial or temporal neighboring blocks.

In HTM-4.0.1, to derive the DV from the temporal neighboring blocks in a temporal collocated picture, a PU-level adaptive collocated picture selection method is used to select the collocated picture for each prediction unit (PU):

* For AMVP, the temporal collocated picture is indicated by the reference index of the current PU.
* For the first PU of a coding unit (CU) that selects merge mode, the temporal collocated picture is indicated by the reference index of the left PU.
* For non-first PUs of the CU that selects merge mode, the temporal collocated picture is indicated by reference index equal to 0.

Under the current scheme, each PU may adaptively derive the DV from different collocated pictures as shown in Figure 1(a). However, since temporal reference data usually reside in DRAM, the decoder must read data of multiple consecutive LCUs in a single burst to satisfy the memory demands, and it must fetch this data before the coding tree unit (CTU) is processed to reduce the latency of memory access. Allowing the collocated picture to change at PU level may break this pipeline architecture, which can be a performance burden for decoders due to DRAM access limitations.



(a) (b)

**Figure 1. Collocated picture selection for DV derivation: (a) PU-level adaptive selection (HTM4.0.1); (b) unified collocated picture (proposed constraint)**

# Proposed schemes

In this contribution, we propose to apply a constraint to the selection of the temporal collocated picture in DV derivation to reduce the overhead for memory access. As shown in Figure 1(b), the constraint is that the temporal collocated picture for DV derivation from a disparity compensated prediction (DCP) coded block and the temporal collocated picture for DV derivation from a DV-motion compensated prediction (DV-MCP) coded block shall be the same for all PUs within one picture. Based on the proposed constraint, two different schemes are proposed and elaborated in the following subsections.

## Scheme 1:

In this scheme, the collocated picture in DV derivation is signaled using additional syntax in slice header in the same manner as the signaling of the collocated picture for temporal motion vector predictor (TMVP). Note that, at the encoder, the reference picture having the most DVs is set as the collocated picture for DV derivation. Besides, we include the DVs from merge coded DV-MCP blocks for DV derivation.

## Scheme 2:

In this scheme, the temporal collocated picture for DV derivation follows the one signaled in slice header for TMVP. The positions for spatial neighbour blocks are the same with those used in HTM4.0.1 as depicted in Figure 2(a). For spatial neighboring block A1, its temporal collocated block TA1 is also checked in addition to RB and Ctr position as depicted in Figure 2(b). The temporal collocated block is the block in the collocated picture which has the same position as that of the block in the current picture.

 

(a) (b)

**Figure 2. Positions of (a) spatial neighboring blocks; (b) temporal neighboring blocks**

The search on DCP blocks is processed in the order: A1, B1, B0, A0, B2, RB, Ctr, TA1; If DV is not found after checking DCP coded blocks, it continues to search DV from DV-MCP blocks in the order: A0, A1, TA1, B0, B1, B2, RB, Ctr.

# Experimental results

The proposed methods are integrated into HTM-4.0.1 [1], and all tests were conducted under common test conditions [2]. For both scheme 1 and scheme 2, no BD-rate changes are observed for view 1 and view 2 and also for overall coded and synthesized results. The detailed experimental results of each scheme are given in Table1 and Table 2, respectively.

Table 1. Results of scheme 1

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Video 1 | Video 2 | Video only | Synthesized only | Coded & synthesized | Enc time | Dec time | Ren time |
| Balloons | -0.1% | -0.3% | -0.1% | 0.0% | 0.0% | 100.6% | 99.9% | 99.8% |
| Kendo | 0.3% | 0.6% | 0.2% | 0.1% | 0.1% | 100.1% | 99.8% | 100.1% |
| Newspapercc | -0.3% | -0.3% | -0.1% | -0.1% | -0.1% | 99.8% | 100.1% | 100.4% |
| GhostTownFly | 0.2% | 0.3% | 0.1% | 0.1% | 0.1% | 99.9% | 99.8% | 100.4% |
| PoznanHall2 | 0.8% | 0.0% | 0.1% | 0.1% | 0.1% | 99.6% | 101.0% | 100.4% |
| PoznanStreet | -0.3% | -0.2% | -0.1% | -0.1% | -0.1% | 99.8% | 99.5% | 100.4% |
| UndoDancer | 0.2% | 0.3% | 0.1% | 0.0% | 0.0% | 98.6% | 99.6% | 101.2% |
| 1024x768 | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 100.2% | 100.0% | 100.1% |
| 1920x1088 | 0.2% | 0.1% | 0.0% | 0.0% | 0.0% | 99.5% | 99.9% | 100.6% |
| **average** | **0.1%** | **0.0%** | **0.0%** | **0.0%** | **0.0%** | **99.8%** | **99.9%** | **100.4%** |

Table 2. Results of scheme 2

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Video 1 | Video 2 | Video only | Synthesized only | Coded & synthesized | Enc time | Dec time | Ren time |
| Balloons | -0.1% | 0.0% | 0.0% | 0.0% | 0.0% | 98.9% | 98.2% | 99.8% |
| Kendo | 0.0% | 0.2% | 0.1% | 0.0% | 0.0% | 99.2% | 97.9% | 100.6% |
| Newspapercc | 0.0% | 0.2% | 0.1% | 0.1% | 0.1% | 101.0% | 101.5% | 100.3% |
| GhostTownFly | 0.1% | 0.1% | 0.0% | 0.0% | 0.0% | 103.8% | 85.5% | 87.5% |
| PoznanHall2 | 0.2% | 0.3% | 0.1% | 0.0% | 0.0% | 107.2% | 94.6% | 101.5% |
| PoznanStreet | -0.2% | -0.2% | 0.0% | -0.1% | -0.1% | 101.5% | 92.5% | 90.6% |
| UndoDancer | 0.1% | 0.2% | 0.1% | 0.0% | 0.0% | 100.5% | 96.8% | 101.7% |
| 1024x768 | 0.0% | 0.2% | 0.0% | 0.0% | 0.0% | 99.7% | 99.2% | 100.2% |
| 1920x1088 | 0.0% | 0.1% | 0.0% | 0.0% | 0.0% | 103.2% | 92.2% | 95.1% |
| **average** | **0.0%** | **0.1%** | **0.0%** | **0.0%** | **0.0%** | **101.7%** | **95.2%** | **97.3%** |

# Conclusion

This contribution proposes to unify the temporal collocated picture for DV derivation for all PUs within the same picture. With the proposed constraint, the overhead for memory access caused by PU adaptive collocated picture switching can be avoided. Two schemes are proposed to use different ways to signal the temporal collocated picture for DV derivation. For scheme 1, additional syntax is used to signal the temporal collocated picture for DV derivation. For the scheme 2, the signaled collocated picture for TMVP is also used for the DV derivation. Besides, one additional position at collocated picture is checked to derive the DV. Experimental results show that both schemes introduce no BD rate change for overall coding performance. It is recommended to adopt either one of the proposed two schemes into the next HTM to reduce the memory access overhead while maintaining competitive coding performance.

# Patent rights declaration (s)

**MediaTek Inc. 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).**

# References

1. HTM-4.0.1, <https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware/tags/HTM-4.0.1/>
2. D. Rusanovskyy, K. Müller, A. Vetro, “Common test conditions of 3DV Core Experiments,” Document of Joint Collaborative Team on 3D Video Coding Extension Development, JCT3V-A1100, July, 2012.