|  |  |
| --- | --- |
| **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**  3rd Meeting: Geneva, CH, 17–23 Jan. 2013 | Document: JCT3V-C0022 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **3D-CE2: Summary Report on Disparity Vector Derivation** | | |
| *Status:* | Input Document | | |
| *Purpose:* | Core Experiment Report | | |
| *Author(s) or Contact(s):* | Ying Chen  5775 Morehouse Drive San Diego, CA 92121 USA | Tel: Email: | +1-858-845-6589 [cheny@qti.qualcomm.com](mailto:cheny@qti.qualcomm.com) |
| *Source:* | Qualcomm Incorporated | | |

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

# Abstract

This document summarizes the activities of the Core Experiment (CE) 2 on disparity vector derivation. During the 2nd JCT-3V meeting, contributions have been proposed to both 3D-AVC and 3D-HEVC for disparity vector derivation.

Experiments are related to 3D-ATM and 3D-HTM. Ten companies were involved in this CE. There are two CE2.a proposals and one CE2.a related proposal for 3D-ATM. It is also reported that there are no CE2.h proposals for 3D-HTM. However, there are six CE2.h related proposals for 3D-HTM.

# Participants

|  |  |  |  |
| --- | --- | --- | --- |
| ***Participant*** | ***Contact*** | ***Email*** | ***Type*** |
| ETRI | Gun Bang | [gbang@etri.re.kr](mailto:gbang@etri.re.kr) | P/C |
| KHU | Kyung Yong Kim | [kimky@khu.ac.kr](mailto:kimky@khu.ac.kr) | P/C |
| MediaTek | Jian-Liang Lin  Yi-Wen Chen  Yu-Wen Huang | [jl.lin@mediatek.com](mailto:jl.lin@mediatek.com)  [yiwen.chen@mediatek.com](mailto:yiwen.chen@mediatek.com)  [yuwen.huang@mediatek.com](mailto:yuwen.huang@mediatek.com) | P/C |
| Samsung | Jin Young Lee | [jinyoung79.lee@samsung.com](mailto:jinyoung79.lee@samsung.com) | P/C |
| LG | Jaewon Sung | [jw.sung@lge.com](file:///C:\Users\cheny\AppData\Local\Temp\Rar$DI00.838\jw.sung@lge.com) | P/C |
| Qualcomm | Jewon Kang  Li Zhang | [jewonk@qti.qualcomm.com](mailto:jewonk@qti.qualcomm.com)  [lizhang@qualcomm.com](mailto:lizhang@qualcomm.com) | P/C |
| INRIA | Laurent Guillo | [Laurent.guillo@inria.fr](mailto:Laurent.guillo@inria.fr) |  |
| NTT | Shinya Shimizu | [shimizu.shinya@lab.ntt.co.jp](mailto:shimizu.shinya@lab.ntt.co.jp) | C |
| Orange | Elie Mora  Joel Jung | [elie.mora@orange.com](mailto:elie.mora@orange.com)  [joelb.jung@orange.com](mailto:joelb.jung@orange.com) |  |
| Nokia | Dmytro Rusanovskyy | [Dmytro.Rusanovskyy@nokia.com](mailto:Dmytro.Rusanovskyy@nokia.com) | P/C |

(P = proponent, C = cross checker)

# Tools under investigation

## CE2.a: Disparity Vector Derivation in ATM

The scope of this CE for 3D-AVC is to simplify the disparity vector derivation process by e.g., less memory access to the depth map and to potentially improve the coding efficiency.

Two CE proposals are described below.

### JCT3V-C0122: Simplification of disparity vector derivation

This contribution proposes a simplified disparity vector derivation method for saving the memory access bandwidth. The implementation of the proposed process can be achieved with sharing a depth sample to derivate the disparity vector in all sub-blocks inside each MB. It is indicated that the proposed approach reduces searching and comparison operations to find maximum depth samples in every sub-blocks and additionally reduces of the encoder complexity.

Negligible coding loss was observed by changing the disparity vector value from the maximum value of the four corners of the depth map to the value of one corner (e.g., bottom-right).

Results for non-CTC (Common Test Condition) cases, wherein texture and depth view components have the same resolution, are also provided, as shown in Table 1.

This proposal is cross verified by Nokia (JCT3V-C0174), Samsung (JCT3V-C0213) and Qualcomm (JCT3V-C0226).

### JCT3V-C0133: Memory access reduction for skip mode from the base view

In this contribution, it is proposed to derive the inter-view candidate from a single corresponding block located by a disparity vector derived for list 0 and list 1 to reduce the memory access bandwidth for the inter-view motion data fetch, instead of locating two different blocks for list 0 and list 1 separately as in the current design.

This proposal is cross verified by NTT (JCT3V-C0184).

## CE2.h: Disparity Vector Derivation in HTM

No input contributions. The two proposals included in this CE during the 2nd JCT-3V meeting in Shanghai have been discontinued.

# Experiments and results

## CE2.a: Disparity Vector Derivation in ATM

It is considered that these two proposals, although both targeting at complexity and memory access reduction, are designed to address different aspects of the disparity vector derivation. Therefore, these two proposals are not compared in this report. Results of these two proposals are presented in Table 1 and Table 2 respectively.

JCT3V-C0122 provides negligible coding loss (0.01% to 0.04%) with reduced memory accesses, under CTC and symmetric texture/depth resolution configuration.

JCT3V-C0133 reduces the number of function calls for disparity derivation with no loss (even very small gain) under CTC.

Table 1. Summary of the results of JCT3V-C0122

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Texture Coding | | Depth Coding | | Total (Coded PSNR) | | Total  (Synthesized PSNR) | | Complexity estimate (ratio to anchor) | |
|  | dBR | dPSNR | dBR | dPSNR | dBR | dPSNR | dBR | dPSNR | Enc. time | Dec. time |
| Test.1 H | 0.06 | 0.00 | 0.00 | 0.00 | 0.05 | 0.00 | 0.03 | 0.00 | 100% | 100% |
| Test.1 F | 0.08 | 0.00 | 0.00 | 0.00 | 0.07 | 0.00 | 0.03 | 0.00 | 100% | 100% |
| Test.2 H | 0.01 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | 0.01 | 0.00 | 99% | 100% |
| Test.2 F | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100% | 101% |
| Test.3 H | 0.05 | 0.00 | 0.00 | 0.00 | 0.05 | 0.00 | 0.02 | 0.00 | 99% | 99% |
| Test.3 F | 0.08 | 0.00 | 0.00 | 0.00 | 0.07 | 0.00 | 0.04 | 0.00 | 100% | 101% |

Table 2. Summary of the results of JCT3V-C0133.



## CE2.h: Disparity Vector Derivation in HTM

No input contributions.

# Related proposals

Input documents coordinated under this CE include one CE2.a related proposal and six CE2.h related proposals.

## CE2.a related proposals for 3D-ATM

Only one CE2.a proposal (JCT3V-C0134 from MediaTek and Samsung) is proposed.

### JCT3V-C0134: MB-level depth-to-DV conversion in ATM

In the current 3D-ATM, it is proposed that the depth-to-DV conversion needs to be performed multiple times in motion vector prediction to derive DVs for various partition blocks in one MB.

In this document, it is proposed that all partition blocks of Inter mode within the same MB share the same DV as the one used for the inter-view MVP derivation in Skip/Direct mode. Therefore, the disparity derivation can be done in the MB-level instead of block-level.

The depth-to-DV conversion is still performed once by deriving the DV from a maximum depth value of four corner samples of the associated depth block.

This proposal was crosschecked by ETRI (JCT3V-Cxxxx).

Table 3. Summary of the results of JCT3V-C0134.



## CE2.h related proposals for 3D-HTM

Five input contributions are further improvements or simplifications for the current disparity vector derivation in CE, under the current common test condition (CTC).

One input document is for the flexible coding order (FCO) scenario, which is currently not part of CTC.

### List of input documents for CE2.h related

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Participants** | **Doc No.** | **Title** | **Type** | **Cross check** |
| Samsung | [JCT3V-C0097](http://phenix.it-sudparis.eu/jct2/doc_end_user/current_document.php?id=536) | 3D-CE2.h related results on disparity vector derivation | CTC | JCT3V-C0197 |
| Qualcomm | [JCT3V-C0050](http://phenix.it-sudparis.eu/jct2/doc_end_user/current_document.php?id=488) | 3D-CE2.h related: Enhanced disparity vector derivation | CTC | JCT3V-C0105 |
| MediaTek &PKU | [JCT3V-C0135](http://phenix.it-sudparis.eu/jct2/doc_end_user/current_document.php?id=576) | 3D-CE2.h related: Restriction on the temporal blocks for memory bandwidth reduction in DV derivation | CTC | JCT3V-C0188 |
| MediaTek, HIT&PKU | [JCT3V-C0141](http://phenix.it-sudparis.eu/jct2/doc_end_user/current_document.php?id=582) | 3D-CE2.h related: Improved DV searching order | CTC | JCT3V-C0206 |
| LG | [JCT3V-C0117](http://phenix.it-sudparis.eu/jct2/doc_end_user/current_document.php?id=556) | 3D-CE2.h related: Motion Aware Temporal Disparity Vector Derivation | CTC | JCT3V-C0207 |
| Nokia | [JCT3V-C0170](http://phenix.it-sudparis.eu/jct2/doc_end_user/current_document.php?id=616) | CE2.h-related: Flexible Coding Order (FCO) and depth-first-based disparity derivation | FCO | JCT3V-C0219 |

### Description of CE2.h CTC related proposals

The five proposals are related to coding efficiency improvements and complexity reductions as well as clean-ups for the current Neighboring Block based Disparity Vector derivation (NBDV) scheme.

#### JCT3V-C0097: Results on disparity vector derivation

This document proposes a solution which contains two aspects: one is to improve the coding efficiency of the NBDV, while the other is to further simplify the NBDV. This solution provides an overall loss of 0.1%. The two aspects of the proposals are described as follows:

1. Priority order change for more efficient NBDV: In 3D-HTM, when a first available disparity motion vector of the neighboring blocks is identified, the disparity motion vector is returned as the final disparity vector. In 3D-HTM, disparity motion vectors from spatial neighboring blocks are firstly checked, followed by those in the temporal neighboring blocks. This proposal propose to change the order so that the disparity motion vectors in the temporal blocks are firstly checked, followed by those in spatial neighboring blocks.
2. Removal of derived disparity vectors (DDV): In 3D-HTM, in addition to the motion vectors of neighboring blocks, the derived disparity vectors of the neighboring blocks may also be used to derive the disparity vector. Such vectors are derived disparity vectors DDVs (also known as DV-MCP). It is proposed that checking of DDVs are removed thus they do not need to be stored.

#### JCT3V-C0050: Enhanced disparity vector derivation

Multiple disparity candidates based derivation: this document proposes a solution which improves the coding efficiency of the NBDV. Instead of always getting the first available disparity vector, multiple (up to three) disparity vector candidates can be generated and used to derive the final disparity vector. Depending on the number of candidates, the mathematical function to derive the final disparity vector can be either maximum or median operation.

#### JCT3V-C0135: Simplifications on disparity vector derivation

This proposal provides three simplifications for NBDV:

1. Consistent checking order of temporal blocks for all views: in NBDV, the checking order or the temporal blocks may be different based on whether the current non-base view is a left view or right view. It is proposed the checking order should be the same for all non-base views and the bottom-right block is checked earlier than the center block.
2. Remove the top-left temporal block (SW only): due to the reason that the 3D-HTM software was based on an earlier version of HM, the so-called top-left block might still be checked in some cases in the current NBDV solution, although this is not present in the 3D-HEVC draft text. It is proposed to remove the checking of top-left block.
3. Prohibit the access of the bottom-right temporal block residing at a lower CTU row: when a bottom-right block is in a lower CTU row of the co-located CTU, it is considered as unavailable.

#### JCT3V-C0141: Improved DV searching order

Priority order change for more efficient NBDV: the proposal presents a priority order of the searched neighboring blocks, so that the motion vectors in temporal neighboring blocks are first checked, followed by those in spatial neighboring blocks.

It is reported that proposal is the same as JCT3V-C0097, part #1 (denoted as P1 in table 4).

#### JCT3V-C0117: Motion Aware Temporal Disparity Vector Derivation

Temporal neighbor blocks adaptive changing in NBDV: It is proposed that when one of the spatial neighbour blocks is referencing the candidate picture (a temporal reference picture used for disparity vector derivation), the locations of the temporal neighbour blocks are changed by the amount of motion vector. The disparity vector derivation process works as follows

* Modified step 1): During step1, temporal motion vector of spatial neighbors and the POC value of the reference picture that the temporal motion vector is referencing are saved to a temporal array mvCand[] and pocCand[].
* Modified step2): The Step 2) is processed after the following two steps: 2-1 and 2-2. For each candidate picture whose POC is pocCandPic, the following applies.
* Step 2-1) Motion vector estimation:
  + Among the mvCand[i] (i=0,…,numCandMv-1), the first mvCand[i] whose corresponding pocCand[i] is equal to pocCandPic is searched and set to mvEst.
  + If the mvEst is not found and current candidate picture is the 2nd candidate picture, then the mvCand[0] is scaled (as is done for MVP) and set to mvEst.
  + If the mvEst is not found, mvEst is set to (0,0)
* Step 2-2) Modification of temporal neighbour block positions:
  + If the mvEst is found, the positions of CT and BR are changed as follows:
    - CT' = CT + mvEst, BT' = BR + mvEst
* Step 3) is applied to both CT and BR block positions by changing them to CT' and BR'.

#### Comparison for the above proposals

Proposals (including different aspects of proposals) are considered in two categories: coding efficiency category and simplification category.

##### Coding efficiency proposals

Table 4. Summary of the coding efficiency proposals.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Proposals** | **Video 0** | **Video 1** | **Video 2** | **video only** | **synthesized only** | **coded & synthesized** |
| JCT3V-C0097P1 | 0.0% | -0.3% | -0.1% | -0.1% | -0.1% | -0.1% |
| JCT3V-C0050 | 0.0% | -0.5% | -0.4% | -0.2% | -0.1% | -0.1% |
| JCT3V-C0141 | 0.0% | -0.3% | -0.1% | -0.1% | -0.1% | -0.1% |
| JCT3V-C0117 | 0.0% | -0.2% | -0.1% | -0.1% | -0.0% | -0.0% |

##### Complexity reduction proposals

Table 5. Summary of the simplification proposals

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Proposals** | **Video 0** | **Video 1** | **Video 2** | **video only** | **synthesized only** | **coded & synthesized** |
| JCT3V-C0097P2 | 0.0% | 0.3% | 0.6% | 0.2% | 0.1% | 0.1% |
| JCT3V-C0135P1 | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| JCT3V-C0135P2 | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| JCT3V-C0135P3 | 0.0% | 0.0% | 0.3% | 0.1% | 0.0% | 0.0% |
| JCT3V-C0135 | 0.0% | 0.1% | 0.2% | 0.1% | 0.0% | 0.0% |

### Description of the CE2.h FCO related proposal

#### JCT3V-C0170: FCO and depth-first-based disparity derivation

This proposal studies six methods for disparity vector derivation.

Method 1: It uses synthesized depth value of the non-base view for disparity vector derivation. It may require warping and hole-filing which may be done in the picture level.

Method 2: Depth estimates based solution. It is reported that method 2 is the disparity derivation in the HTM software in a version earlier than HTM 3.

Method 3: It is the NBDV in the current HTM software.

Method 4: It sounds that method 4 is a simplification of the current NBDV, while not considering depth information from any depth view components. This method reduces number of parameters required for validation during the disparity derivation. This requires further clarification. Results for method 4 are not present in the first version of JCT3V-C0170.

Method 5: It is as proposed in JCT3V-B0090 (the same as JCT3V-C0131) and considered under CE1.h. This method further considers depth information of the base view on top of NBDV.

Method 6: The disparity derivation is straightforward: accessing depth samples which being coded earlier and converting them to disparity through a look up table search. This method requires FCO.

It is reported that method 2, 3 and 4 are compatible to texture only multiview codec, while the other methods are not stereo/multivew compatible.

Method 1 and method 2 may require methods similar to view synthesis prediction so the complexity increase and implementation efforts are considered significant.

Method 3 and 4 have relatively higher complexity compared to method 6, due to the checks of neighboring blocks. As shown in method 4, potential complexity reduction is possible.

Method 5 may potentially increase the coding efficiency under the context of view synthesis prediction.

Method 6 only requires depth to disparity conversion thus may be considered as low complex. It requires flexible coding order and is not stereo/multiview compatible.

# Recommendations

To review the relevant input contributions.