|  |  |
| --- | --- |
| **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-C0112 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **3D-CE1.h related : Adaptive method for Depth-oriented Neighboring Block Disparity Vector** | | |
| *Status:* | Input Document | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Jiwook Jung Jaewon Sung Sehoon Yea | Tel: Email: | +82-2-6912-6477 jiwook.jung@lge.com |
| *Source:* | LG Electronics | | |

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

# Abstract

In the last meeting, Depth Oriented Neighboring Block Disparity Vector (DoNBDV) was proposed. The disparity vector candidates for merge and AMVP could be refined by the method. In the proposal, they were generated from the maximum value of the coded depth. However, It is not clear why the maximum value of the depth was used for the derivation of disparity. In this contribution, the Most Frequent Disparity (MFD) value is used adaptively for some cases instead of the disparity derived from the maximum value of the depth block. These methods are applied to the AMVP mode and the merge mode. Compared to DoNBDV, experimental results reportedly show -0.6% and 0.1% BD-Rate changes for video 1 and video 2, -0.1% BD-Rate gains is achieved for coded and synthesized views. Compared to HTM-5.0.1, experimental results reportedly show -2.0% and -1.5% BD-Rate changes for video 1 and video 2 with encoding time 102.0% and decoding time 102.8%.

# Introduction

In 3D-HEVC, the disparity vector can be used for inter-view merge/AMVP, residual prediction, …,etc. Disparity vector candidates can be refined by associated depth. In JCT3V-B0090[1], NBDV was refined using the maximum of virtual depth. It was called as DoNBDV (Depth-oriented Neighboring Block Disparity Vector). In 3DV-ATM track, similar method was adopted for disparity vector derivation when disparity vector candidate was not available.[2] The disparity vector derived by the maximum depth(MaxD) has the benefit that the object associated with the maximum depth value might not be occluded. However, the maximum depth can be inappropriate for the representative value of the depth block in some cases. For example, when the differences between maximum depth value and most of depth samples in the depth block are relatively big, the derived disparity vector can be inefficient. The Most Frequent Disparity (MFD) value can be used as an alternative for a disparity vector candidate. In this contribution, an adaptive method selecting representative value of depth block is presented.

# Proposed Method

When the differences between maximum depth value and most of depth samples in the depth block are relative big, the Most Frequent Disparity (MFD) derived from the depth block can be used as an alternative for disparity vector. The virtual depth block is retrieved by the same method as JCT3V-B0090 [1]. MFD value and DoNBDV can be derived from the virtual depth block by the following processes.

## Proposed Method1

Derivation process of MFD and MaxD

1. Find the location of the virtual depth block using NBDV (method in current 3D-HTM).
2. Convert depth value in the virtual depth block to disparity and build the histogram of the disparity.
3. Search for the MFD value of the disparity histogram and MaxD.

Adaptive selection process for DoNBDV

1. Read the number of MaxD from the disparity histogram. 🡪 NumMaxD
2. Occlusion area is calculated by the below equation.
3. If (NumMaxD + ) >= ((BlkWidth \* BlkHeight)>>1), MaxD is used for DoNBDV
4. Else, MFD is used for DoNBDV.

|  |
| --- |
| AreaOcc=(abs(MaxD-MFD)>> 2)\*BlkHeight  if ( NumMaxD + AreaOcc < (BlkWidth \* BlkHeight)>>1 )  DoNBDV = MFD  else  DoNBDV = MaxD |

## Proposed Method2

Derivation process of MFD and MaxD is the same as the proposed method1.

Adaptive selection process for DoNBDV

1. If the difference between original NBDV and MFD is smaller than the difference between NBDV and MaxD, then the MFD is used for DoNBDV.
2. Else, MaxD is used for DoNBDV.

|  |
| --- |
| if (abs( MFD - NBDV ) < abs (MaxD - NBDV )  DoNBDV = MFD  else  DoNBDV = MaxD |

## Proposed method3 (Simplification of proposed method2)

It can be relatively complex to build histogram and search for the maximum value and MFD value in each block. Therefore, complexity reduction is necessary. Those values can be extracted for only selected pixels, not all the pixels. The selected pixels are sampled by the positions in the block. For example, the pixels in a 8x8 block can be selected as figure2.

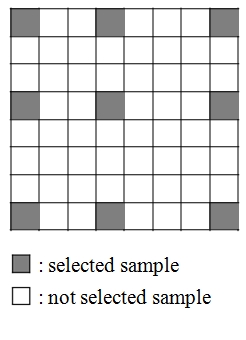


Figure2. Example of selected samples in a 8x8 block

In this 8x8 block case, The values for DoNBDV can be extracted using only 9 pixels, not using whole 64 pixels in the block.

## Proposed method4 (Simplification of conventional DoNBDV)

For method4, only MaxD (No MFD) is used for DoNBDV as JCT3V-B0090. However, the maximum depth value is selected from the 5 pixels as figure3.

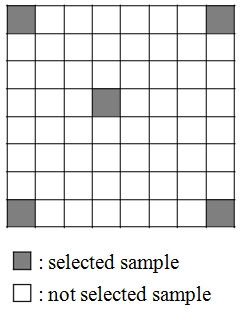


Figure3. Example of selected samples in a 8x8 block

# Experimental results

The methods were integrated into HTM-5.1. Four test was conducted.

## Proposed Method1

Table1. BD-rate changes of the proposed method1, Anchor : HTM-5.1

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | video 0 | video 1 | video 2 | video only | synthesized only | coded & synthesized | enc time | dec time |
| Balloons | 0.0% | -1.5% | 0.0% | -0.3% | -0.2% | -0.3% | 102.5% | 104.3% |
| Kendo | 0.0% | -2.0% | -1.1% | -0.6% | -0.5% | -0.5% | 101.8% | 102.9% |
| Newspapercc | 0.0% | -0.9% | -0.7% | -0.3% | -0.3% | -0.3% | 103.1% | 100.3% |
| GhostTownFly | 0.0% | -4.5% | -4.0% | -1.2% | -0.9% | -1.0% | 99.6% | 98.2% |
| PoznanHall2 | 0.0% | -1.3% | -1.5% | -0.6% | -0.3% | -0.4% | 98.2% | 97.9% |
| PoznanStreet | 0.0% | -1.8% | -1.9% | -0.6% | -0.5% | -0.5% | 101.5% | 105.0% |
| UndoDancer | 0.0% | -3.4% | -2.6% | -1.1% | -0.7% | -0.8% | 101.3% | 109.1% |
| 1024x768 | 0.0% | -1.5% | -0.6% | -0.4% | -0.4% | -0.4% | 102.5% | 102.5% |
| 1920x1088 | 0.0% | -2.7% | -2.5% | -0.9% | -0.6% | -0.7% | 100.1% | 102.4% |
| **average** | **0.0%** | **-2.2%** | **-1.7%** | **-0.7%** | **-0.5%** | **-0.6%** | **101.1%** | **102.5%** |

## Proposed Method2

Table2. BD-rate changes of the proposed method2, Anchor : HTM-5.1

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | video 0 | video 1 | video 2 | video only | synthesized only | coded & synthesized | enc time | dec time |
| Balloons | 0.0% | -1.3% | -0.2% | -0.3% | -0.2% | -0.2% | 102.5% | 109.3% |
| Kendo | 0.0% | -1.7% | -1.4% | -0.6% | -0.5% | -0.5% | 101.8% | 110.3% |
| Newspapercc | 0.0% | -1.0% | -0.7% | -0.3% | -0.3% | -0.3% | 103.2% | 104.8% |
| GhostTownFly | 0.0% | -4.3% | -3.7% | -1.2% | -0.8% | -1.0% | 100.0% | 101.1% |
| PoznanHall2 | 0.0% | -0.7% | -2.0% | -0.6% | -0.3% | -0.4% | 98.2% | 104.8% |
| PoznanStreet | 0.0% | -1.6% | -1.7% | -0.6% | -0.4% | -0.5% | 101.4% | 107.5% |
| UndoDancer | 0.0% | -2.8% | -1.8% | -0.9% | -0.6% | -0.7% | 101.3% | 103.6% |
| 1024x768 | 0.0% | -1.3% | -0.7% | -0.4% | -0.3% | -0.4% | 102.5% | 108.1% |
| 1920x1088 | 0.0% | -2.3% | -2.3% | -0.8% | -0.5% | -0.6% | 100.2% | 104.2% |
| **average** | **0.0%** | **-1.9%** | **-1.6%** | **-0.6%** | **-0.5%** | **-0.5%** | **101.2%** | **105.9%** |

## Proposed Method3

Table3. BD-rate changes of the proposed method3, Anchor : HTM-5.1

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | video 0 | video 1 | video 2 | video only | synthesized only | coded & synthesized | enc time | dec time |
| Balloons | 0.0% | -1.3% | -0.3% | -0.3% | -0.3% | -0.3% | 101.1% | 100.4% |
| Kendo | 0.0% | -1.8% | -1.5% | -0.6% | -0.6% | -0.6% | 100.6% | 104.7% |
| Newspapercc | 0.0% | -0.8% | -0.7% | -0.3% | -0.3% | -0.3% | 101.3% | 101.0% |
| GhostTownFly | 0.0% | -4.1% | -3.4% | -1.1% | -0.8% | -0.9% | 98.6% | 99.0% |
| PoznanHall2 | 0.0% | -0.7% | -1.4% | -0.5% | -0.2% | -0.3% | 96.1% | 101.3% |
| PoznanStreet | 0.0% | -1.5% | -1.7% | -0.6% | -0.4% | -0.4% | 99.2% | 103.2% |
| UndoDancer | 0.0% | -2.4% | -1.5% | -0.8% | -0.6% | -0.7% | 99.6% | 100.9% |
| 1024x768 | 0.0% | -1.3% | -0.8% | -0.4% | -0.4% | -0.4% | 101.0% | 102.0% |
| 1920x1088 | 0.0% | -2.2% | -2.0% | -0.7% | -0.5% | -0.6% | 98.4% | 101.1% |
| **average** | **0.0%** | **-1.8%** | **-1.5%** | **-0.6%** | **-0.5%** | **-0.5%** | **99.5%** | **101.5%** |

## Proposed Method4

Table4. BD-rate changes the proposed method4, Anchor : HTM-5.0.1

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | video 0 | video 1 | video 2 | video only | synthesized only | coded & synthesized | enc time | dec time |
| Balloons | 0.0% | -1.2% | -1.3% | -0.5% | -0.4% | -0.4% | 100.9% | 99.6% |
| Kendo | 0.0% | -1.8% | -2.0% | -0.8% | -0.7% | -0.7% | 100.3% | 98.5% |
| Newspapercc | 0.0% | -0.7% | -0.8% | -0.3% | -0.3% | -0.3% | 101.3% | 103.1% |
| GhostTownFly | 0.0% | -3.8% | -3.5% | -1.1% | -0.8% | -0.9% | 98.5% | 97.0% |
| PoznanHall2 | 0.0% | -0.1% | -2.1% | -0.5% | -0.3% | -0.3% | 95.9% | 100.9% |
| PoznanStreet | 0.0% | -1.3% | -1.6% | -0.5% | -0.4% | -0.4% | 99.4% | 97.1% |
| UndoDancer | 0.0% | -2.4% | -1.5% | -0.8% | -0.4% | -0.5% | 99.5% | 102.4% |
| 1024x768 | 0.0% | -1.2% | -1.3% | -0.5% | -0.5% | -0.5% | 100.8% | 100.4% |
| 1920x1088 | 0.0% | -1.9% | -2.2% | -0.7% | -0.5% | -0.5% | 98.3% | 99.3% |
| **average** | **0.0%** | **-1.6%** | **-1.8%** | **-0.6%** | **-0.5%** | **-0.5%** | **99.4%** | **99.8%** |

Table5. BD-rate changes the proposed method4, Anchor : JCT3V-C0131

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | video 0 | video 1 | video 2 | video only | synthesized only | coded & synthesized | enc time | dec time |
| Balloons | 0.0% | -0.1% | -0.3% | -0.1% | -0.1% | -0.1% |  |  |
| Kendo | 0.0% | -0.2% | -0.1% | -0.1% | -0.1% | -0.1% |  |  |
| Newspapercc | 0.0% | -0.2% | -0.1% | 0.0% | 0.0% | 0.0% |  |  |
| GhostTownFly | 0.0% | 0.1% | -0.1% | 0.0% | 0.0% | 0.0% |  |  |
| PoznanHall2 | 0.0% | 0.0% | 0.3% | 0.1% | 0.1% | 0.1% |  |  |
| PoznanStreet | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |  |  |
| UndoDancer | 0.0% | 0.0% | -0.1% | 0.0% | 0.0% | 0.0% |  |  |
| 1024x768 | 0.0% | -0.2% | -0.2% | -0.1% | -0.1% | -0.1% |  |  |
| 1920x1088 | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |  |  |
| **average** | **0.0%** | **-0.1%** | **-0.1%** | **0.0%** | **0.0%** | **0.0%** |  |  |

# Conclusion

In this contribution, methods to improve DoNBDV are proposed. The disparity from the maximum of virtual depth block can be inefficient for representative value of the block in some cases. The Most Frequent Disparity value is selected as an alternative adaptively in method 1&2 . Method 3&4 are the simplifications of two DoNBDV methods. Although average coding gains of methods are not significant, there are noticeable coding gain for some cases.

It is recommended that one of proposed methods is adopted into 3D-HEVC.

# Reference

1. Y.-L. Chang, C.-L. Wu, Y.-P. Tsai, S. Lei (MediaTek), "3D-CE5.h related: Depth-oriented Neighboring Block Disparity Vector (DoNBDV) with virtual depth," Joint Collaborative Team on 3D Video Coding Extension Development (JCT-3V) of ITU-T VCEG and ISO/IEC MPEG JCT3V-B0090, Shanghai, China, October, 2012.
2. Jin Young Lee, Tadashi Uchiumi, Jaejoon Lee, Yoshiya Yamamoto, Du-Sik Park, " 3D-CE5.a results on joint proposal for an improved depth-based motion vector prediction method by Samsung and Sharp", ISO/IEC JTC1/SC29/WG11, M24824, Geneva, May 2012.

# Patent rights declaration(s)

**LG Electronics 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).**