|  |  |
| --- | --- |
| **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**  1st Meeting: Stockholm, SE, 16–20 July 2012 | Document: JCT2-A0098 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **3D-CE6.h related: Depth Modeling Mode (DMM) 3 simplification for HTM** | | |
| *Status:* | Input Document | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Xin Zhao Ying Chen  Li Zhang Marta Karczewicz  5775 Morehouse Drive San Diego, CA 92121 USA | Tel: Email: | +86-10-5776-0696 [xinzhao@qualcomm.com](mailto:xinzhao@qualcomm.com)  +1-858-845-6589 [cheny@qualcomm.com](mailto:cheny@qualcomm.com)  +1-858-651-6660 [lizhang@qualcomm.com](mailto:lizhang@qualcomm.com) |
| *Source:* | Qualcomm Incorporated | | |

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

# Abstract

This contribution proposes a method to simplify Depth Modeling Mode (DMM) 3 in HTM. In DMM mode 3, Wedgelet pattern is not signalled, but derived by an exhaustive search from a huge number of candidates at the decoder. In this proposal, it is firstly proposed that DMM mode 3 is disabled when the top-left 4×4 block in Co-located Texture Luma Block (CTLB) is not coded as Intra. Furthermore, it is proposed that the Intra mode within the CTLB is used to derive the starting point for Wedgelet pattern searching with a fixed search range. It is asserted that the proposal significantly reduces the number of Wedgelet candidates to be searched at the decoder for DMM mode 3, without changing the coding efficiency of the current DMM.

# Introduction

Depth map is characterized by sharp edges and large areas of constant samples. To better accommodate the characteristics of depth map, in DMM modes, a depth prediction unit (PU) is partitioned into two regions. There are two DMM partitioning methods, namely Wedgelet and Contour respectively, as shown in Table 1. Each partition of a PU is predicted by a constant value. A Wedgelet partition pattern can be either explicitly signaled (DMM mode 1), predicted by spatially neighboring blocks (mode 2), or derived by CTLB (mode 3 and mode 4). As depicted in Fig.1, for a Wedgelet pattern, a depth block is partitioned into two regions by a straight line (i.e., partition 0/1 where pixels are labeled by 0/1), and for a Contour pattern, a depth block can be partitioned into two irregular regions.

 

Fig. : Example of Wedgelet pattern (left) and Contour pattern (right)

In DMM, depending on the type of the DMM pattern and how it is signaled, four modes [1] are defined in DMM as listed in Table 1.

Table : Four modes defined in DMM

|  |  |  |  |
| --- | --- | --- | --- |
|  | Explicit Signaling | Predicted in Spatial | Derived by Texture |
| Wedgelet pattern | Mode 1 | Mode 2 | Mode 3 |
| Contour pattern |  |  | Mode 4 |

The usage of the four modes depends on the PU size, which is specified in Table 2.

Table 2: Usage of DMM modes for different PU sizes

|  |  |
| --- | --- |
| **PU size** | **Available DMM modes** |
| 4x4 | Mode 1, Mode 3 |
| 8x8 | Mode 1, Mode 2, Mode 3, Mode 4 |
| 16x16 | Mode 1, Mode 2, Mode 3, Mode 4 |
| 32x32 | Mode 1, Mode 2, Mode 3, Mode 4 |
| 64x64 | None |

Table 3: Number of available Wedgelet patterns for different PU sizes in HTM3.1

|  |  |
| --- | --- |
| **PU size** | **Available DMM modes** |
| 4x4 | 86 |
| 8x8 | 782 |
| 16x16 | 1394 |
| 32x32 | 1503 |
| 64x64 | None |

In DMM mode 3, the Wedgelet pattern is not signalled or predicted, but derived by exhaustively searching all available Wedgelet patterns (up to around 1500, as listed in Table 3) on the reconstructed CTLB at the decoder. During each iteration, the distortion based on one Wedgelet pattern is calculated. The average value of the components in CTLB with the corresponding components in the Wedgelet pattern labeled by “0” is calculated as *P0*, and the average value of the components in the CTLB with the corresponding components in the Wedgelet pattern labeled by “1” is calculated as *P1*. Then a prediction block *P* is generated with the same size of the CTLB, the component of *P* is valued by *P0* if the corresponding component in Wedgelet pattern is labeled with 0, otherwise, it is valued by *P1*. After that, the distortion is calculated as the sum of squared difference between each component of CTLB and *P*. When the exhaustive search terminates, the Wedgelet pattern with minimum distortion value is identified as the pattern applied on current depth block. Therefore, there is a significant increase of decoder complexity using DMM mode 3.

The silhouette of a texture region and that of the corresponding depth region is correlated. Wedgelet patterns are used to present the main edge information of one block and can be classified into several sets based on the edge information. Therefore, there is correlation between the prediction angle of the Intra mode of texture block and the Wedgelet partition of the corresponding depth block.

# Proposed Solution

To reduce the complexity of DMM mode 3, in this contribution, it is proposed to enable DMM mode 3 only when the top-left 4×4 block in CTLB is Intra-coded, otherwise it is disabled. When DMM mode 3 is enabled, only a selected list of Wedgelet patterns is used as the search candidates.

## Wedgelet index list generation for each Intra prediction mode

After Wedgelet patterns are created as in the current HTM3.1, the Wedgelet index list is generated for each Intra prediction mode (IPM). To generate the Wedgelet list, each Wedgelet pattern *idxW* is first mapped to an IPM *optIdx* by applying the following:

– For each IPM *i* from 2 to 34, (*Hi*, *Vi*) is obtained based on Table 4. And a value *D* is calculated for each IPM by the below arithmetic:

*D*[*i*] = | *Vi*×(*Xs*-*Xe*) - *Hi*×(*Ye*-*Ys*) |

– Map the Wedgelet pattern *idxW* to the IPM *idxOpt* which minimizes *D* among all IPMs. One example of mapping a Wedgelet pattern to IPM 22is shown in Fig.2.

The Wedgelet index list for IPM equal to 0 and 1 is simply set to contain only one Wedgelet index 0, corresponding to the pattern that has top-left sample in one partition and all the remaining pixels in the other partition. Then for each IPM *pidx*, all the Wedgelet patterns with mapped IPM *widx* satisfying | *widx* - *pidx* | ≤ *tr*, are included in a Wedgelet list denoted by *WdgIdxLst*[*pidx*], where *tr* is a set to 2 in this proposal.



Fig. 2: Example of mapping from a Wedglet pattern to Intra prediction mode 22.

Note that similar to Wedgelet pattern generation, this process is done once per coded video sequence.

After generation of each pattern, the start point and end point of a line splitting the PU is slightly shifted to avoid the case e.g., when start point and end point is at the same location. The new points are mainly to make it more robust and accurate for *D*[*i*] calculation.

Table 4: Values of *V* and *H* for angular Intra prediction modes

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Intra Mode Index | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| *H* | 32 | 32 | 32 | 32 | 32 | 32 | 32 | 32 | 32 | 32 | 32 | 32 | 32 | 32 | 32 | 32 | 32 |
| *V* | 32 | 26 | 21 | 17 | 13 | 9 | 5 | 2 | 0 | -2 | -5 | -9 | -13 | -17 | -21 | -26 | -32 |
| Intra Mode Index | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 |  |
| *H* | 26 | 21 | 17 | 13 | 9 | 5 | 2 | 0 | 2 | 5 | 9 | 13 | 17 | 21 | 26 | 32 |  |
| *V* | -32 | -32 | -32 | -32 | -32 | -32 | -32 | 32 | 32 | 32 | 32 | 32 | 32 | 32 | 32 | 32 |  |

## Identify a Wedgelet pattern for DMM mode 3

To apply the Wedgelet pattern list, i.e., identify a Wedgelet pattern for a depth block using DMM mode 3, the following steps are applied,

Let*TexIntraMode* be the IPM of the top-left 4×4 block in CTLB, search in the *WdgIdxLst*[*TexIntraMode*] the best Wedgelet pattern similar to the current exhaustive search in HTM.

# Experimental Results

The proposed method is tested under the common test condition [2], and the anchor is HTM3.1. The results are reported in Table 5 for 3-view test case. Since the proposed method only changes the depth map coding, only the results for “synthesized only” and “coded & synthesized” are reported.

Table 5: BD rate performance of proposed method for 3-view test case

|  |  |  |
| --- | --- | --- |
|  | synthesized only | coded & synthesized |
| Balloons | 0.0% | 0.0% |
| Kendo | 0.0% | 0.0% |
| Newspapercc | -0.3% | -0.2% |
| GhostTownFly | 0.0% | 0.0% |
| PoznanHall2 | 0.0% | 0.0% |
| PoznanStreet | 0.0% | 0.0% |
| UndoDancer | 0.1% | 0.1% |
| 1024x768 | -0.1% | 0.0% |
| 1920x1088 | 0.0% | 0.0% |
| average | -0.01% | 0.00% |

As it is reported in Table 5, there is almost no change on the overall BD rates compared with HTM3.1 for both “synthesized only” and “coded & synthesized” measurements.

# Complexity analysis

In this section, the complexity of proposed method is compared to HTM3.1. In Table 6, the average runtime ratios of the proposed method to HTM3.1 are listed. It is reportedly shown that average 3.7% of the decoding time is saved by the proposed method, while there are almost no changes on the encoding and decoding time between proposed method and HTM3.1.

Table 6: Runtime comparisons between proposed method and HTM3.1

|  |  |  |  |
| --- | --- | --- | --- |
|  | **enc time** | **dec time** | **ren time** |
| Balloons | 104.2% | 95.6% | 97.1% |
| Kendo | 103.5% | 97.5% | 99.8% |
| Newspapercc | 102.5% | 92.6% | 99.8% |
| GhostTownFly | 98.0% | 94.8% | 98.7% |
| PoznanHall2 | 107.9% | 102.6% | 104.3% |
| PoznanStreet | 94.2% | 96.8% | 101.3% |
| UndoDancer | 100.9% | 94.4% | 96.6% |
| 1024x768 | 103.4% | 95.2% | 98.9% |
| 1920x1088 | 100.1% | 97.1% | 100.2% |
| average | 101.5% | 96.3% | 99.6% |

Table 7: Number of decoder searches, HTM vs proposal

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Original** | **Proposed** | **Ratio** |
| Balloons | 359934 | 30960 | 8.8% |
| Kendo | 310637 | 25357 | 8.6% |
| Newspapercc | 1143154 | 105508 | 9.9% |
| GhostTownFly | 1464827 | 179428 | 12.5% |
| PoznanHall2 | 199554 | 21972 | 11.5% |
| PoznanStreet | 1144044 | 104830 | 9.7% |
| UndoDancer | 962885 | 110615 | 11.7% |
| 1024x768 | 604575 | 53942 | 9.1% |
| 1920x1088 | 942827 | 104211 | 11.4% |
| average | 797862 | 82667 | 10.4% |

In Table 7, the numbers of decoder searches of the current HTM and the proposed method are given for each sequence and on average, only 10% of the patterns are searched with the proposed method.

# Conclusion

The proposed solution for DMM mode 3 reduces the number of Wedgelet patterns to be searched by 10.4%, with almost no coding loss. Experimental results reportedly show that the proposal can reach roughly average 3.7% decoding time saving.

# Reference

1. “Test Model under Consideration for HEVC based 3D video coding v3.0,” ISO/IEC JTC1/SC29/WG11 MPEG2011/N12744, Geneva, Switzerland, April 2012.
2. “Common test conditions for 3DV experimentation,” ISO/IEC JTC1/SC29/WG11 MPEG2011/N12745, Geneva, Switzerland, April 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).**