|  |  |
| --- | --- |
| **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-A0032 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **3D-CE6.h related results on depth intra prediction** | | |
| *Status:* | Input document | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Byung Tae Oh,  Jaejoon Lee,  Du-Sik Park, | Email: | [byung.oh@samsung.com](mailto:byung.oh@samsung.com)  jaejoon1.lee@samsung.com  dusikpark@samsung.com |
| *Source:* | Samsung Electronics Co., Ltd | | |

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

# Abstract

Depth intra prediction is proposed to improve the depth coding gains. Coding experiments for CE6.h as defined in N12746 using a version of HTM-v3.1 including proposed scheme have been conducted and are reported in this document. The results reports 0.4% BD-rate gain is achieved. Additionally an alternative configuration is also presented, but results for this configuration have not been cross evaluated.

# Proposed algorithm description

The depth modelling mode (DMM) which is integrated in HTM-v3.1 targets more accurate intra prediction with additional information for depth map. It consists of four modes, in which wedgelet and contourlet are applied. In wedgelet, a straight line as an additional information segments each block, while contourlet uses corresponding texture information for block partition. They can achieve more accurate prediction, but usually fails to exactly detect the block boundary or edge information. Instead, the proposed plane-segmentation based intra prediction (PSIP) scheme exactly detects the block edge, and spends more bits for side information to further reduce the prediction residual.

The main idea of the proposed scheme is segmentation of each block into *k* regions, and applying different prediction schemes for each *k* (=2) plane. For example, a 4×4 depth map block with edge is shown in Fig. 1. Here, **R** = {*R*1 , ..., *R*9} represents the available neighborhood pixels for prediction, and **C** = {*C*1, …, *C*16} represents the current target pixels to be encoded. Then, two representative values, say, **P** = {P1, P*2*} (P1≤ P2) are decided from neighborhood pixels **R** by simple k-means like method. Then, target pixels are predicted by its closest representative value out of **P**, i.e.,

 (2)

where  is predicted pixel value.

Equation (2) produces the N×N binary data, called prediction map, containing which pixel is mapped to which representative value as in Fig. 2, and it is important to find out the efficient lossless coding method for the binary prediction map. In the proposed scheme, a pattern code for each 1×N row is considered, where the size of pattern code table will be 2N. In order to reduce the size of pattern code table, a code including more than two bit-transition is eliminated as shown in Fig. 3. Likewise, column-wise pattern code can be also considered with signaling 1-bit flag.

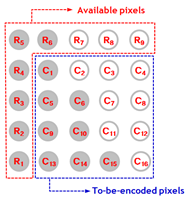


Fig. 1. Example of *k*-region segmentation (*k*=2)

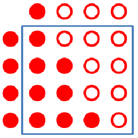


Fig. 2. Example of prediction map for Fig. 1.

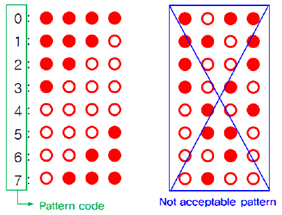


Fig.3. Example of pattern table for 4×4 block.

Finally, code difference is encoded instead of code itself to reduce the redundancy again. The difference will be computed as

 (3)

where MOD indicates modulo operator, and *si* represents *i*th row pattern code. *s*0 means a pattern code for neighbourhood pixels. For example of a prediction map in Fig. 4, the to-be-encoded pattern code {*s*1, *s*2, *s*3, *s*4} = {1, 3, 4, 4}, and *s*0=0 by pattern table in Table 1. Instead, its difference, i.e., {Δ*s*1, Δ*s*2, Δ*s*3, Δ*s*4 } = {+1, +2, +1, 0} is stored.

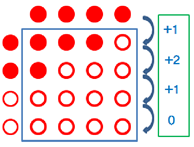


Fig. 4. Example of prediction maps.

# Coding experiments

In order to maximize the PSIP scheme, the proposed PSIP mode is used with DMM mode 1, since other DMM mode (2-4) can be covered by DMM mode 1 and the proposed PSIP mode. It is implemented in HTM-v3.1, and the comparison to anchor is given in Table 1.

Table 1: coding results for PSIP with DMM mode 1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | synthesized only | coded & synthesized | enc time | dec time |
| Balloons | -0.2% | -0.2% | 100.3% | 100.8% |
| Kendo | -0.2% | -0.1% | 100.0% | 100.4% |
| Newspapercc | -0.8% | -0.6% | 100.1% | 100.4% |
| GhostTownFly | -0.2% | -0.1% | 100.3% | 97.8% |
| PoznanHall2 | -0.5% | -0.4% | 101.0% | 98.9% |
| PoznanStreet | -0.2% | -0.1% | 100.7% | 98.3% |
| UndoDancer | -0.8% | -0.5% | 100.4% | 100.5% |
| **average** | **-0.4%** | **-0.3%** | **100.4%** | **99.6%** |

# Alternative configuration (not cross checked)

In alternative test, the PSIP mode is used with DMM mode 1, 3 and 4.

Table 2: coding results for PSIP with DMM mode 1, 3, 4

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | synthesized only | coded & synthesized | enc time | dec time |
| Balloons | -0.2% | -0.2% | 101.8% | 100.0% |
| Kendo | -0.1% | -0.1% | 101.0% | 102.4% |
| Newspapercc | -0.9% | -0.6% | 99.9% | 108.5% |
| GhostTownFly | -0.2% | -0.2% | 101.0% | 103.4% |
| PoznanHall2 | -0.2% | -0.1% | 101.4% | 101.5% |
| PoznanStreet | -0.1% | -0.1% | 101.2% | 101.8% |
| UndoDancer | -0.9% | -0.7% | 101.6% | 102.7% |
| **average** | **-0.4%** | **-0.2%** | **101.1%** | **102.9%** |

# Conclusion

Depth intra prediction has been suggested. Coding results for two different tests with the proposed scheme have been presented. As results, it reportedly achieves 0.4% BD-gain for synthesized view only, and 0.3% BD-gain for coded & synthesized views.

# Patent right declaration(s)

**Samsung Electronics Co., Ltd. may have IPR 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).**