|  |  |
| --- | --- |
| **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-B0051 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **3D-CE5.h related: Advanced residual prediction for multiview coding** | | |
| *Status:* | Input Document | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Li Zhang Ying Chen Xiang Li Marta Karczewicz  5775 Morehouse Drive San Diego, CA 92121 USA | Tel: Email: | 1-858-651-6660 [lizhang@qti.qualcomm.com](mailto:lizhang@qti.qualcomm.com)  1-858-845-6589 [cheny@qti.qualcomm.com](mailto:cheny@qti.qualcomm.com) |
| *Source:* | Qualcomm Incorporated | | |

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

# Abstract

Inter-view residual prediction is enabled in the current HTM design to code the residue of dependent texture views more efficiently. In this proposal, an advanced residual prediction (ARP) is proposed to further improve the coding efficiency of inter-view residual prediction. In ARP, to ensure high correlation between residues of two views, motion of the current block of picture in current view is applied to the corresponding block in a reference view picture to generate residual in the base view to be used for inter-view residual prediction. Moreover, an adaptive weighting factor is applied to the residue signal so that the prediction error is further reduced. Compared to the current HTM design, the proposed method achieves compression efficiency gain of 0.9% for coded views in terms of BD rate.

1. **Introduction**

In the current design of 3D-HEVC, inter-view residual prediction is enabled in a way that the corresponding block in the base view is firstly located based on the disparity vector and then the residue of the corresponding block is used as a predictor of the residue of current coding unit (CU). The residue of the corresponding block is added to the motion compensated signal and only the resulting difference signal is transform coded and transmitted. In other words, the residue generated from motion compensation is further predicted by the residue from the base view, before it is transformed, quantized and written in the bitstream.

Although the current inter-view residual prediction further improves the overall quality of prediction, it has two drawbacks. First, it uses the reference view residues to directly predict those in current view regardless whether the residual block is obtained with the same motion as that in the current view. Clearly, when the two blocks are with different motion, the correlation between them is relatively low so that the prediction performance will be sub-optimal. Second, reconstructed residues are used in prediction, which introduces undesired quantization error in prediction and further degrades the performance.

1. **Proposed Method**

To further improve the coding efficiency, it is proposed that motion of current view is applied to the corresponding block in a reference view to have a better prediction performance. Moreover, additional weighting factor is introduced to further compensate quality difference between different views.

Figure 1 illustrates the prediction structure of ARP in multiview video coding. As shown in Figure 1, Dc represents the current block in the current view (view 1), Bc, and Dr denote the representation of Dc in the reference view (view 0) at time Tj and Dc’s temporal prediction from the same view (view 1) at time Ti. VD denotes the motion from Dc to Dr. Since Dc and Bc are actually projections of the same object in two different views, these two blocks should share the same motion information. Therefore, Bc’s temporal prediction Br in view 0 at time Ti can be located from Bc by applying the motion information of VD. The residual of Bc with motion information of VD is then multiplied by a weighting factor and used as the residual predictor for current residual.

## 1

Figure 1. Prediction structure of the proposed ARP in 3D-HEVC.



Figure 2. Relationship among current block, reference block and motion compensated block

## Decoding process of ARP

Main procedures of the proposed ARP at the decoder side can be described as follows:

1. Obtain a disparity vector as specified in the current 3D-HEVC, pointing to a target reference view. Then, in the picture of the reference view within the same access unit, the corresponding block is located by the disparity vector.
2. Re-use the motion information of the current block to derive the motion information for the reference block. Apply motion compensation for the corresponding block based the same motion vector of current block and derived reference picture in the reference view for the reference block, to derive a residue block. The relationship among current block, corresponding block and motion compensated block is shown in Figure 2. The reference picture in the reference view (V0) which has the same POC (Picture Order Count) value as the reference picture of current view (Vm) is selected as the reference picture of the corresponding block.
3. Apply the weighting factor to the residue block to get a weighted residue block and add the values of the weighted residue block to the predicted samples.

## Weighting factor

Three weighting factors are used in ARP, i.e., 0, 0.5 and 1. The one leading to minimal rate-distortion cost for the current CU is selected as the final weighting factor and the corresponding weighting index (0, 1 and 2 which correspond to weighting factor 0, 1, and 0.5, respectively) is transmitted in the bitstream at the CU level. All PU predictions in one CU share the same weighting factor. When the weighting factor is equal to 0, ARP is not used for the current CU.

1. **Compression Performance**

This section provides simulation results of the proposed ARP in comparison with the 3D-HTM anchor. The platform 3D-HTM 4.0 [1] is utilized and the proposed method is implemented on it. All the simulation tests are performed based on the common test conditions [2]. Since disparity vector will be utilized in the proposed ARP, different methods for generating the disparity vector have been tested for ARP.

Table 1 provides the results based on the disparity vector derivation method used in common test condition [3]. As shown in Table 1, the overall average bitrate reduction is 0.7%, 0.5%, 0.5% for decoded texture views, synthesized views, coded and synthesized views, respectively.

As the advanced residual prediction may be sensitive to the correctness of disparity vectors, more efficient disparity vector derivation method named NBDV as proposed in [4] is further utilized to evaluate the performance of ARP and the coding gain compared to anchor is tabulated in Table 2.

As shown in Table 2, when more efficient disparity vector derivation is used, additional coding gain could be observed, i.e., 0.9%, 0.7%, 0.7% for decoded texture views, synthesized views, coded and synthesized views, respectively. Note the coding performance difference between NBDV and anchor (0.07%) for the texture is less than the gap between 0.7% and 0.9%, which is more precisely 0.18%.

Table 1: Coding gain of ARP with [3] with respect to anchor for 3-view case

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Video 1 | Video 2 | Video only | Synthesized only | Coded & synthesized | Enc time | Dec time | Ren time |
| Balloons | -2.5% | -2.6% | -0.9% | -0.7% | -0.7% | 121.7% | 111.0% | 108.4% |
| Kendo | -3.3% | -3.4% | -1.2% | -0.9% | -0.9% | 117.4% | 113.9% | 107.1% |
| Newspapercc | -1.8% | -1.6% | -0.6% | -0.3% | -0.4% | 123.7% | 101.3% | 113.9% |
| GhostTownFly | -1.8% | -2.2% | -0.5% | -0.3% | -0.3% | 125.2% | 108.4% | 105.8% |
| PoznanHall2 | -1.0% | -0.4% | -0.3% | -0.4% | -0.4% | 127.7% | 109.3% | 109.3% |
| PoznanStreet | -2.8% | -2.9% | -0.8% | -0.7% | -0.7% | 130.1% | 108.0% | 109.1% |
| UndoDancer | -2.3% | -2.6% | -0.6% | -0.4% | -0.4% | 128.7% | 112.4% | 107.3% |
| 1024x768 | -2.5% | -2.5% | -0.9% | -0.7% | -0.7% | 120.9% | 108.6% | 109.8% |
| 1920x1088 | -2.0% | -2.0% | -0.5% | -0.4% | -0.5% | 127.9% | 109.5% | 107.8% |
| **average** | **-2.2%** | **-2.2%** | **-0.7%** | **-0.5%** | **-0.5%** | **124.9%** | **109.1%** | **108.7%** |

Table 2: Coding gain of ARP with NBDV with respect to anchor for 3-view case

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Video 1 | Video 2 | Video only | Synthesized only | Coded & synthesized | Enc time | Dec time | Ren time |
| Balloons | -3.2% | -3.0% | -1.1% | -0.8% | -0.9% | 121.3% | 110.2% | 105.8% |
| Kendo | -3.7% | -4.3% | -1.5% | -1.1% | -1.1% | 116.8% | 110.9% | 100.3% |
| Newspapercc | -2.7% | -1.9% | -0.8% | -0.5% | -0.6% | 123.4% | 97.9% | 103.6% |
| GhostTownFly | -1.6% | -1.9% | -0.4% | -0.3% | -0.3% | 117.6% | 103.7% | 98.3% |
| PoznanHall2 | -1.3% | -1.2% | -0.5% | -0.6% | -0.6% | 121.0% | 108.2% | 101.9% |
| PoznanStreet | -3.5% | -4.0% | -1.1% | -0.9% | -1.0% | 126.1% | 107.0% | 104.9% |
| UndoDancer | -2.6% | -2.6% | -0.7% | -0.4% | -0.5% | 119.6% | 106.5% | 97.9% |
| 1024x768 | -3.2% | -3.1% | -1.1% | -0.8% | -0.9% | 120.4% | 106.2% | 103.2% |
| 1920x1088 | -2.3% | -2.4% | -0.7% | -0.6% | -0.6% | 121.0% | 106.3% | 100.7% |
| **average** | **-2.7%** | **-2.7%** | **-0.9%** | **-0.7%** | **-0.7%** | **120.8%** | **106.3%** | **101.7%** |

1. **References**
2. 3D-HTM version 4.0: <https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware/tags/HTM-4.0/>.
3. D. Rusanovskyy, K. Müller, A. Vetro, "Common Test Conditions of 3DV Core Experiments," ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JCT3V-A1100, 1st Meeting: Stockholm, SE, 16–20 July 2012.
4. J. Sung, M. Koo, S. Yea, "3D-CE5.h: Simplification of disparity vector derivation for HEVC-based 3D video coding," ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JCT3V-A0126, 1st Meeting: Stockholm, SE, 16–20 July 2012.
5. L. Zhang, Y. Chen, M. Karczewicz, "CE5.h: Disparity vector generation results," ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JCT3V-A0097, 1st Meeting: Stockholm, SE, 16–20 July 2012.
6. **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).**