|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  14th Meeting: Vienna, AT, 25 July – 2 Aug. 2013 | Document: JCTVC-N0187 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Performance investigation on using H-ILP picture for reference index based SHVC** | | |
| *Status:* | Input Document | | |
| *Purpose:* | Information | | |
| *Author(s) or Contact(s):* | Xiaoyu Xiu, Yan Ye, Yuwen He, Yong He  9710 Scranton R-D, #250 San Diego, CA 92121 USA | Tel: Email: | 1-858-210-4830 [Xiaoyu.Xiu@InterDigital.com](mailto:Xiaoyu.Xiu@InterDigital.com)  [Yan.Ye@InterDigital.com](mailto:Yan.Ye@InterDigital.com)  [Yuwen.He@InterDigital.com](mailto:Yuwen.He@InterDigital.com)  [Yong.He@InterDigital.com](mailto:Yong.He@InterDigital.com) |
| *Source:* | InterDigital Communications, Inc. | | |

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

# Abstract

The reference index based solution of current SHVC Test Mode (SHM-2.0) is built upon a multi-loop decoding scheme, where reconstructed base layer (BL) picture (after up-sampling if necessary) is used for the inter-layer prediction (ILP) of enhancement layer (EL) coding. One hybrid inter-layer prediction (H-ILP) picture, which is generated using the motion information and the residue information of BL picture and the texture of EL temporal reference picture, is proposed to replace the conventional ILP picture for the single-loop decoding of SNR scalability of SHVC in [1]. This contribution investigates the performance of H-ILP picture based multi-loop decoding by using both ILP picture and H-ILP picture as inter-layer reference pictures for EL prediction. As a part of the investigation, bilinear filter is applied to up-sample the residue of BL picture for the generation of H-ILP picture in the spatial scalability cases. Our investigation shows that H-ILP picture can further improve the performance of SHVC reference index based framework by providing 2.1%, 2.4% and 1.6% BD-rate savings on average for RA, LDB and LDP respectively.

# Introduction

In [1], a SNR single-loop decoding scheme is proposed by replacing normal ILP picture with H-ILP picture. H-ILP picture is generated by using BL motion information to perform motion compensation (MC) on EL temporal reference pictures. The experimental results in [1] show that H-ILP is comparable to normal ILP picture with regards to the efficiency of EL inter-layer prediction, while significantly reducing the decoding complexity and memory access. The main objective of this contribution is to investigate the performance of using H-ILP in the multi-loop decoding based SHVC. More specifically, H-ILP picture is incorporated into the current SHVC reference index based framework by using both ILP picture and H-ILP picture as the inter-layer reference pictures for EL prediction. In addition, bilinear filter is applied to up-sample the residue of BL picture for the generation of H-ILP picture in the spatial scalability cases.

# Technical description

## H-ILP picture generation

For each block of H-ILP picture, its sample values are generated by adding the residue of the BL collocated block (up-sampled if BL resolution is different from EL resolution) to the motion compensated component which is generated using the motion information of the collocated BL block and the texture information of EL temporal reference pictures. Figure 1 illustrates the H-ILP picture generation process where the collocated BL block is uni-predicted.



Figure 1. Example of generating H-ILP picture

Let denote the block located at in the H-ILP picture, and denote the block at the corresponding collocated position in the BL picture. denotes the residue signal of . is obtained by adding the up-sampled to the motion compensated prediction component which is generated using the motion vector (MV) of (scaled if the resolutions of the BL picture and the EL picture are different) and the EL temporal reference pictures. More specifically, when is uni-predicted, is predicted using the equation (1):

) (1)

where is up-sampling operation. is the EL temporal reference picture indicated by which is the scaled MV of . When is bi-predicted, is predicted according to the equation (2):

(2)

where and are the EL temporal reference picture in reference picture lists L0 and L1, as indicated by ( and ( which are the scaled MVs of in L0 and L1 respectively.

To obtain the residue signal of H-ILP block in Figure 1, bilinear filter is applied to calculate the residue value at fractional BL pixel positions.

## EL reference picture list placement

shows the reference picture list construction process with the consideration of H-ILP picture. More specifically, for EL B-slice, ILP picture and H-ILP picture are placed at the end of the reference picture lists L0 and L1 respectively. For EL P-slice, H-ILP picture is placed after all EL temporal reference pictures and ILP picture.



Figure 2. Illustration of EL reference picture list placement

In addition, for the BL pictures that are intra-coded, H-ILP will not be available given that there is no BL motion information. In this case, only ILP picture is used for the EL inter-layer prediction.

# Experimental results

The H-ILP picture based multi-loop scheme as discussed in Section 2 is implemented based on the current SHM-2.0 [2] and tested under the common test conditions specified in [3]. The results of reference index based SHM-2.0 are used as anchor for BD-rate calculation. For the results presented in this section, uncompressed BL motion information is used to generate H-ILP pictures. Both encoding time and decoding time are measured from a heterogeneous cluster system.

Table 1. Average BD-rate results compared to SHM-2.0 reference index based anchor

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **RA HEVC 2x** | | | **RA HEVC 1.5x** | | | **RA HEVC SNR** | | |
|  | Y | U | V | Y | U | V | Y | U | V |
| Class A | -2.1% | -4.8% | -4.8% |  |  |  | -1.9% | -5.5% | -6.0% |
| Class B | -1.8% | -4.1% | -4.7% | -2.2% | -5.2% | -5.8% | -2.0% | -5.2% | -6.3% |
| **Overall (Test vs Ref)** | -1.9% | -4.3% | -4.7% | -2.2% | -5.2% | -5.8% | -2.0% | -5.2% | -6.2% |
| **Overall (Test vs single layer)** | 16.9% | 27.5% | 25.7% | 13.6% | 22.3% | 21.7% | 12.1% | 25.0% | 25.8% |
| **Overall (Ref vs single layer)** | 19.2% | 33.3% | 32.0% | 16.2% | 28.8% | 29.1% | 14.4% | 32.1% | 34.1% |
| **EL only (Test vs Ref)** | -3.4% | -5.8% | -6.2% | -5.7% | -8.7% | -9.4% | -3.5% | -7.0% | -8.0% |
| Enc Time[%] | 114.3% | | | 119.2% | | | 122.6% | | |
| Dec Time[%] | 103.4% | | | 121.1% | | | 125.8% | | |
| Enc Mem[%] | #DIV/0! | | | #DIV/0! | | | #DIV/0! | | |
| BL Match | Matched | | | Matched | | | Matched | | |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **LD-B HEVC 2x** | | | **LD-B HEVC 1.5x** | | | **LD-B HEVC SNR** | | |
|  | Y | U | V | Y | U | V | Y | U | V |
| Class A | -2.3% | -4.8% | -5.0% |  |  |  | -2.2% | -5.1% | -5.5% |
| Class B | -2.0% | -4.2% | -4.6% | -2.8% | -5.6% | -6.0% | -2.2% | -5.1% | -6.0% |
| **Overall (Test vs Ref)** | -2.1% | -4.4% | -4.7% | -2.8% | -5.6% | -6.0% | -2.2% | -5.1% | -5.8% |
| **Overall (Test vs single layer)** | 25.8% | 32.8% | 33.1% | 21.3% | 25.6% | 27.7% | 21.5% | 27.8% | 31.2% |
| **Overall (Ref vs single layer)** | 28.5% | 39.0% | 39.7% | 24.8% | 33.0% | 35.9% | 24.3% | 34.7% | 39.5% |
| **EL only (Test vs Ref)** | -3.6% | -5.9% | -6.1% | -6.4% | -9.1% | -9.6% | -3.7% | -6.7% | -7.5% |
| Enc Time[%] | 107.2% | | | 108.3% | | | 106.4% | | |
| Dec Time[%] | 98.1% | | | 108.8% | | | 103.9% | | |
| Enc Mem[%] | #DIV/0! | | | #DIV/0! | | | #DIV/0! | | |
| BL Match | Matched | | | Matched | | | Matched | | |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **LD-P HEVC 2x** | | | **LD-P HEVC 1.5x** | | | **LD-P HEVC SNR** | | |
|  | Y | U | V | Y | U | V | Y | U | V |
| Class A | -1.5% | -3.3% | -3.2% |  |  |  | -1.7% | -4.1% | -4.2% |
| Class B | -1.4% | -2.6% | -3.1% | -1.8% | -3.5% | -3.8% | -1.4% | -3.0% | -3.5% |
| **Overall (Test vs Ref)** | -1.4% | -2.8% | -3.1% | -1.8% | -3.5% | -3.8% | -1.5% | -3.3% | -3.7% |
| **Overall (Test vs single layer)** | 24.8% | 34.1% | 34.8% | 20.6% | 28.2% | 30.5% | 21.5% | 30.1% | 34.2% |
| **Overall (Ref vs single layer)** | 26.6% | 38.0% | 39.1% | 22.8% | 32.8% | 35.6% | 23.4% | 34.6% | 39.4% |
| **EL only (Test vs Ref)** | -2.5% | -3.8% | -4.2% | -4.3% | -6.0% | -6.3% | -2.6% | -4.5% | -4.9% |
| Enc Time[%] | 106.7% | | | 107.9% | | | 109.2% | | |
| Dec Time[%] | 94.0% | | | 100.8% | | | 104.9% | | |
| Enc Mem[%] | #DIV/0! | | | #DIV/0! | | | #DIV/0! | | |
| BL Match | Matched | | | Matched | | | Matched | | |

As shown in Table 1, compared to SHM-2.0 anchor, the H-ILP based scheme can reportedly provide average {Y, U, V} BD-rate savings of {2.1%, 4.9%, 5.0%}, {2.4%, 5.0%, 5.5%} and {1.6%, 3.2%, 3.5%} for RA, LDB and LDP configurations, respectively, The complete BD-rate results can be found in the spreadsheet in the package.

# Conclusion

This contribution investigates the performance of using H-ILP for multi-loop decoding which uses both ILP picture and H-ILP picture as inter-layer reference pictures of EL prediction. The investigation results show that H-ILP picture can provide additional BD-rate savings of 2.1%, 2.4% and 1.6% BD-rate savings on average for RA, LDB and LDP respectively, compared to the SHM-2.0 anchor.

# References

1. X. Xiu, Y. Ye, Y. W. He and Y. He, “Single-loop decoding of SNR scalability for reference index based SHVC”, Vienna, Austria, July, 2013.
2. J. Chen, J. Boyce, Y. Ye and M. M. Hannuksela, “SHVC Working Draft 2”, JCTVC document JCTVC-M1008, Incheon, Korea, April, 2013.
3. X. Li, J. Boyce, P. Onno and Y. Ye, “Common Test Conditions and Software Reference Configurations for the Scalable Test Model”, JCTVC document JCTVC-M1009, Incheon, Korea, April, 2013.