|  |  |
| --- | --- |
| **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-N0186 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Single-loop decoding based SNR scalability for SHVC** | | |
| *Status:* | Input Document | | |
| *Purpose:* | Proposal | | |
| *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

This contribution describes a single-loop decoding scheme for SNR scalability of reference index based SHVC. In the current reference index based SHVC, the reconstructed base layer (BL) picture (after up-sampling if necessary) is used for the inter-layer prediction (ILP) of enhancement layer (EL) coding. This implies that multiple motion compensation (MC) operations have to be performed for all dependent layers. In this contribution, a single-loop decoding scheme is achieved by introducing an alternative ILP picture by applying BL motion information on EL temporal reference pictures. In addition, BL residue is added to the alternative ILP picture for quality improvement. As both BL motion information and BL residue could be obtained without the full reconstruction of BL picture, the single-loop decoding requirement is fulfilled by replacing the conventional ILP picture with the alternative ILP picture for EL prediction. Compared to the existing single-loop decoding approaches studied in AHG16, the proposed scheme does not require any low-level changes to single-layer HEVC.

Experimental results shows that compared to SHM 2.0 reference index based multi-loop decoding scheme, the proposed sing-loop decoding scheme can reportedly reduce the average memory access by 34% for block-based implementation and 32% for picture-based implementation, with average BD-rate increase of 1.3% and 1.5% for RA and LDB configurations respectively. In addition, the average decoding time is reduced by 25%. If the combined prediction between the alternative ILP picture and EL temporal reference picture is disallowed for EL prediction, the average decoding time reduction is increased to 28% and the worst case of memory access is the same as single-layer HEVC with BD-rate increase of 2.2% and 3.5%.

# Introduction

The current SHVC solution is built upon the reference index based multi-loop decoding framework where ILP picture generated from BL picture is used as one additional reference picture along with the temporal reference pictures for EL picture coding [1]. This approach implies that the BL picture has to be fully reconstructed in order to decode the EL picture. This requires an SHVC decoder to perform the MC operation for each coding layer, which results in a high complexity on EL decoding. Though the complexity increase of the current multi-loop decoding scheme could be acceptable for the spatial scalability cases, it may bring a more significant decoding complexity especially the memory access bandwidth increase for SNR scalability given the same resolution of BL pictures and EL pictures. Therefore, single-loop decoding scheme is desirable for the SNR scalability of the current SHVC design.

# Proposed method

In this contribution, a SNR single-loop decoding scheme is proposed for the reference index based framework of SHVC. The desired single-loop advantage of the proposed scheme is achieved by replacing the conventional ILP picture generated from reconstructed BL picture by an alternative ILP picture. The alternative ILP picture is generated by performing MC operation with BL motion information and EL temporal reference pictures. Then, the BL residue signal is added to the motion compensated EL signal to further improve the quality of the alternative ILP picture. Given that the alternative ILP picture is generated using the information from both BL and EL, we thus call this new type of ILP pictures as hybrid inter-layer prediction (H-ILP) picture in the rest of the contribution.

## H-ILP picture generation

For each block of the H-ILP picture, its sample values are generated by adding the residue of the BL collocated block 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. The framework of H-ILP picture generation process is shown in Figure 1 where the collocated BL block is uni-predicted.



Figure 1. Example of generating H-ILP picture

Let and denote the blocks located at in the H-ILP picture and the BL picture respectively, and denote the residue signal of . is obtained by adding to the motion compensated prediction component which is generated using the motion information of and the EL temporal reference pictures. More specifically, when is uni-predicted, is predicted using the equation (1):

) (1)

where is the EL temporal reference picture indicated by the motion vector 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 the forward motion vector ( and the backward motion vector ( of

On the other hand, if the collocated BL block is intra-coded and has no motion information, the corresponding block in the H-ILP picture is directly copied from the reconstructed samples of the collocated BL block. However, in this case, additional restrictions should be applied to ensure that the H-ILP block can be constructed without reconstructing any inter-coded BL blocks. First, the constrained intra prediction is enabled when encoding the BL pictures such that no inter-coded samples are employed for intra prediction of BL pictures. Second, all the BL intra-coded samples that are used to generate H-ILP picture are obtained prior to the de-blocking filter process and the SAO process of BL, which ensures that no filtering is needed across the boundaries between inter-coded BL samples and intra-coded BL samples in order to generate H-ILP picture.

## Single-loop decoding scheme based on H-ILP picture

shows the proposed single-loop scheme using H-ILP pictures for the reference index based SHVC. Compared to the current multi-loop based SHVC solution, the proposed single-loop scheme uses H-ILP picture to replace the previous ILP picture. The H-ILP generation process takes the BL motion information, the BL residue information and the EL temporal reference pictures as inputs, and outputs H-ILP picture according to process specified in Section 2.1. Because the BL motion information and the BL residue information can be directly obtained by parsing the BL bit-stream and applying inverse quantization and inverse transform to the quantized BL residual, the decoding of the EL pictures can be operated without the need to perform the MC operation and the loop filtering operations to obtain the fully reconstructed BL pictures. In addition, when compared to the current single-loop decoding approaches studied in AHG16 [2, 3], the proposed single-loop decoding scheme is naturally compatible with the reference index design of the current SHVC, as it does not require any low-level changes to the existing single-layer HEVC decoder design.



Figure 2. The proposed single-loop decoding scheme

## Constraint on EL inter-prediction

In the proposed single-loop decoding scheme, as H-ILP picture is used as one additional reference picture along with temporal reference pictures to predict EL picture, one bi-predicted EL PU could be generated by combining two prediction components, one from the H-ILP picture and the other from a temporal EL reference picture. And, the H-ILP prediction component itself may be bi-predicted and consequently require two MC operations. Therefore, at the worst case, one bi-predicted EL PU could maximally require three MC operations (two MC operations to generate the bi-predicted H-ILP prediction signal and one MC operation to combine the EL temporal prediction with the bi-predicted H-ILP prediction). This could impose an undesirable bandwidth increase on the decoding of the EL video, as one single motion compensation loop can at most support two MC operations. Therefore, in this constraint, it is disallowed to combine the signal of H-ILP picture with the signal from EL temporal reference picture when the H-ILP prediction signal itself is bi-predicted.

# Coding performance results

The proposed single-loop decoding scheme is implemented based on the current SHM2.0 and tested under the common test conditions specified in [4]. The results of reference index based SHM2.0 with constrained intra prediction activated are used as multi-loop decoding anchor for BD-rate calculation. For the performance results presented in this section, uncompressed BL motion information is used to generate H-ILP pictures.

The decoding timing results are reliable by measuring from the same PC without output of decoded YUV data, while the encoding timing results are inaccurate which are measured from a heterogeneous cluster system.

## Coding performance without EL inter-prediction constraint

The summarized BD-rate results of the proposed single-loop decoding scheme compared to SHM2.0 multi-loop anchor are shown in Table 1. As can be seen, the proposed sing-loop scheme can significantly reduce the decoding time by 25% on average by skipping the computationally complex MC operation and loop filtering to fully reconstruct the BL pictures, while only increases the luma BD-rate on average by 1.3% and 1.5% for RA and LDB configurations respectively compared to the multi-loop results. Meanwhile, the proposed single-loop scheme reportedly achieves 7.5% and 9.8% BD-rate savings for U and V components respectively. The corresponding complete BD-rate results can be found in the spreadsheet named “SHM-2.0-SLoop-vs-SHM-2.0-Refidx.xls” in the package.

Table 1. BD-rates of single-loop decoding compared with SHM2.0 multi-loop anchor

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **RA HEVC SNR** | | | **LD-B HEVC SNR** | | |
|  | Y | U | V | Y | U | V |
| Class A | 1.4% | -8.6% | -9.3% | 1.9% | -8.0% | -9.0% |
| Class B | 1.3% | -7.4% | -10.3% | 1.4% | -7.0% | -9.8% |
| **Overall (Test vs Ref)** | 1.3% | -7.7% | -10.0% | 1.5% | -7.3% | -9.6% |
| **Overall (Test vs single layer)** | 17.2% | 22.5% | 21.0% | 27.9% | 26.8% | 27.6% |
| **Overall (Ref vs single layer)** | 15.8% | 33.7% | 35.8% | 26.0% | 37.5% | 42.6% |
| **EL only (Test vs Ref)** | 2.6% | -7.4% | -9.9% | 3.1% | -6.4% | -8.9% |
| Enc Time[%] | 90.3% | | | 89.1% | | |
| Dec Time[%] | 74.0% | | | 77.2% | | |
| BL Match | Matched | | | Matched | | |

## Coding performance with EL inter-prediction constraint

Table 2 summarizes the comparison between the BD-rates of the proposed single-loop decoding scheme when the EL inter-prediction constraint as specified in Section 2.3 is enabled and that of SHM2.0 multi-loop anchor. It can be seen that the EL inter-prediction constraint could increase the average {Y, U, V} BD-rates by {0.9%, 0.7%, 0.7%} and {2%, 2.1%, 2.1%} for RA and LDB configurations respectively, while provides another 3% decoding time reduction compared to single-loop decoding results presented in Table 1. The corresponding complete BD-rate results can be found in the spreadsheet named “SHM-2.0-SLoop-PredCond-vs-SHM-2.0-Refidx.xls” in the package.

Table 2. BD-rates of single-loop decoding with EL inter-prediction constraint compared with SHM2.0 multi-loop anchor

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **RA HEVC SNR** | | | **LD-B HEVC SNR** | | |
|  | Y | U | V | Y | U | V |
| Class A | 1.8% | -8.0% | -8.8% | 3.3% | -6.4% | -7.6% |
| Class B | 2.3% | -6.5% | -9.5% | 3.6% | -4.7% | -7.4% |
| **Overall (Test vs Ref)** | 2.2% | -7.0% | -9.3% | 3.5% | -5.2% | -7.5% |
| **Overall (Test vs single layer)** | 18.2% | 23.5% | 22.0% | 30.4% | 29.6% | 30.5% |
| **Overall (Ref vs single layer)** | 15.8% | 33.7% | 35.8% | 26.0% | 37.5% | 42.6% |
| **EL only (Test vs Ref)** | 3.6% | -6.4% | -9.0% | 6.1% | -3.4% | -5.9% |
| Enc Time[%] | 95.8% | | | 94.5% | | |
| Dec Time[%] | 71.9% | | | 73.9% | | |
| BL Match | Matched | | | Matched | | |

# Complexity assessment

To better understand the complexity impact of the proposed single-loop decoding scheme, the methodology of complexity assessment from AHG17 [5] is also applied to the proposed single-loop method. The proposed single-loop decoding algorithm is integrated into the modified SHM2.0 generated by applying the complexity assessment patch file [4]. Both the worst case complexity and the average complexity of the proposed single-loop decoding scheme are studied and compared with the multi-loop SHM2.0 as well as single-layer HEVC decoding. The worst case complexity provides a theoretical measurement of the peak memory access, i.e., memory bandwidth, when decoding BL and EL pictures. The average complexity collects the actual memory access data when decoding the real BL and EL bit-streams. Combined with the decoding time as presented in Section 3, the statistics presented in this section could provide a good estimate on the complexity of the proposed single-loop decoding scheme.

## Worst case complexity

Table 3 shows the worst case of memory access of both single-loop decoding scheme and multi-loop decoding scheme compared to single-layer HEVC. The number over 100% means the tested method needs more memory bandwidth than HEVC for hardware design.

Table 3. Worst memory bandwidth comparison with single-layer HEVC

|  |  |  |
| --- | --- | --- |
| Tested method | Block-based implementation | Picture-based implementation |
| SHM2.0 multi-loop decoding | 200% | 218% |
| SHM2.0 single-loop decoding | 193% | 208% |
| SHM2.0 single-loop decoding with inter-prediction constraint | 100% | 208% |

As shown in Table 3, for picture-based implementation, both the single-loop decoding scheme and the multi-loop decoding scheme could use as many as two times the memory bandwidth of the single-layer HEVC, given that all the blocks in the inter-layer reference picture (i.e., H-ILP picture for single-loop decoding scheme and ILP picture for multi-loop decoding scheme) are always generated in picture-based implementation even if they are never referred by the EL picture. Alternatively, for block-based implementation, the blocks in the inter-layer reference picture are only generated when they are referred by the EL picture. Even that, the block-based memory access of single-loop decoding is still very high (193%), which is close to the corresponding number of multi-loop decoding scheme (200%). After analyzing the data, it is found that this peak memory access comes from EL 4x8 uni-predicted PU which makes reference to one H-ILP bi-predicted block. The problem can be solved by applying the EL inter-prediction constraint, which makes the memory bandwidth of the proposed single-loop decoding scheme equal to that of single-layer HEVC.

## Average complexity

The summarized average memory access of the proposed single-loop decoding scheme compared to that of SHM2.0 multi-loop anchor are shown in Table 4. As in [5], the actual memory access is calculated for 3 different memory architectures (Pure, DDR2 and DDR3).

As can be seen in Table 4, the single-loop decoding scheme can significantly lower the number of average memory access by 36%, 33% and 32% on average for Pure, DDR2 and DDR3 respectively, for block-based implementation. When considering picture-based implementation, the corresponding average memory access savings are 31%, 31% and 32%. Moreover, enabling EL inter-prediction constraint could bring another additional 4% average memory access saving for block-based implementation.

Table 4. Average memory access comparison between single-loop decoding and multi-loop decoding

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Tested method | Block-based implementation | | | Picture-based implementation | | |
| Pure | DDR2 | DDR3 | Pure | DDR2 | DDR3 |
| SHM2.0 single-loop decoding | 64% | 67% | 68% | 69% | 69% | 68% |
| SHM2.0 single-loop decoding with inter-prediction constraint | 60% | 63% | 63% | 69% | 69% | 68% |

The detailed memory access results can be found in the following spreadsheets attached in the package.

“JCTVC-SLOOP-Complexity-BLK.xls” contains the memory access results of block-based implementation of the proposed single-loop decoding scheme.

“JCTVC-SLOOP-Complexity-PIC.xls” contains the memory access results of picture-based implementation of the proposed single-loop decoding scheme.

“JCTVC-SLOOP-PredCond-Complexity-BLK.xls” contains the memory access results of block-based implementation of the proposed single-loop decoding scheme with EL inter-prediction constraint enabled.

“JCTVC-SLOOP-PredCond-Complexity-PIC.xls” contains the memory access results of picture-based implementation of the proposed single-loop decoding scheme with EL inter-prediction constraint enabled.

# Conclusion

This contribution describes a single-loop decoding scheme for SNR scalability of reference index based SHVC by replacing ILP picture with H-ILP picture for EL inter-layer prediction, and H-ILP picture is generated by adding BL residue to the motion compensated EL signal using BL motion information. The proposed single-loop decoding scheme is compatible with the current reference index based framework of SHVC without involving any low-level changes to single-layer HEVC. Experimental results show that compared to SHM 2.0 reference index based multi-loop scheme, the proposed sing-loop scheme can reportedly reduce the average memory access by 34% for block-based implementation and 32% for picture-based implementation, with average BD-rate increase of 1.3% and 1.5% for RA and LDB configurations. In addition, the average decoding time is reduced by 25% compared to multi-loop decoding scheme.

# Patent rights declaration(s)

**InterDigital Communications Inc. 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).**

# References

1. J. Chen, J. Boyce, Y. Ye and M. M. Hannuksela, “SHVC Working Draft 2”, JCTVC document JCTVC-M1008, Incheon, Korea, April, 2013.
2. M. Wien, J. Boyce, M. Budagavi, K. Mishra and K. Ugur,“JCT-VC AHG report: Single-Loop Scalability (AHG16)”, JCTVC document JCTVC-M0016, Incheon, Korea, April, 2013.
3. C. Feldmann, F. Jäger and M. Wien, “[AHG16] Analysis of Single-loop SNR Scalability using Binary Residual Refinement Coding”, JCTVC document JCTVC-M0176, Incheon, Korea, April, 2013.
4. 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.
5. E. Alshina and E. Francois, “BoG Report on SHVC Complexity Assessment”, JCTVC document JCTVC-M0455, Incheon, Korea, April, 2013.