|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  15th Meeting: Geneva, CH, 23 Oct. – 1 Nov. 2013 | Document: JCTVC-O0162 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **SCE2 2.2: Single-loop decoding based SNR scalability for SHVC** | | |
| *Status:* | Input Document to JCT-VC | | |
| *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 document reports the performance of the single-loop decoding scheme for SNR scalability of SHVC, which was originally proposed in JCTVC-N0186 and included in SCE2. In the current SHVC design, 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 JCTVC-N0186, a single-loop decoding scheme is proposed by introducing an alternative ILP picture by applying BL motion information on EL temporal reference pictures. In addition, BL residue could be further 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.

Experimental results shows that compared to SHM 3.0.1 anchor, the proposed sing-loop decoding scheme can reportedly reduce the average memory access by 35%, with the average BD-rate increase of 2.7% for RA/LDB. When temporal scalability is enabled, the same reduced memory access can also be achieved, with average BD-rate increase of 3.0% for RA/LDB, compared to SHM 3.0.1 temporal scalability anchor. In addition, when the compressed BL motion information is used to generate the alternative picture without using BL residue, the average savings of the memory access is increased to 44% and 43% for non-temporal scalability case and temporal scalability case, respectively.

# Introduction

In JCTVC-N0186 [1], a single-loop decoding scheme is proposed for the SNR scalability of SHVC. It generates an alternative ILP picture by applying BL motion information on EL temporal reference pictures, and replaces the conventional ILP picture with the alternative ILP picture for the EL coding. Moreover, BL residue information could be additionally applied to the alternative ILP picture to further improve the ILP efficiency of the EL. It is included in SCE2 2.2 [2] for a thorough testing and rigorous evaluation.

This document reports the experimental results of the proposed single-loop decoding technique, evaluated using the testing conditions specified by SCE2 [2], the AHG 17 complexity assessment methodology [3] and the pixel-level complexity assessment methodology [4]. In addition, the BD-rate performance and the complexity of the proposed single-loop decoding scheme are also studied, when the compressed BL motion information is applied to generate the alternative ILP picture without BL residue enhancement.

# Technical description

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. Additionally, the BL residue signal could be 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, this new type of ILP pictures is thus named as hybrid inter-layer prediction (H-ILP) picture for the description in the rest of the document.

## H-ILP picture generation

For each block of the H-ILP picture, its sample values are generated by applying motion compensation on the texture of EL temporal reference pictures based on the motion information of the collocated BL block. In addition, the residue of the BL collocated block could be added to the motion compensated component to enhance the quality of the H-ILP picture. 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

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.

It should be mentioned that the motion information used to generate H-ILP pictures can be mapped either from the original (un-compressed) motion information or from the compressed motion information of the BL. Moreover, the BL residue enhancement can also be enabled or disabled for the generation of H-ILP pictures. The corresponding BD-rate performance and complexity data, using un-compressed/compressed BL motion information and with/without BL residue, will be presented in Section 3.

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

Figure 2 shows the proposed single-loop scheme by using H-ILP pictures based on the current SHVC design. Compared to the multi-loop based SHVC solution, the proposed single-loop scheme uses the H-ILP picture, generated according to the process specified in Section 2.1, to replace the previous ILP picture. Because the BL motion information and the BL residue information can be directly obtained by parsing the BL bit-stream, decoding of the EL pictures can be operated without the need to perform the MC operation and the loop filtering operations in the BL. In addition, the proposed single-loop decoding scheme is naturally compatible with the design of the current SHVC, as it does not require any low-level changes to the existing single-layer HEVC decoding process.



Figure 2 Proposed single-loop decoding scheme

## Constraint on EL inter-prediction

In the proposed single-loop decoding scheme, as H-ILP picture is used as the inter layer 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. Since the H-ILP prediction component itself may be bi-predicted and consequently require two MC operations, in 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 more 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. Therefore, in the proposed scheme, when an H-ILP prediction signal itself is bi-predicted, we disallow it to be further used in bi-prediction with an EL temporal reference.

# Experimental results

In this section, the proposed single-loop decoding scheme is experimentally verified with different configurations as defined in Table 1, where Test 1, Test 2 and Test 3 are conducted with temporal scalability being disabled, and Test 4, Test 5 and Test 6 are conducted with temporal scalability being enabled. All the BD-rate results and the complexity results are obtained based on the testing conditions specified by SCE2 [2], the AHG 17 complexity assessment methodology [3] and the pixel-level complexity assessment methodology [4]. Only the summarized simulation results and complexity information will be presented in this section, while the complete results can be found in the attached excel files.

Table 1 Configurations of tested cases

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Case | Temporal scalability | Compressed BL motion information | BL residue | EL Bi-prediction constraint | Anchor |
| Test 1 | N | N | Y | N | SHM 3.0.1 |
| Test 2 | N | N | Y | Y | SHM 3.0.1 |
| Test 3 | N | Y | N | N | SHM 3.0.1 |
| Test 4 | Y | N | Y | N | SHM3.0.1 temporal scalability |
| Test 5 | Y | N | Y | Y | SHM3.0.1 temporal scalability |
| Test 6 | Y | Y | N | N | SHM 3.0.1 temporal scalability |

## BD-rate performance

Table 2 summarizes the average BD-rate performance of both RA and LDB configurations for the different test cases. It can be observed that compared to SHM 3.0.1 anchor the proposed single-loop decoding scheme (Test 1) only increases the luma BD-rate on average by 2.6% and 2.9% for RA and LDB respectively, while achieving average -6.0% and -8.3% BD-rate (that is BD-rate reductions) for U and V components. If the EL inter-layer prediction constraint (Test 2) is enabled, it could increase the {Y, U, V} BD-rate by {1.4%, 1.5%, 1.4%} on average compared to Test 1. In addition, if BL compressed motion information is used and BL residue information is disabled for the generation of H-ILP pictures (Test 3), the corresponding BD-rate increases compared to SHM 3.0.1 anchor are 11.6% and 12.7% for RA and LDB respectively.

When it comes to the temporal scalability case, the proposed single-loop decoding scheme (Test 4) increases the luma BD-rate by 2.5% and 3.6% for RA and LDB. In addition, EL inter-layer prediction constraint (Test 5) gives another average {Y, U, V} BD-rate increase by {1.0%, 1.0%, 1.0%}. Finally, when considering the case of using BL compressed motion information and no BL residue information, the corresponding BD-rate increases are 11.7% and 15.8% for RA and LDB respectively.

Table 2 Summarized BD-rate performance of tested cases (compared to SHM3.0.1 anchor/SHM3.0.1 temporal scalability anchor)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Case | Configuration | Y | U | V |
| Test 1 | RA | 2.6% | -6.6% | -8.9% |
| LDB | 2.9% | -5.3% | -7.6% |
| Average | 2.7% | -6.0% | -8.3% |
| Test 2 | RA | 3.4% | -5.9% | -8.2% |
| LDB | 4.9% | -3.1% | -5.5% |
| Average | 4.1% | -4.5% | -6.9% |
| Test 3 | RA | 11.6% | 1.8% | -0.7% |
| LDB | 12.7% | 4.5% | 1.8% |
| Average | 12.1% | 3.1% | 0.5% |
| Test 4 | RA | 2.5% | -6.9% | -9.2% |
| LDB | 3.6% | -6.1% | -9.1% |
| Average | 3.0% | -6.5% | -9.2% |
| Test 5 | RA | 3.4% | -6.1% | -8.5% |
| LDB | 4.6% | -4.8% | -7.9% |
| Average | 4.0% | -5.5% | -8.2% |
| Test 6 | RA | 11.7% | 1.6% | -0.9% |
| LDB | 15.8% | 5.3% | 1.9% |
| Average | 13.7% | 3.4% | 0.5% |

## AHG 17 complexity assessment

Table 3 and Table 4 summarizes the average complexity data and the worst case complexity data of different tested cases respectively, based on the complexity assessment methodology provided by AHG 17 [3].

As can be seen in Table 3, when considering the non-temporal scalability case, the single-loop decoding scheme (Test 1) can significantly lower the number of average memory access by 35%, 32% and 31% on average for Pure, DDR2 and DDR3 respectively, compared to the average memory access data of SHM3.0.1. In addition, enabling EL inter-prediction constraint (Test 2) could bring another additional 4% average memory access saving on top of the average memory access savings of Test 1. Finally, when H-ILP pictures are generated using BL compressed motion information and no BL residue information (Test 3), the corresponding memory access savings are increased to 44%, 41% and 41% on average for Pure, DDR2 and DDR3 respectively. Similar average memory access savings can be observed when temporal scalability is applied.

Table 3 Summarized AHG17 complexity savings of tested cases (compared to SHM3.0.1 anchor/SHM3.0.1 temporal scalability anchor)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Case | Configuration | Pure | DDR2 | DDR3 | Mults |
| Test 1 | RA | 37% | 33% | 33% | 39% |
| LDB | 33% | 30% | 30% | 32% |
| Average | 35% | 32% | 31% | 35% |
| Test 2 | RA | 40% | 37% | 36% | 43% |
| LDB | 38% | 36% | 35% | 39% |
| Average | 39% | 36% | 36% | 41% |
| Test 3 | RA | 44% | 41% | 40% | 40% |
| LDB | 44% | 42% | 41% | 37% |
| Average | 44% | 41% | 41% | 39% |
| Test 4 | RA | 37% | 37% | 33% | 39% |
| LDB | 33% | 29% | 29% | 31% |
| Average | 35% | 33% | 31% | 35% |
| Test 5 | RA | 40% | 37% | 36% | 42% |
| LDB | 37% | 34% | 33% | 36% |
| Average | 38% | 35% | 35% | 39% |
| Test 6 | RA | 44% | 41% | 40% | 40% |
| LDB | 43% | 41% | 40% | 35% |
| Average | 43% | 41% | 40% | 37% |

As shown in Table 4, without enabling the EL inter-layer prediction constraint (Test 1 and Test 4), the worst case memory access of the single-loop decoding is 193% of that of single-layer HEVC, which is very close to the worst-case number for multi-loop decoding (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 single-loop decoding scheme equal to that of single-layer HEVC.

Table 4 Summarized complexity data of tested cases in worst case (compared to single-layer HEVC anchor)

|  |  |  |
| --- | --- | --- |
| Case | Memory access | Mults |
| Test 1 | 193% | 130% |
| Test 2 | 100% | 100% |
| Test 3 | 193% | 130% |
| Test 4 | 193% | 130% |
| Test 5 | 100% | 100% |
| Test 6 | 193% | 130% |

## Pixel-level complexity assessment

Table 5 summarizes the selected average pixel-level complexity data of different tested cases, based on the pixel-level complexity assessment methodology provided by [4].

Similar to the AHG17 complexity results presented in Section 3.2, pixel-level complexity data also shows a significant complexity reduction by using the proposed single-loop decoding method, compared to the SHM3.0.1 multi-loop decoding based anchor. Specifically, for Test 1, the numbers of the motion compensated pixels, de-blocking filtered pixels and SAO processed pixels are reduced by around 47%, 47% and 15% respectively. When comparing Test 1 and Test 2, EL inter-layer prediction constraints could further reduce the number of motion compensated pixels by another 7% but have smaller impact on the number of pixels processed by de-blocking and SAO. The complexity of Test 3 is similar to that of Test 2 in terms of the motion compensated pixels. However, Test 3 can significantly reduce the number of inverse transformed pixels by 33% compared to SHM3.0.1, because there is no need to reconstruct BL residue information to generate the H-ILP picture.

Table 5 Summarized pixel-level complexity savings of tested cases (compared to SHM3.0.1 anchor)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Case | Configuration | MV Full | MV Half | MV Quarter | Inv Trans Y | De-blocking Y | SAO Y |
| Test 1 | RA | 55.7% | 49.2% | 48.4% | 6.1% | 43.6% | 11.8% |
| LDB | 50.5% | 37.6% | 37.7% | 9.9% | 49.8% | 18.6% |
| Average | 53.1% | 43.4% | 43.1% | 8.0% | 46.7% | 15.2% |
| Test 2 | RA | 60.1% | 54.3% | 53.8% | 7.1% | 44.6% | 10.2% |
| LDB | 57.7% | 48.6% | 48.5% | 15.3% | 50.1% | 16.7% |
| Average | 58.9% | 51.4% | 51.1% | 11.2% | 47.3% | 13.4% |
| Test 3 | RA | 55.6% | 48.1% | 47.3% | 29.0% | 40.9% | 20.4% |
| LDB | 53.6% | 40.1% | 40.1% | 36.6% | 47.5% | 29.2% |
| Average | 54.6% | 44.1% | 43.7% | 32.8% | 44.2% | 24.8% |

# Conclusion

In this proposal, the performance of the single-loop decoding technique for the SNR scalability of SCE2 2.2 is investigated in different configurations. By avoiding the motion compensation and the loop-filter processing in the BL decoding process, significant complexity reduction is achieved by the proposed single-loop decoding scheme. Moreover, the coding performance of the single-loop decoding is kept comparable to that of multi-loop decoding based solution. It is suggested to adopt the proposed single-loop decoding method in the reference software for the SNR scalability.

# 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. X. Xiu, Y. Ye, Y. W. He and Y. He, “AHG16: Single-loop decoding of SNR scalability for refIdx based SHVC”, JCTVC document JCTVC-N0186, Vienna, Austria, August 2013.
2. M. Wien, K. Rapaka and X. Xiu, “HEVC Scalable Extension Core Experiment SCE2: Key pictures and single-loop decoding”, JCTVC document JCTVC-N1102, Vienna, Austria, August 2013.
3. [M. Budagavi](mailto:madhukar@ti.com), [E. Alshina](mailto:elena_a.alshina@samsung.com), [J. Dong](mailto:Jie.Dong@InterDigital.com), [E. François](mailto:edouard.francois@crf.canon.fr), [J. Kang](mailto:jungwon@etri.re.kr%20), [X. Li](mailto:lxiang@qti.qualcomm.com) and [A. Tabatabai](mailto:Ali.Tabatabai@am.sony.com), “JCT-VC AHG report: SHVC complexity assessment (AHG17)”, JCTVC document JCTVC-N0017, Vienna, Austria, August 2013.
4. [Christian Feldmann](mailto:feldmann@ient.rwth-aachen.de), [Fabian Jäger](mailto:jaeger@ient.rwth-aachen.de), [Mathias Wien](mailto:wien@ient.rwth-aachen.de), “[AHG16] Analysis of Single-Loop SNR Scalability using Binary Residual Refinement Coding”, JCTVC document JCTVC-M0176, Incheon, Korea, April 2013.
5. J. Chen, J. Boyce, Y. Ye and M. M. Hannuksela, “SHVC Working Draft 3”, JCTVC document JCTVC-N1008, Vienna, Austria, August 2013.