|  |  |
| --- | --- |
| **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-O0169 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **SCE2: Cross-verification of key picture concept and single loop decoding from Aachen** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Xiaoyu Xiu, Yan Ye  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) |
| *Source:* | InterDigital Communications, Inc. | | |

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

# Abstract

This document reports the cross-verification results of JCTVC-O0145 on key picture concept and single loop decoding. The software provided from the proponents is studied and is compliant with technical description in JCTVC-O0145. The experimental results and complexity assessment results are also verified. In the proposed method, as the EL pictures are used to predict the BL pictures, drifting errors are involved when only decoding the BL bit-stream. The corresponding BL drifting errors are also verified.

Based on the current software implementation of JCTVC-O0145, one block-level change is observed that in order to reduce the BL quality loss caused by the drifting error of BL temporal motion vector prediction, the motion information of BL reference picture is used to predict the motion information of BL non-key picture. More specially, instead of using the block of the collocated picture (which is an EL reference picture) for BL TMVP, the motion information of the BL block that corresponds to the EL collocated block is used. Given that EL pictures are not utilized to predict BL non-key pictures, this change is a block-level change to the current HEVC TMVP logic. Also, for the current complexity data using the AHG 17 methodology, the memory access of fetching BL transform coefficients has not been included.

# Introduction

In JCTVC-O0145 included in SCE2 [2], the concept of key pictures is studied for single-loop decoding and. More specifically, the BL pictures are classified to categories: key pictures and non-key pictures. The first picture in each GOP is key picture, and can only be predicted from other BL key pictures. All other BL pictures in each GOP are non-key pictures, and are predicted using the reconstructed EL pictures. Given that BL non-key pictures use reconstructed EL pictures for inter prediction, drifting error could be introduced when only decoding the BL bit-stream.

In JCTVC-O0145, in order to fulfill the single-loop decoding requirement, the following constraints need to be applied to the BL coding:

1. The de-blocking filter and SAO are enabled in BL coding. However, the un-filtered BL pixels (without de-blocking filter and SAO) are utilized for the prediction of EL pictures.
2. The constrained intra prediction is enabled in the BL coding.
3. In order to reduce the BL quality loss caused by the drifting error of BL temporal motion vector prediction, the motion information of BL reference picture is used to predict the motion information of BL non-key picture.
4. A bi-prediction constraint could be applied to disable the bi-predicted EL PUs which use bi-predicted BL pixels as reference.

# Problems

Based on the current software design of JCTVC-O0145, the following points could be related with changes to the decoder design. Especially, the second point could modify the low-level operation of the TMVP derivation process of HEVC.

1. Constraint 1 needs the decoder to cache the un-filtered BL pixels for the prediction of the EL, which needs additional memory buffer for the storage of the un-filtered pixels.
2. Constraint 3 could be regarded as block-level change to the current HEVC TMVP design. In key picture based structure in JCTVC-O0145, only EL pictures are used as references for the prediction of non-key BL pictures, which means only EL pictures are available to be used as the collocated picture of BL TMVP derivation process. Therefore, by the current TMVP logic of HEVC, there is no way for one BL non-key picture to access the motion information of any other BL reference picture for its TMVP derivation.

In addition, for the complexity assessment using the AHG 17 methodology, it is found that the memory accesses of the BL transform coefficients are not counted for the single-loop decoding. Currently, AHG 17 complexity patch only considers the memory access of motion compensation related operations. For single-loop decoding based scheme, as BL pictures are not completely reconstructed, partial BL information, such as motion vectors and transform coefficients, should be stored for decoding EL pictures. Compared to the memory access of motion vector, the memory access of transform coefficients are not negligible, which should be also counted when considering the complexity of EL decoding.

# Simulation results

## BD-rate performance

Table 1 Single-loop decoding without EL prediction constraint

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  | **BD-PSNR (dB)** | | | | | | |
| **RA HEVC SNR** | | |  | **RA HEVC SNR BL ONLY** | | |  | **RA HEVC SNR** | | |  | **RA HEVC SNR BL ONLY** | | |
| Y | U | V |  | Y | U | V |  | Y | U | V |  | Y | U | V |
| 2.1% | -6.6% | -6.9% |  | 10.0% | 3.8% | 4.0% |  | -0.0872 | 0.1749 | 0.1417 |  | -0.4142 | -0.0829 | -0.0747 |
| 1.9% | -5.3% | -7.8% |  | 8.8% | 5.2% | 5.1% |  | -0.0418 | 0.0674 | 0.1190 |  | -0.2258 | -0.0694 | -0.0938 |
| 2.0% | -5.7% | -7.6% |  | 9.1% | 4.8% | 4.8% |  | -0.0548 | 0.0981 | 0.1255 |  | -0.2796 | -0.0732 | -0.0883 |
| 16.6% | 23.7% | 23.1% |  |  |  |  |  |  |  |  |  |  |  |  |
| 14.4% | 32.1% | 34.2% |  |  |  |  |  |  |  |  |  |  |  |  |
| 2.1% | -6.2% | -8.3% |  |  |  |  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  | **BD-PSNR (dB)** | | | | | | |
| **LD-B HEVC SNR** | | |  | **LD-B HEVC SNR BL ONLY** | | |  | **LD-B HEVC SNR** | | |  | **LD-B HEVC SNR BL ONLY** | | |
| Y | U | V |  | Y | U | V |  | Y | U | V |  | Y | U | V |
| 2.7% | -4.6% | -5.4% |  | 15.3% | 5.0% | 3.5% |  | -0.1325 | 0.1117 | 0.1097 |  | -0.6290 | -0.1026 | -0.0632 |
| 2.4% | -3.2% | -5.5% |  | 15.5% | 6.5% | 5.6% |  | -0.0562 | 0.0378 | 0.0707 |  | -0.4117 | -0.0871 | -0.1174 |
| 2.5% | -3.6% | -5.5% |  | 15.5% | 6.1% | 5.0% |  | -0.0780 | 0.0589 | 0.0818 |  | -0.4737 | -0.0915 | -0.1019 |
| 27.3% | 29.2% | 31.0% |  |  |  |  |  |  |  |  |  |  |  |  |
| 24.3% | 34.7% | 39.5% |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.5% | -5.8% | -7.7% |  |  |  |  |  |  |  |  |  |  |  |  |

Table 2 Single-loop decoding with EL prediction constraint

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  | **BD-PSNR (dB)** | | | | | | |
| **RA HEVC SNR** | | |  | **RA HEVC SNR BL ONLY** | | |  | **RA HEVC SNR** | | |  | **RA HEVC SNR BL ONLY** | | |
| Y | U | V |  | Y | U | V |  | Y | U | V |  | Y | U | V |
| 2.1% | -6.6% | -6.9% |  | 10.0% | 3.8% | 4.0% |  | -0.0872 | 0.1749 | 0.1417 |  | -0.4142 | -0.0829 | -0.0747 |
| 1.9% | -5.3% | -7.8% |  | 8.8% | 5.2% | 5.1% |  | -0.0418 | 0.0674 | 0.1190 |  | -0.2258 | -0.0694 | -0.0938 |
| 2.0% | -5.7% | -7.6% |  | 9.1% | 4.8% | 4.8% |  | -0.0548 | 0.0981 | 0.1255 |  | -0.2796 | -0.0732 | -0.0883 |
| 16.6% | 23.7% | 23.1% |  |  |  |  |  |  |  |  |  |  |  |  |
| 14.4% | 32.1% | 34.2% |  |  |  |  |  |  |  |  |  |  |  |  |
| 2.1% | -6.2% | -8.3% |  |  |  |  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  | **BD-PSNR (dB)** | | | | | | |
| **LD-B HEVC SNR** | | |  | **LD-B HEVC SNR BL ONLY** | | |  | **LD-B HEVC SNR** | | |  | **LD-B HEVC SNR BL ONLY** | | |
| Y | U | V |  | Y | U | V |  | Y | U | V |  | Y | U | V |
| 2.7% | -4.6% | -5.4% |  | 15.3% | 5.0% | 3.5% |  | -0.1325 | 0.1117 | 0.1097 |  | -0.6290 | -0.1026 | -0.0632 |
| 2.4% | -3.2% | -5.5% |  | 15.5% | 6.5% | 5.6% |  | -0.0562 | 0.0378 | 0.0707 |  | -0.4117 | -0.0871 | -0.1174 |
| 2.5% | -3.6% | -5.5% |  | 15.5% | 6.1% | 5.0% |  | -0.0780 | 0.0589 | 0.0818 |  | -0.4737 | -0.0915 | -0.1019 |
| 27.3% | 29.2% | 31.0% |  |  |  |  |  |  |  |  |  |  |  |  |
| 24.3% | 34.7% | 39.5% |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.5% | -5.8% | -7.7% |  |  |  |  |  |  |  |  |  |  |  |  |

## AHG 17 complexity

Table 3 AHG 17 complexity of single-loop decoding without EL prediction constraint

|  |  |  |  |
| --- | --- | --- | --- |
| **RA HEVC SNR** | | |  |
| Pure | DDR2 | DDR3 | Mults |
| 59% | 60% | 62% | 68% |
| 59% | 61% | 62% | 67% |
| **59%** | **60%** | **62%** | **68%** |

|  |  |  |  |
| --- | --- | --- | --- |
| **LD-B HEVC SNR** | | |  |
| Pure | DDR2 | DDR3 | Mults |
| 62% | 62% | 63% | 77% |
| 63% | 63% | 64% | 75% |
| **63%** | **63%** | **63%** | **76%** |

Table 4 AHG 17 complexity of single-loop decoding with EL prediction constraint

|  |  |  |  |
| --- | --- | --- | --- |
| **RA HEVC SNR** | | |  |
| Pure | DDR2 | DDR3 | Mults |
| 58% | 59% | 60% | 66% |
| 57% | 59% | 60% | 64% |
| **57%** | **59%** | **60%** | **65%** |

|  |  |  |  |
| --- | --- | --- | --- |
| **LD-B HEVC SNR** | | |  |
| Pure | DDR2 | DDR3 | Mults |
| 61% | 61% | 62% | 75% |
| 61% | 62% | 63% | 72% |
| **61%** | **62%** | **62%** | **73%** |

## Pixel-level complexity

Table 5 Pixel-level complexity of single-loop decoding without EL prediction constraint for RA configuration

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | BL Decoding | | |  | EL decoding (including required BL operations) | | |
|  | SHM vs HM | Prop vs HM | Prop vs SHM |  | SHM vs HM | Prop vs HM | Prop vs SHM |
| Pix Inter | 100,29% | 102,09% | 101,79% |  | 209,98% | 172,38% | 82,09% |
| Pix Intra | 95,89% | 70,48% | 73,50% |  | 86,83% | 73,48% | 84,62% |
| Pix MV Full Sum | 100,69% | 100,34% | 99,65% |  | 237,55% | 121,39% | 51,10% |
| Pix MV Half | 100,06% | 102,65% | 102,59% |  | 188,44% | 121,39% | 64,42% |
| Pix MV Quarter | 100,08% | 104,98% | 104,90% |  | 186,55% | 122,66% | 65,75% |
| Pix MV Full BL |  |  |  |  |  |  | 5,17% |
|  |  |  |  |  |  |  |  |
| Pix Inv Transf Y | 98,72% | 95,54% | 96,78% |  | 163,17% | 160,58% | 98,41% |
| Pix Deblocking Y | 99,34% | 95,65% | 96,28% |  | 195,09% | 128,09% | 65,66% |
| Pix SAO Y | 97,35% | 145,43% | 149,39% |  | 143,96% | 147,14% | 102,21% |

Table 6 Pixel-level complexity of single-loop decoding without EL prediction constraint for LDB configuration

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | BL Decoding | | |  | EL decoding (including required BL operations) | | |
|  | SHM vs HM | Prop vs HM | Prop vs SHM |  | SHM vs HM | Prop vs HM | Prop vs SHM |
| Pix Inter | 100,00% | 101,14% | 101,14% |  | 204,31% | 170,86% | 83,63% |
| Pix Intra | 100,00% | 61,98% | 61,98% |  | 90,58% | 74,44% | 82,18% |
| Pix MV Full Sum | 100,00% | 96,97% | 96,97% |  | 229,01% | 178,13% | 77,78% |
| Pix MV Half | 100,00% | 110,82% | 110,82% |  | 163,93% | 121,87% | 74,34% |
| Pix MV Quarter | 100,00% | 106,97% | 106,97% |  | 168,96% | 120,09% | 71,08% |
| Pix MV Full BL |  |  |  |  |  |  | 134,16% |
|  |  |  |  |  |  |  |  |
| Pix Inv Transf Y | 100,00% | 89,49% | 89,49% |  | 159,17% | 150,89% | 94,80% |
| Pix Deblocking Y | 100,00% | 92,27% | 92,27% |  | 191,84% | 127,85% | 66,64% |
| Pix SAO Y | 100,00% | 126,49% | 126,49% |  | 145,46% | 149,85% | 103,02% |

Table 7 Pixel-level complexity of single-loop decoding without EL prediction constraint for RA configuration

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | BL Decoding | | |  | EL decoding (including required BL operations) | | |
|  | SHM vs HM | Prop vs HM | Prop vs SHM |  | SHM vs HM | Prop vs HM | Prop vs SHM |
| Pix Inter | 100,29% | 102,09% | 101,79% |  | 209,98% | 165,84% | 78,98% |
| Pix Intra | 95,89% | 70,50% | 73,52% |  | 86,83% | 74,71% | 86,04% |
| Pix MV Full Sum | 100,69% | 101,04% | 100,36% |  | 237,55% | 179,94% | 75,75% |
| Pix MV Half | 100,06% | 102,80% | 102,74% |  | 188,44% | 114,58% | 60,80% |
| Pix MV Quarter | 100,08% | 104,84% | 104,76% |  | 186,55% | 114,92% | 61,60% |
| Pix MV Full BL |  |  |  |  |  |  | 217,53% |
|  |  |  |  |  |  |  |  |
| Pix Inv Transf Y | 98,72% | 95,57% | 96,82% |  | 163,17% | 161,15% | 98,76% |
| Pix Deblocking Y | 99,34% | 31,87% | 32,08% |  | 195,09% | 125,26% | 64,21% |
| Pix SAO Y | 97,35% | 62,56% | 64,26% |  | 143,96% | 149,87% | 104,11% |

Table 8 Pixel-level complexity of single-loop decoding without EL prediction constraint for LDB configuration

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | BL Decoding | | |  | EL decoding (including required BL operations) | | |
|  | SHM vs HM | Prop vs HM | Prop vs SHM |  | SHM vs HM | Prop vs HM | Prop vs SHM |
| Pix Inter | 100,00% | 101,14% | 101,14% |  | 204,31% | 170,86% | 83,63% |
| Pix Intra | 100,00% | 61,98% | 61,98% |  | 90,58% | 74,44% | 82,18% |
| Pix MV Full Sum | 100,00% | 96,97% | 96,97% |  | 229,01% | 123,37% | 53,87% |
| Pix MV Half | 100,00% | 110,82% | 110,82% |  | 163,93% | 121,87% | 74,34% |
| Pix MV Quarter | 100,00% | 106,97% | 106,97% |  | 168,96% | 120,09% | 71,08% |
| Pix MV Full BL |  |  |  |  |  |  | 2,10% |
|  |  |  |  |  |  |  |  |
| Pix Inv Transf Y | 100,00% | 89,49% | 89,49% |  | 159,17% | 150,89% | 94,80% |
| Pix Deblocking Y | 100,00% | 92,27% | 92,27% |  | 191,84% | 127,85% | 66,64% |
| Pix SAO Y | 100,00% | 126,49% | 126,49% |  | 145,46% | 149,85% | 103,02% |

# References

1. [Christian Feldmann](mailto:feldmann@ient.rwth-aachen.de), [Mathias Wien](mailto:wien@ient.rwth-aachen.de), “[SCE2] Key picture concept and single loop decoding”, JCTVC document JCTVC-O0145, Geneva, Switzerland, October 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.