|  |  |  |  |
| --- | --- | --- | --- |
| **Joint Collaborative Team on 3D Video Coding Extensions**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  8th Meeting: Valencia, ES, 29 March – 4 April 2014 | Document: JCT3V-H0064 |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **CE2-related: Improvement disparity vector on temporal ARP and chroma 4x4 off** | | |
| *Status:* | Input Document to JCT-3V | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Tomohiro Ikai  1-9-2 Nakase, Mihama-ku, Chiba-shi, Chiba 261-8520  JAPAN | Tel: Email: | +81-43-299-8526 [ikai.tomohiro@sharp.co.jp](mailto:ikai.tomohiro@sharp.co.jp) |
| *Source:* | Sharp Corporation | | |

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

# Abstract

This proposal proposes a combination of the disparity improvement part of JCT3V-G0119 and chroma 4x4 residual prediction off of JCT3V-G0033. Since the disparity improvement could compensate the loss of chroma 4x4 off, it is asserted that the combination shows better trade-off between coding efficiency and complexity. The experiment result shows that the BD-rate gain is 0.0 %, 0.0 % and 0.0 % in video, total video and synthesis respectively.

# Introduction

The proponent found that SubPU part of JCT3V-G0119 shows negligible gain under chroma 4x4 off condition (shown in Table 1 as “simple combination”) but considered that the disparity improvement part of JCT3V-G0119 remains effective.

Therefore combination of disparity improvement part of JCT3V-G0119 and chroma 4x4 residual prediction off of G0033 was tested and proposed.

Table 1: Summary of proposal

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | video 0 | video 1 | video 2 | video PSNR / video bitrate | video PSNR / total bitrate | synth PSNR / total bitrate | enc time | dec time | ren time |
| **Proposed combination** | **0.0%** | **0.1%** | **0.0%** | **-0.02%** | **-0.04%** | **0.00%** | **100.1%** | **101.2%** | **100.2%** |
| **chroma 4x4 removal**  **G0033** | **0.0%** | **0.4%** | **0.3%** | **0.08%** | **0.07%** | **0.08%** | **100.0%** | **101.1%** | **100.4%** |
| **Simple combination**  **G0033+G0121** |  |  |  |  |  |  |  |  |  |

# Proposal

The contribution consists two parts:

## Omit chroma 4x4 residual prediction

It is proposed to omit chroma residual prediction in 4x4 chroma block.

* If nPbW is greater than 8, the modified prediction samples predSamplesLXCb[ x ][ y ] with x = 0..( nPbW /2 ) − 1 and y = 0..( nPbH /2 )−1 are derived as specified in the following:
  1. predSamplesLXCb[ x ][ y ] = predSamplesLXCb[ x ][ y ] +   
      ( ( rpSamplesLXCb[ x ][ y ] − rpRefSamplesLXCb[ x ][ y ] )  >>  shiftVal ) (I‑227)
* If nPbW is greater than 8, the modified prediction samples predSamplesLXCr[ x ][ y ] with x = 0..( nPbW /2 ) − 1 and y = 0..( nPbH /2 ) − 1 are derived as specified in the following:
  1. predSamplesLXCr[ x ][ y ] = predSamplesLXCr[ x ][ y ] +   
      ( ( rpSamplesLXCr[ x ][ y ] − rpRefamplesLXCr[ x ][ y ] )  >>  shiftVal ) (I‑228)

## Use more accurate disparity vector for temporal ARP in SPI blocks

It is proposed to use more accurate disparity vector instead of DoNBDV/NBDV for temporal ARP in SPI blocks. This proposal additionally derives disparity vector (mvT) in SubPU utilized case and uses it instead of DoNBDV/NBDV (MVDisp). This modification only affects SubPU and temporal ARP (ivRefFlag==0) case.

I.8.5.3.3.7 Bilinear sample interpolation and residual prediction process

…

* If ivRefFlag is equal to 0 and RpRefIdxLX is not equal to −1, the variable availFlag is set equal to 1, the variable refIdxLX is set equal to RpRefIdxLX and the residual prediction motion vector scaling process as specified in subclause is invoked with the prediction list utilization variable equal to X, the motion vector mvLX, and the RefPicListX[ refIdxLX ] and as inputs and modified mvLX as output.
* ~~Otherwise, when ivRefFlag is equal to 1, the following applies:~~
* The derivation process for a motion vector from a reference block for residual prediction as specified in subclause is invoked with ( xP, yP ), nPbW and nPbH, RefPicListX[ refIdxLX ], and mvLX as inputs, and availFlagT, view order index refVIdx, motion vector mvT and prediction list utilization variable Y as outputs.
* If ivRefFlag is equal to 1, availFlag is set equal to availFlagT

When availFlag is equal to 1 and iv\_res\_pred\_weight\_idx is not equal to 0, the following applies:

* Depending on ivRefFlag, the variables rpPic, rpRefPic, mvRp and curRefIdx are derived as specified in the following:
  + If ivRefFlag is equal to 0 and availFlag, the following applies:
    - Let rpPic be the picture with PicOrderCnt( rpPic ) equal to PicOrderCntVal and ViewIdx equal to availFlagT ? refVIdx: RefViewIdx[ xP ][ yP ].
    - Let rpRefPic be the picture with PicOrderCnt( rpRefPic ) equal to RefPicListX[ RpRefIdxLX ] ) and ViewIdx equal to availFlagT ? refVIdx: RefViewIdx[ xP ][ yP ],
    - The variable mvRp is set equal to availFlagT ? mvT: MvDisp[ xP ][ yP ].
    - The variable curRefIdx is set equal to RpRefIdxLX.
  + Otherwise (ivRefFlag is equal to 1), the following applies:
    - Let rpPic be the picture RefPicListY[ RpRefIdxLY ].
    - Let rpRefPic be the picture with PicOrderCnt( rpRefPic ) equal to PicOrderCnt( rpPic ) and ViewIdx equal to RefViewIdx[ xP ][ yP ]
    - The variable mvRp is set equal to mvT.
    - The variable currRefIdx is set equal to RpRefIdxLY.

I.8.5.3.3.7.4 Derivation process for a motion vector from a reference block for residual prediction

Inputs to this process are:

* a luma location ( xP, yP ) of the top-left luma sample of the current prediction unit relative to the top-left luma sample of the current picture,
* variables nPbW and nPbH specifying the width and the height, respectively, of the current prediction unit,
* a reference picture refPic,
* a motion vector mvDisp

Outputs of this process are:

* a flag availFlag
* an view order index refVIdx,
* a motion vector mvT
* prediction list utilization variable Y.

The variable availFlag is set to 0 and the reference luma location ( xRef, yRef ) in refPicLX is derived by

* 1. xRef = Clip3( 0, PicWidthInSamplesL – 1, xP + ( nPSW >> 1 ) + ( ( mvDisp[ 0 ] + 2 ) >> 2 ) ) (‑231)  
     yRef = Clip3( 0, PicHeightInSamplesL – 1, yP + ( nPSH >> 1 ) + ( ( mvDisp[ 1 ] + 2 ) >> 2 ) ) (‑232)

Let variable refCU and refPU be the coding unit and prediction unit that cover the luma location ( xRef, yRef ) in refPic, respectively.

When the variable PredMode for the coding unit refCU is equal to MODE\_SKIP or MODE\_INTER, the following ordered steps apply for X in the range of 0 to 1, inclusive:

* + The variable refPredFlagLX is set equal to the prediction utilization flag predFlagLX of the prediction unit refPU.
  + When availFlag is equal to 0 and refPredFlagLX is equal to 1, the following applies:
    - Let refPicListRefX be the reference picture list X of refPic.
    - Let mvLX and refIdxLX be the motion vector and reference index of the prediction unit refPU corresponding to refPicListRefX, respectively.
    - When ~~refPicListRefX[ refIdxLX ] is a temporal reference picture of refPic~~ DiffPicOrderCnt( currPic, refPic) is equal to 0, DiffPicOrderCnt( refPic, refPicListRefX[ refIdxLX ]) is unequal to 0,and RpRefIdxLX is not equal to –1, availFlag is set to 1, refVIdx is set equal to the view order index of refPic, Y is set equal to X and the residual prediction motion vector scaling process as specified in subclause I.8.5.3.3.7.3 is invoked with the prediction list utilization variable equal to X, the motion vector mvLX, and the reference picture refPicListRefX[ refIdxLX ] as the inputs, and the output being mvT.
    - When DiffPicOrderCnt( refPic, refPicListRefX[ refIdxLX ]) is equal to 0, DiffPicOrderCnt( currPic, refPic) is unequal to 0, availFlag is set to 1, refVIdx is set equal to the view order index of refPicListRefX[ refIdxLX ], mvT is set equal to mvLX and Y is set equal to X.

# Complexity impact

The modification of disparity improvement incurs one inter-layer access at most 8 bytes (4 bytes for MV, 2 bytes for refIdx, 2 bytes for ViewId) per 8x8 PU, while chroma 4x4 off discards access from 64 bytes (4x4x2(arp)x2(Cb/Cr)x1(uni)) to 200 bytes (5x5x2(arp)x2(Cb/Cr)x2(bi)) per 8x8 PU.

Therefore it can be said that this combination retains the bandwidth reduction benefit without loss.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| anchor (HTM101) |  | CU | PU | Mult | Add | Bandwidth | Mult [%] | Add [%] | Bandwidth [%] |
| ARP | 64x64 | 64x64 | 36.4 | 24.2 | 9.4 | 64% | 50% | 93% |
| 32x32 | 32x32 | 36.8 | 24.4 | 9.8 | 64% | 50% | 97% |
| 16x16 | 16x16 | 37.5 | 24.8 | 10.6 | 66% | 51% | 105% |
| 8x8 | 8x8 | 39.0 | 25.5 | 12.4 | 68% | 53% | 123% |
| 4x4 chroma off (Hxxxx) |  | CU | PU | Mult | Add | Bandwidth | Mult [%] | Add [%] | Bandwidth [%] |
| ARP+G0033 | 64x64 | 64x64 | 36.4 | 24.2 | 9.4 | 64% | 50% | 93% |
| 32x32 | 32x32 | 36.8 | 24.4 | 9.8 | 64% | 50% | 97% |
| 16x16 | 16x16 | 37.5 | 24.8 | 10.6 | 66% | 51% | 105% |
| 8x8 | 8x8 | 30.0 | 19.0 | 9.3 | 53% | 39% | **92%** |
| 4x4  chroma off + improvement DV (Hxxxx) |  | CU | PU | Mult | Add | Bandwidth | Mult [%] | Add [%] | Bandwidth [%] |
| ARP+G0033 | 64x64 | 64x64 | 36.4 | 24.2 | 9.4 | 64% | 50% | 93% |
| 32x32 | 32x32 | 36.8 | 24.4 | 9.8 | 64% | 50% | 97% |
| 16x16 | 16x16 | 37.5 | 24.8 | 10.6 | 66% | 51% | 105% |
| 8x8 | 8x8 | 30.0 | 19.0 | 9.4 | 53% | 39% | **93%** |

# Simulation results

## Proposal

Table 2 Experimental results of proposal (anchor: HTM101)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | video 0 | video 1 | video 2 | video PSNR / video bitrate | video PSNR / total bitrate | synth PSNR / total bitrate | enc time | dec time | ren time |
| Balloons | 0.0% | 0.8% | 0.6% | 0.20% | 0.16% | 0.18% | 100.8% | 102.9% | 101.5% |
| Kendo | 0.0% | 0.1% | 0.1% | -0.04% | -0.10% | -0.01% | 100.3% | 102.3% | 99.9% |
| Newspaper\_CC | 0.0% | 0.1% | 0.1% | 0.03% | -0.01% | 0.02% | 100.0% | 99.1% | 99.9% |
| GT\_Fly | 0.0% | 0.2% | 0.0% | -0.02% | -0.01% | 0.03% | 100.1% | 99.9% | 100.7% |
| Poznan\_Hall2 | 0.0% | -0.5% | -0.8% | -0.24% | -0.24% | -0.17% | 100.1% | 100.8% | 99.6% |
| Poznan\_Street | 0.0% | -0.2% | -0.1% | -0.05% | -0.05% | -0.05% | 99.9% | 100.8% | 99.7% |
| Undo\_Dancer | 0.0% | -0.1% | -0.1% | -0.03% | -0.03% | -0.01% | 99.9% | 102.6% | 100.6% |
| Shark | 0.0% | 0.0% | 0.2% | -0.02% | -0.02% | -0.02% | 100.0% | 101.6% | 100.1% |
| 1024x768 | 0.0% | 0.3% | 0.3% | 0.06% | 0.02% | 0.06% | 100.4% | 101.4% | 100.4% |
| 1920x1088 | 0.0% | -0.1% | -0.2% | -0.07% | -0.07% | -0.04% | 100.0% | 101.1% | 100.1% |
| **average** | **0.0%** | **0.1%** | **0.0%** | **-0.02%** | **-0.04%** | **0.00%** | **100.1%** | **101.2%** | **100.2%** |

## Additional results

Table 3 Results of G0033 only (anchor: HTM101)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | video 0 | video 1 | video 2 | video PSNR / video bitrate | video PSNR / total bitrate | synth PSNR / total bitrate | enc time | dec time | ren time |
| Balloons | 0.0% | 1.0% | 1.0% | 0.30% | 0.26% | 0.28% | 100.2% | 99.5% | 100.8% |
| Kendo | 0.0% | 0.6% | 0.5% | 0.16% | 0.11% | 0.22% | 100.1% | 103.5% | 100.4% |
| Newspaper\_CC | 0.0% | 0.3% | 0.2% | 0.11% | 0.11% | 0.11% | 100.2% | 101.8% | 101.4% |
| GT\_Fly | 0.0% | 0.1% | 0.1% | 0.00% | -0.01% | 0.00% | 100.0% | 100.6% | 100.3% |
| Poznan\_Hall2 | 0.0% | 0.1% | 0.0% | 0.03% | 0.01% | -0.05% | 100.1% | 100.6% | 100.5% |
| Poznan\_Street | 0.0% | 0.0% | 0.2% | 0.03% | 0.03% | 0.02% | 99.7% | 99.5% | 99.6% |
| Undo\_Dancer | 0.0% | 0.1% | 0.1% | 0.01% | 0.01% | 0.00% | 99.6% | 102.6% | 100.0% |
| Shark | 0.0% | 0.4% | 0.4% | 0.05% | 0.02% | 0.04% | 100.1% | 100.5% | 99.9% |
| 1024x768 | 0.0% | 0.7% | 0.6% | 0.19% | 0.16% | 0.20% | 100.2% | 101.6% | 100.9% |
| 1920x1088 | 0.0% | 0.2% | 0.2% | 0.02% | 0.01% | 0.00% | 99.9% | 100.8% | 100.0% |
| **average** | **0.0%** | **0.4%** | **0.3%** | **0.08%** | **0.07%** | **0.08%** | **100.0%** | **101.1%** | **100.4%** |

Table 4 Results of simple combination of G0119 and G0033 (anchor: HTM101)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | video 0 | video 1 | video 2 | video PSNR / video bitrate | video PSNR / total bitrate | synth PSNR / total bitrate | enc time | dec time | ren time |
| Balloons |  |  |  | TBD |  |  |  |  |  |
| Kendo |  |  |  |  |  |  |  |  |  |
| Newspaper\_CC |  |  |  |  |  |  |  |  |  |
| GT\_Fly |  |  |  |  |  |  |  |  |  |
| Poznan\_Hall2 |  |  |  |  |  |  |  |  |  |
| Poznan\_Street |  |  |  |  |  |  |  |  |  |
| Undo\_Dancer |  |  |  |  |  |  |  |  |  |
| Shark |  |  |  |  |  |  |  |  |  |
| 1024x768 |  |  |  |  |  |  |  |  |  |
| 1920x1088 |  |  |  |  |  |  |  |  |  |
| **average** |  |  |  |  |  |  |  |  |  |

Table 5 Results of G0119 only (anchor: HTM101)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | video 0 | video 1 | video 2 | video PSNR / video bitrate | video PSNR / total bitrate | synth PSNR / total bitrate | enc time | dec time | ren time |
| Balloons | 0.0% | -0.6% | -0.7% | -0.22% | -0.21% | -0.24% |  |  |  |
| Kendo | 0.0% | -0.4% | -0.5% | -0.19% | -0.21% | -0.17% |  |  |  |
| Newspaper\_CC | 0.0% | -0.3% | -0.4% | -0.10% | -0.12% | -0.12% |  |  |  |
| GT\_Fly | 0.0% | -0.2% | -0.3% | -0.05% | -0.04% | 0.02% |  |  |  |
| Poznan\_Hall2 | 0.0% | -0.8% | -1.1% | -0.32% | -0.33% | -0.35% |  |  |  |
| Poznan\_Street | 0.0% | -0.4% | -0.3% | -0.08% | -0.08% | -0.07% |  |  |  |
| Undo\_Dancer | 0.0% | -0.4% | -0.6% | -0.12% | -0.12% | -0.12% |  |  |  |
| Shark | 0.0% | -0.5% | -0.5% | -0.09% | -0.08% | -0.10% |  |  |  |
| 1024x768 | 0.0% | -0.4% | -0.5% | -0.17% | -0.18% | -0.18% |  |  |  |
| 1920x1088 | 0.0% | -0.5% | -0.5% | -0.13% | -0.13% | -0.12% |  |  |  |
| **average** | **0.0%** | **-0.5%** | **-0.5%** | **-0.15%** | **-0.15%** | **-0.14%** |  |  |  |

# Conclusion

Because the proposal addresses the worst case memory bandwidth concerns with no average loss in CTC, it is recommended to adopt this method in 3D-HEVC.

# Patent rights declaration

**Sharp Corporation 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] T. Ikai, “CE4-related: ARP simplification”, JCT3V-G0033, JCT-3V 7th Meeting: San Jose, USA, 11 Jan. – 17 Jan. 2013

[2] L. Zhang, Y. Chen, M. Karczewicz, Q. Yu, S. Ma, “CE4: Further improvements on advanced residual prediction”, JCT3V-G0121, JCT-3V 7th Meeting: San Jose, USA, 11 Jan. – 17 Jan. 2013