|  |  |
| --- | --- |
| **Joint Collaborative Team on 3D Video Coding Extensions**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  5th Meeting: Vienna, AT, 27 July – 2 Aug. 2013 | Document: JCT3V-E0044\_r1 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **ARP simplification** | | |
| *Status:* | Input Document to JCT-3V | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Tomohiro Ikai  Yoshiya Yamamoto  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 presents a simplification method for advanced residual prediction (ARP) aiming to remove line buffer in context derivation and to reduce the number of weight candidates. The proposal uses two weights instead of the current three weights and restrict ARP enabling mode to merge mode with merge\_idx < 3 (disable for other merge block and AMVP block). The experimental result reportedly shows 0.0 %, 0.0% and 0.0% BD-rate change in texture, video and synth respectively and encoding time decreased by about 8 %.

# Introduction

Currently advanced residual prediction (ARP) refers top and left neighbouring block’s ARP weight to derive context, which needs additional line buffer. In addition, ARP uses two weights in all merge mode and AMVP mode, which needs a lot of encoder searches. The usage of two weight means three searches are needed to decide a weight. Small number of candidate is important in merge mode because accurate estimation is needed to achieve a good coding efficiency. This proposal presents a simplification method to remove line buffer in advanced residual prediction (ARP) context derivation and reduce the number of weight candidates.

# Proposal

## Line buffer removal

Line buffer is eliminated by removing top (condA) neighbouring block reference in ARP context derivation.

Table H‑31 – Specification of ctxIdxInc using left and above syntax elements

|  |  |  |  |
| --- | --- | --- | --- |
| **Syntax element** | **condL** | **condA** | **ctxIdxInc** |
| iv\_res\_pred\_weight\_idx | iv\_res\_pred\_weight\_idx [ xL ][ yL ] | ~~iv\_res\_pred\_weight\_idx [ xA ][ yA ]~~ | ( condL && availableL ) ~~+ ( condA && availableA )~~ |

## Number of weights candidates reduction and mode restriction

I the proposal, the number of weight decreases from 3 to 2. By this change the following benefit is achieved.

* the number of weight decreases from two to one
* shift operation removes in ARP prediction
* the number of context decreases from 8 to 4
* binarization text can be removed

Additionally ARP is further disabled when merge\_idx is equal to or larger than 3 or AMVP mode.

* ARP is enabled (merge\_flag == 1 && merge\_idx < 3)
* ARP is disabled (merge\_flag == 1 && merge\_idx >= 3)
* ARP is disabled (merge\_flag == 0 or AMVP)

The number of merge&weight candidates reduced by half (18 to 9) as shown Table 1.

Table 1 Comparison of the number of merge&weights candidate

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Merge | | | | | | AMVP |
|  | merge\_idx  =0 | merge\_idx  =1 | merge\_idx  =2 | merge\_idx  =3 | merge\_idx  =4 | merge\_idx  =5 | total  in merge |  |
| HTM70 | 3 | 3 | 3 | 3 | 3 | 3 | 18  (=3\*6) | 3 |
| Proposal | 2 | 2 | 2 | 1 | 1 | 1 | 9  (=2\*3+3) | 0 |

|  |  |
| --- | --- |
| coding\_unit( x0, y0, log2CbSize , ctDepth) { | **Descriptor** |
| if( transquant\_bypass\_enable\_flag ) { |  |
| **cu\_transquant\_bypass\_flag** | ae(v) |
| } |  |
| if( slice\_type != I ) |  |
| **skip\_flag**[ x0 ][ y0 ] | ae(v) |
| if( skip\_flag[ x0 ][ y0 ] ) { |  |
| prediction\_unit( x0, y0, log2CbSize ) |  |
| if ( iv\_res\_pred\_flag && TempRefPicInListsFlag && merge\_idx[ x0 ][ y0 ] < 3) |  |
| iv\_res\_pred\_weight\_flag | ae(v) |
| .. |  |
| } |  |
| else { |  |
| … |  |
| } else { |  |
| if( PartMode = = PART\_2Nx2N ) { |  |
| prediction\_unit( x0, y0, nCbS, nCbS ) |  |
| if ( iv\_res\_pred\_flag && TempRefPicInListsFlag && (merge\_flag[ x0 ][ y0 ] && merge\_idx[ x0 ][ y0 ] < 3)) |  |
| iv\_res\_pred\_weight\_flag | ae(v) |
| } |  |

**iv\_res\_pred\_weight\_flag** specifies the index of the weighting factor used for residual prediction. iv\_res\_pred\_weight\_idx equal to 0 specifies that residual prediction is not used for the current coding unit. iv\_res\_pred\_weight\_idx equal to 1 specifies that residual prediction is used for the current coding unit. When not present, the value of iv\_res\_pred\_weight\_flag is inferred to be equal to 0.

~~The variable shiftVal is set equal to ( iv\_res\_pred\_weight\_idx − 1 ).~~

The modified prediction samples predSamplesLXL[ x ][ y ] with x = 0..( nPbW ) − 1 and y = 0..( nPbH ) − 1 are derived as specified in the following:

* 1. predSamplesLXL[ x ][ y ] = predSamplesLXL[ x ][ y ] +   
      ( ( currIvSamplesLXL[ x ][ y ] − refIvSamplesLXL[ x ][ y ] )  ~~>>  shiftVal~~) (H‑202)

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 ] +   
      ( ( currIvSamplesLXCb[ x ][ y ] − refIvSamplesLXCb[ x ][ y ] )  ~~>>  shiftVal~~) (H‑203)

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:

predSamplesLXCr[ x ][ y ] = predSamplesLXCr[ x ][ y ] +   
 ( ( currIvSamplesLXCr[ x ][ y ] − refIvSamplesLXCr[ x ][ y ] )  ~~>>  shiftVal~~) (H‑204)

Table H‑12 – Association of ctxIdx and syntax elements for each initializationType in the initialization process

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Syntax element** | **ctxIdxTable** | **initType** | | |
| **0** | **1** | **2** |
| coding\_unit( ) depth\_mode\_parameters( ) | depth\_intra\_mode |  | 0..7 | 8..15 | 16..23 |
| wedge\_full\_tab\_idx |  | 0 | 1 | 2 |
| wedge\_predtex\_tab\_idx |  | 0 | 1 | 2 |
| dmm\_delta\_end\_flag dmm\_delta\_end\_abs\_minus1 |  | 0 | 1 | 2 |
| dmm\_dc\_flag |  | 0 | 1 | 2 |
| dmm\_dc\_abs |  | 0 | 1 | 2 |
| edge\_code |  | 0 | 1 | 2 |
| iv\_res\_pred\_weight\_flag |  |  | 0..1  ~~0..3~~ | 2..3  ~~4..7~~ |
| ic\_flag |  |  | 0 | 1 |
| edge\_dc\_flag |  | 0 | 1 | 2 |
| edge\_dc\_abs |  | 0 | 1 | 2 |
| sdc\_residual\_flag |  | 0 | 1 | 2 |
| sdc\_residual\_abs\_minus1 | prefix:  suffix: na | prefix:0 suffix: na | prefix:1 suffix: na | prefix:0 suffix: na |

Table H‑18 – Values of variable initValue for iv\_res\_pred\_weight\_idx ctxIdx

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Initialization variable** | **iv\_res\_pred\_weight\_idx** | | | | | | | |
| **0** | **1** | **2** | **3** | **~~4~~** | **~~5~~** | **~~6~~** | **~~7~~** |
| **initValue** | 154 | 154 | 154 | 154 | ~~154~~ | ~~154~~ | ~~154~~ | ~~154~~ |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ~~iv\_res\_pred\_weight\_idx~~ | ~~1~~ | ~~TU, cMax = 2~~ | ~~0~~ |  | ~~0~~ |
| ~~2~~ | ~~0~~ |  | ~~1~~ |

# Simulation results

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1 Simulation results | 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% | -0.1% | -0.1% | -0.1% | -0.1% | -0.1% | 92.4% | 101.4% | 99.5% |
| Kendo | 0.0% | 0.2% | 0.0% | 0.1% | 0.1% | 0.0% | 91.8% | 96.9% | 97.4% |
| Newspaper\_CC | 0.0% | 0.7% | 0.6% | 0.3% | 0.3% | 0.3% | 92.5% | 105.4% | 100.1% |
| GT\_Fly | 0.0% | -0.1% | -0.3% | 0.0% | 0.0% | -0.1% | 94.7% | 99.8% | 100.5% |
| Poznan\_Hall2 | 0.0% | 0.1% | 0.2% | 0.0% | 0.0% | 0.0% | 91.6% | 100.8% | 99.8% |
| Poznan\_Street | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 91.0% | 102.3% | 100.7% |
| Undo\_Dancer | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 92.3% | 103.3% | 100.9% |
| 1024x768 | 0.0% | 0.2% | 0.2% | 0.1% | 0.1% | 0.1% | 92.2% | 101.2% | 99.0% |
| 1920x1088 | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 92.4% | 101.6% | 100.5% |
| **average** | **0.0%** | **0.1%** | **0.1%** | **0.0%** | **0.0%** | **0.0%** | **92.3%** | **101.4%** | **99.8%** |

Test 1: Liner buffer removal

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 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% | 0.1% | 0.1% | 0.04% | 0.04% | 0.01% | 99.8% | 99.7% | 99.3% |
| Kendo | 0.0% | 0.3% | 0.3% | 0.12% | 0.14% | 0.14% | 99.4% | 98.5% | 99.6% |
| Newspaper\_CC | 0.0% | 0.2% | 0.1% | 0.06% | 0.05% | 0.14% | 100.1% | 101.7% | 100.5% |
| GT\_Fly | 0.0% | 0.3% | 0.1% | 0.02% | 0.03% | 0.07% | 99.8% | 100.0% | 100.3% |
| Poznan\_Hall2 | 0.0% | 0.3% | 0.2% | 0.08% | 0.07% | 0.08% | 99.5% | 100.9% | 99.9% |
| Poznan\_Street | 0.0% | 0.0% | 0.1% | 0.01% | 0.00% | 0.01% | 99.4% | 100.1% | 100.7% |
| Undo\_Dancer | 0.0% | 0.1% | 0.0% | 0.01% | 0.00% | 0.03% | 99.6% | 101.2% | 99.8% |
| 1024x768 | 0.0% | 0.2% | 0.2% | 0.07% | 0.07% | 0.10% | 99.8% | 100.0% | 99.8% |
| 1920x1088 | 0.0% | 0.2% | 0.1% | 0.03% | 0.02% | 0.05% | 99.6% | 100.6% | 100.2% |
| **average** | **0.0%** | **0.2%** | **0.1%** | **0.05%** | **0.05%** | **0.07%** | **99.7%** | **100.3%** | **100.0%** |

Test 2: 2 weight

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 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% | -0.3% | -0.2% | -0.07% | -0.07% | -0.11% | 95.4% | 100.9% | 99.3% |
| Kendo | 0.0% | 0.2% | -0.1% | 0.08% | 0.11% | 0.03% | 95.2% | 96.8% | 97.1% |
| Newspaper\_CC | 0.0% | 0.8% | 0.8% | 0.36% | 0.35% | 0.47% | 95.8% | 107.7% | 100.6% |
| GT\_Fly | 0.0% | 0.2% | -0.2% | 0.04% | 0.05% | -0.03% | 96.6% | 99.4% | 99.3% |
| Poznan\_Hall2 | 0.0% | -0.2% | 0.2% | 0.11% | 0.12% | 0.07% | 94.1% | 102.0% | 98.0% |
| Poznan\_Street | 0.0% | 0.2% | 0.3% | 0.15% | 0.15% | 0.12% | 94.1% | 102.3% | 98.0% |
| Undo\_Dancer | 0.0% | -0.1% | 0.0% | 0.01% | 0.00% | 0.04% | 94.8% | 98.7% | 97.0% |
| 1024x768 | 0.0% | 0.2% | 0.2% | 0.12% | 0.13% | 0.13% | 95.5% | 101.8% | 99.0% |
| 1920x1088 | 0.0% | 0.1% | 0.1% | 0.07% | 0.08% | 0.05% | 94.9% | 100.6% | 98.1% |
| **average** | **0.0%** | **0.1%** | **0.1%** | **0.09%** | **0.10%** | **0.08%** | **95.1%** | **101.1%** | **98.5%** |

Test 3: Liner buffer removal + 2 weight (Test1 + Test2)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 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% | -0.1% | -0.1% | -0.05% | -0.04% | -0.09% | 95.4% | 100.5% | 99.9% |
| Kendo | 0.0% | 0.1% | -0.1% | 0.06% | 0.08% | 0.06% | 95.8% | 98.1% | 100.0% |
| Newspaper\_CC | 0.0% | 0.5% | 0.6% | 0.24% | 0.23% | 0.31% | 95.4% | 103.1% | 99.4% |
| GT\_Fly | 0.0% | 0.0% | -0.3% | 0.01% | 0.01% | -0.05% | 96.8% | 98.2% | 99.7% |
| Poznan\_Hall2 | 0.0% | 0.0% | 0.6% | 0.13% | 0.13% | 0.12% | 95.1% | 100.6% | 99.7% |
| Poznan\_Street | 0.0% | -0.1% | 0.1% | 0.03% | 0.04% | 0.03% | 94.7% | 102.2% | 100.0% |
| Undo\_Dancer | 0.0% | 0.0% | -0.1% | 0.00% | 0.00% | 0.05% | 95.7% | 100.3% | 100.1% |
| 1024x768 | 0.0% | 0.2% | 0.1% | 0.09% | 0.09% | 0.09% | 95.5% | 100.6% | 99.7% |
| 1920x1088 | 0.0% | 0.0% | 0.1% | 0.04% | 0.04% | 0.04% | 95.6% | 100.4% | 99.9% |
| **average** | **0.0%** | **0.1%** | **0.1%** | **0.06%** | **0.06%** | **0.06%** | **95.6%** | **100.4%** | **99.8%** |

# Conclusion

This proposal presents a simplification method for advanced residual prediction (ARP) aiming to remove line buffer in context derivation and to reduce the number of weight candidates. The experimental result reportedly shows 0.0 %, 0.0% and 0.0% BD-rate change in texture, video and synth respectively and encoding time decreased by about 8 %.

# References

[1] L. Zhang, Y. Chen, X. Li, M. Karczewicz, “CE4: Advanced residual prediction for multiview coding” JCT3V-D0177, JCT3V 4th Meeting: Incheon, KR, 20–26 Apr. 2013

# Patent rights declaration(s)

**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).**