|  |  |
| --- | --- |
| **Joint Collaborative Team on 3D Video Coding Extensions**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  4th Meeting: Incheon, KR, 20–26 Apr. 2013 | Document: JCT3V-D0063 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **CE5 related: Simplified residual prediction with luma clipping and chroma removal** | | |
| *Status:* | Input Document | | |
| *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

At the last meeting, the complexity of residual prediction was discussed and simpler solution of JCT3V-C0138 was adopted. This contribution proposes further simplifications based on the former decision. Specifically, the luma residual samples are clipped in 8 bit range to reduce memory size for storing residual signal. In addition, the chroma residual prediction is completely removed considering the negligible gain in chroma component doesn’t justify its complexity. The experimental results show the proposed simplification affects no significant bd-rate difference compared to HTM50 with chroma residual bug-fix.

# Proposal

**Clipping of luma residual samples**

* Luma residual samples are clipped in the range of 8bit [-128, 127] before residual interpolation.

**Removal of chroma residual prediction**

* We propose to remove chroma residual prediction from the spec.

# Proposed Text

The changes related to luma residual clipping are highlighted with Green.

The changes related to chroma residual removal are highlighted with Yellow.

*Based on JCT3V-C1005\_spec\_d0.doc [2]*

When ResPredFlag[ xC + xB ][ yC + yB ] is equal to 1, the inter-view residual prediction process as specified in subclause H.8.5.2.2.6 is invoked with the luma locations ( xC, yC ) and ( xC + xB, yC + yB ), the size of the current luma coding block nCS, the variables nPSW and nPSH, and the arrays predSamplesL~~, predSamplesCb, and predSamplesCr~~ as the inputs and the outputs are modified versions of the arrays predSamplesL~~, predSamplesCb, and predSamplesCr~~.

**H.8.5.2.2.6** **Inter-view residual prediction process**

The process is only invoked if res\_pred\_flag is equal to 1.

Inputs to this process are:

* a luma location ( xC, yC ) specifying the top-left sample of the current luma coding block relative to the top left luma sample of the current picture,
* 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,
* a variable nCS specifying the size of the current luma coding block,
* variables nPSW and nPSH specifying the width and the height, respectively, of the current prediction unit,prediction list utilization flags, predFlagL0 and predFlagL1,
* a (nPSW)x(nPSH) array predSamplesL of luma prediction samples,
* ~~two (nPSW / 2)x(nPSH / 2) arrays predSamples~~~~Cb~~ ~~and predSamples~~~~Cr~~ ~~of chroma prediction samples.~~

Output of this process are:

* a modified version of the (nPSW)x(nPSH) array predSamplesL,
* ~~a modified versions of the (nPSW / 2)x(nPSH / 2) arrays predSamples~~~~Cb~~ ~~and predSamples~~~~Cr~~~~.~~

The derivation process for a disparity vector as specified in subclause H.8.5.4 is invoked with the luma locations ( xC, yC ) and ( xP, yP ), the coding block size nCS, the variables nPSW and nPSH, the partition index partIdx and the variable deriveFromDepthFlag being equal to 0, as the inputs and the outputs are the view order index refViewIdx, the flag availableDV and the disparity vector mvDisp.[Ed. (GT) partIdx is missing as input to this subclause.].

[Ed. (GT). In software refViewIdx is set equal to 0. Might be better to reuse disparity and refViewIdx from inter-view motion vector prediction. ].

Let refResSamplesL be the (PicWidthInSamplesL)x(PicHeightInSamplesL) array of luma residual samples ResSamplesL: of the view component with ViewIdx equal to refViewIdx. ~~Let refResSamples~~~~Cb~~ ~~and refResSamples~~~~Cr~~ ~~be the (PicWidthInSamples~~~~L~~~~/ 2)x(PicHeightInSamples~~~~L~~~~/ 2) arrays of Cb and Cr residual samples ResSamples~~~~Cb~~ ~~and ResSamples~~~~Cr~~~~, respectively, for inter-coded coding units for the view component with ViewIdx equal to refViewIdx.~~

When the flag availableDV is equal to 0 the whole decoding process of this sub-clause terminates.

For y proceeding over the values 0..(nPSH – 1) and x proceeding over the values 0..(nPSW – 1), the following ordered steps apply.

* 1. The variables xR0, xR1, yR0, yR1 and w0, w1, w2, and w3 are derived by
     + 1. xR0 = Clip3( 0, PicWidthInSamplesL – 1, xP + x + (mvDisp[ 0 ] >> 2 ) ) (H‑233)  
          xR1 = Clip3( 0, PicWidthInSamplesL – 1, xP + x + (mvDisp[ 0 ] >> 2 ) + 1 ) (H‑234)  
          yR0 = Clip3( 0, PicHeightInSamplesL – 1, yP + y + (mvDisp[ 1 ] >> 2 ) ) (H‑235)  
          yR1 = Clip3( 0, PicHeightInSamplesL – 1, yP + y + (mvDisp[ 1 ] >> 2 ) + 1 ) (H‑236)  
          w0 = 4 – mvDisp[ 0 ] + ( ( mvDisp[ 0 ] >> 2 ) << 2 ) (H‑237)  
          w1 = mvDisp[ 0 ] − ( ( mvDisp[ 0 ] >> 2 ) << 2 ) (H‑238)  
          w2 = 4 – mvDisp[ 1 ] + ( ( mvDisp[ 1 ] >> 2 ) << 2 ) (H‑239)  
          w3 = mvDisp[ 1 ] − ( ( mvDisp[ 1 ] >> 2 ) << 2 ) (H‑240)
  2. The sample predSamplesL[ x ][ y ] is modified by
     + 1. deltaL1 = ( w0 \* Clip3(-128, 127, refResSamplesL[ xR0 ][ yR0 ]) + w1 \* Clip3(-128, 127, refResSamplesL[ xR1 ][ yR0 ]) + 4 ) >> 3 (H‑241)  
          deltaL2 = ( w0 \* Clip3(-128, 127, refResSamplesL[ xR0 ][ yR1 ]) + w1 \* Clip3(-128, 127, refResSamplesL[ xR1 ][ yR1 ]) + 4 ) >> 3 (H‑242)  
          deltaL = ( w2 \* deltaL1 + w3 \* deltaL2 + 4 ) >> 3 (H‑243)  
          predSamplesL[ x ][ y ] = predSamplesL[ x ][ y ] + deltaL (H‑244)

~~For y proceeding over the values 0..(nPSH / 2 – 1) and x proceeding over the values 0..(nPSW / 2 – 1), the following ordered steps are specified:~~

* 1. ~~The variables xR0, xR1, yR0, yR1 and w0, w1, w2, and w3 are derived by~~
     + 1. ~~xR0 = Clip3( 0, PicWidthInSamples~~~~L~~~~/ 2 – 1, xP / 2 + x + (mvDisp[ 0 ] >> 3 ) ) (‑245)  
          xR1 = Clip3( 0, PicWidthInSamples~~~~L~~~~/ 2 – 1, xP / 2 + x + (mvDisp[ 0 ] >> 3 ) + 1 ) (‑246)  
          yR0 = Clip3( 0, PicHeightInSamples~~~~L~~~~/ 2 – 1, yP / 2 + y + (mvDisp[ 1 ] >> 3 ) ) (‑247)  
          yR1 = Clip3( 0, PicHeightInSamples~~~~L~~~~/ 2 – 1, yP / 2 + y + (mvDisp[ 1 ] >> 3 ) + 1 ) (‑248)  
          w0 = 8 – mvDisp[ 0 ] + ( (mvDisp[ 0 ] >> 3 ) << 3 ) (‑249)  
          w1 = mvDisp[ 0 ] − ( (mvDisp[ 0 ] >> 3 ) << 3 ) (‑250)  
          w2 = 8 – mvDisp[ 1 ] + ( (mvDisp[ 1 ] >> 3 ) << 3 ) (‑251)  
          w3 = mvDisp[ 1 ] − ( (mvDisp[ 1 ] >> 3 ) << 3 ) (‑252)~~
  2. ~~The sample predSamplesCb[ x ][ y ] is modified by~~
     + 1. ~~deltaCb1 = ( w0 \* refResSamplesCb[ xR0 ][ yR0 ] + w1 \* refResSamplesCb[ xR1 ][ yR0 ] + 8 ) >> 4 (‑253)  
          deltaCb2 = ( w0 \* refResSamplesCb[ xR0 ][ yR1 ] + w1 \* refResSamplesCb[ xR1 ][ yR1 ] + 8 ) >> 4 (‑254)  
          deltaCb = ( w2 \* deltaCb1 + w3 \* deltaCb2 + 8 ) >> 4 (‑255)  
          predSamplesCb[ x ][ y ] = predSamplesCb[ x ][ y ] + deltaCb (‑256)~~
  3. ~~The sample predSamplesCr[ x ][ y ] is modified by~~
     + 1. ~~deltaCr1 = ( w0 \* refResSamplesCr[ xR0 ][ yR0 ] + w1 \* refResSamplesCr[ xR1 ][ yR0 ] + 8 ) >> 4 (‑257)  
          deltaCr2 = ( w0 \* refResSamplesCr[ xR0 ][ yR1 ] + w1 \* refResSamplesCr[ xR1 ][ yR1 ] + 8 ) >> 4 (‑258)  
          deltaCr = ( w2 \* deltaCr1+ w3 \* deltaCr2) + 8 ) >> 4 (‑259)  
          predSamplesCr[ x ][ y ] = predSamplesCr[ x ][ y ] + deltaCr (‑260)~~

# Experimental results

The experimental results on CTC is shown in Table 1.

Table 1 Results of luma clipping + removal of chroma residual

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 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% | -0.5% | -0.4% | -0.15% | -0.11% | -0.17% | 100.7% | 96.6% | 98.9% |
| Kendo | 0.0% | -0.2% | -0.1% | -0.07% | -0.06% | -0.02% | 99.8% | 101.3% | 101.1% |
| Newspapercc | 0.0% | -0.2% | -0.1% | -0.03% | -0.01% | -0.03% | 99.2% | 98.4% | 100.8% |
| GhostTownFly | 0.0% | 0.3% | 0.2% | 0.05% | 0.05% | 0.05% | 100.6% | 100.5% | 100.5% |
| PoznanHall2 | 0.0% | -0.3% | -0.1% | -0.08% | -0.08% | -0.04% | 99.8% | 97.1% | 99.9% |
| PoznanStreet | 0.0% | -0.2% | -0.2% | -0.05% | -0.05% | -0.06% | 99.9% | 95.7% | 99.9% |
| UndoDancer | 0.0% | 0.0% | 0.0% | 0.00% | 0.00% | -0.01% | 99.4% | 98.4% | 100.8% |
| 1024x768 | 0.0% | -0.3% | -0.2% | -0.08% | -0.06% | -0.07% | 99.9% | 98.8% | 100.3% |
| 1920x1088 | 0.0% | -0.1% | 0.0% | -0.02% | -0.02% | -0.02% | 100.0% | 97.9% | 100.3% |
| **average** | **0.0%** | **-0.2%** | **-0.1%** | **-0.05%** | **-0.04%** | **-0.04%** | **99.9%** | **98.3%** | **100.3%** |

Table 2 Results of removal of chroma residual

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 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% | -0.5% | -0.4% | -0.15% | -0.11% | -0.17% | 100.7% | 96.6% | 98.9% |
| Kendo | 0.0% | -0.2% | -0.1% | -0.07% | -0.06% | -0.02% | 99.8% | 101.3% | 101.1% |
| Newspapercc | 0.0% | -0.2% | -0.1% | -0.03% | -0.01% | -0.03% | 99.2% | 98.4% | 100.8% |
| GhostTownFly | 0.0% | 0.3% | 0.2% | 0.05% | 0.05% | 0.05% | 100.6% | 100.5% | 100.5% |
| PoznanHall2 | 0.0% | -0.3% | -0.1% | -0.08% | -0.08% | -0.04% | 99.8% | 97.1% | 99.9% |
| PoznanStreet | 0.0% | -0.2% | -0.2% | -0.05% | -0.05% | -0.06% | 99.9% | 95.7% | 99.9% |
| UndoDancer | 0.0% | 0.0% | 0.0% | 0.00% | 0.00% | -0.01% | 99.4% | 98.4% | 100.8% |
| 1024x768 | 0.0% | -0.3% | -0.2% | -0.08% | -0.06% | -0.07% | 99.9% | 98.8% | 100.3% |
| 1920x1088 | 0.0% | -0.1% | 0.0% | -0.02% | -0.02% | -0.02% | 100.0% | 97.9% | 100.3% |
| **average** | **0.0%** | **-0.2%** | **-0.1%** | **-0.05%** | **-0.04%** | **-0.04%** | **99.9%** | **98.3%** | **100.3%** |

We have compared the bug-fixed version and chroma removal version as shown in Table 3. The difference is less than 0.05 % in Total and 0.16 % and 0.17% in video U/V respectively. We think that gain doesn’t justify the complexity.

Table 3 Results of removal of chroma residual (anchor is HTM50 + FIX\_CHROMA\_RESIDUAL\_C0129)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | video 0 | video 1 | video 2 | video only | synthesized only | coded & synthesized | video only U | video only V |
| Balloons | 0.0% | 0.2% | 0.4% | 0.1% | 0.1% | 0.1% | 0.4% | 0.4% |
| Kendo | 0.0% | 0.0% | 0.1% | 0.0% | 0.1% | 0.1% | 0.1% | 0.2% |
| Newspapercc | 0.0% | 0.1% | 0.1% | 0.0% | 0.0% | 0.0% | 0.1% | 0.1% |
| GhostTownFly | 0.0% | -0.1% | 0.2% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| PoznanHall2 | 0.0% | 0.0% | 0.3% | 0.1% | 0.1% | 0.1% | 0.2% | 0.3% |
| PoznanStreet | 0.0% | -0.1% | 0.1% | 0.0% | 0.0% | 0.0% | 0.4% | 0.2% |
| UndoDancer | 0.0% | 0.0% | 0.0% | 0.0% | 0.1% | 0.0% | 0.1% | 0.1% |
| 1024x768 | 0.0% | 0.1% | 0.2% | 0.1% | 0.1% | 0.1% | 0.2% | 0.2% |
| 1920x1088 | 0.0% | 0.0% | 0.1% | 0.0% | 0.0% | 0.0% | 0.2% | 0.1% |
| **average** | **0.00%** | **0.02%** | **0.16%** | **0.03%** | **0.05%** | **0.05%** | **0.16%** | **0.17%** |

Table 4 Results of fix chroma residual (HTM50 + FIX\_CHROMA\_RESIDUAL\_C0129)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 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% | -0.7% | -0.7% | -0.25% | -0.21% | -0.25% | 99.7% | 100.1% | 99.5% |
| Kendo | 0.0% | -0.2% | -0.2% | -0.08% | -0.06% | -0.11% | 99.3% | 100.2% | 100.7% |
| Newspapercc | 0.0% | -0.2% | -0.2% | -0.08% | -0.05% | -0.07% | 98.6% | 102.0% | 101.2% |
| GhostTownFly | 0.0% | 0.3% | 0.0% | 0.04% | 0.04% | 0.02% | 99.3% | 101.3% | 99.8% |
| PoznanHall2 | 0.0% | -0.3% | -0.5% | -0.15% | -0.14% | -0.09% | 99.2% | 98.7% | 99.8% |
| PoznanStreet | 0.0% | -0.2% | -0.3% | -0.04% | -0.04% | -0.06% | 99.6% | 98.1% | 99.8% |
| UndoDancer | 0.0% | 0.0% | 0.0% | 0.00% | 0.00% | -0.02% | 98.8% | 99.7% | 100.2% |
| 1024x768 | 0.0% | -0.4% | -0.4% | -0.14% | -0.11% | -0.14% | 99.2% | 100.8% | 100.5% |
| 1920x1088 | 0.0% | 0.0% | -0.2% | -0.04% | -0.03% | -0.04% | 99.2% | 99.5% | 99.9% |
| **average** | **0.0%** | **-0.2%** | **-0.3%** | **-0.08%** | **-0.07%** | **-0.08%** | **99.2%** | **100.0%** | **100.1%** |

# Conclusion

In this contribution, the luma residual samples are clipped in 8 bit range to reduce memory size for storing residual signal. Also the chroma residual prediction is completely removed considering the negligible gain in chroma component doesn’t justify its complexity. The experimental results show the proposed simplification affects no significant bd-rate difference compared to HTM50 with chroma residual bug-fix.

It is recommended to adopt this method in the next HTM.

# References

[1] J. An, K. Zhang, J.-L. Lin, S. Lei, “3D-CE4.h results on removal of parsing dependency for inter-view residual prediction”, JCT3V-C0138, JCT3V 3rd Meeting: Geneva, CH, 17–23 Jan. 2013

[2] G. Tech, K. Wegner, Y. Chen, S. Yea, “3D-HEVC Test Model 3”, JCT3V-C1005, JCT3V 3rd Meeting: Geneva, CH, 17–23 Jan. 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).**