|  |  |
| --- | --- |
| **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-D0061 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **3D-CE5.h related: Illumination compensation regression improvement and simplification** | | |
| *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

In the current 3D-HEVC HTM, illumination compensation (IC) utilizes the linear regression model in which parameters are derived by least square error (LSE) algorithm. However these LSE cost is known not robust in real world where a lot of outliner and noise exists. In this contribution, the parameter cost, which is optimized to IC, is added. Additionally we simplify the IC prediction process and IC parameter derivation process by using constant shift operation rather than adaptive shift operation and by restricting its application to 2Nx2N and L0 only. The experimental results show the proposed regression improvement shows 0.2 % gain in total video. The simplification doesn’t change the coding efficiency. The combination of the regression improvement and the simplification shows 0.2% gain in total video.

This combination result of this contribution (JCT3V-D0061) and a related contribution by the author (JCT3V-D0062) is reported in JCT3V-D0062.

# Introduction

**Regression improvement**

The current illumination compensation in HTM uses the linear regression model. When we denotes x is input and y is output, the regression model is described as

y = ax + b

In each PU, the parameters are derived with the least square error (LSE) algorithm, where the following error function is minimized.

E (a, b) =

The normal equation is formed as follows.

Then parameter a is solved with

In HTM, the above derivation is computed in integer form.

**Simplification**

Current illumination compensation has several complexity problems:

* adaptive shift operation in IC prediction
* small block size
* potential bi-prediction usage.
* parsing dependency

Specifically, adaptive shift operation is not effectively implemented by SIMD operation. It should be changed to **constant shift operation**.

Smallest block size is 8x16 and 16x8 in luma. If we use **2Nx2N only** we can double the block size.

In most cases, IC is used in only uni-prediction case but it can be used in bi-prediction case. Which means two IC process for L0 and L1 is needed. Considering this problem, **L0 only** restriction is considered useful.

Parsing dependency is addressed in another proposal of JCT3V-Dxxxx [1] by the author.

# Proposal

## Regression improvement

We propose the following error function E (a, b ) in regression in which parameter cost is added.

E (a, b ) =

If we utilized the parameter costs where parameter a should be near 1, we can get more robust prediction parameters.

The following is the revised normal equation.

Then parameter a is solved with

In this proposal, is set equal to

## Simplification

**Constant shift**

We propose that adaptive shift to constant shift (5 bit shits). This change also simplifies the IC parameter derivation process by removing a complex decIcShift operation which is applied on the condition that an intermediate value invPsIcWeight is less than −26 or greater than or equal to 26. Further mode the parameter range is clipped within unsigned 8 bit range.

**2Nx2N only**

We propose that illumination compensation is only applied in 2Nx2N. The minimum PU size is changed into 8x8 rather than 4x8/8x4

**L0 only**

We propose that illumination compensation is only applied in L0 only to remove bi-prediction IC prediction case. This restriction also simplifies anyIvRefPicFlag derivation because it only refer L0 related parameters (predFlagL0, refIdxL0 and RefPicListL0) rather than L0 and L1 related parameters (predFlagLX, refIdxLX and RefPicListLX).

Table 1 Comparison HTM and proposal

|  |  |  |
| --- | --- | --- |
|  | HTM | Proposal |
| Shift operation | Adaptive (up to 13 bit) | Constant (5bit) |
| Parameter range | -64 to 63 (signed 7 bit) | 0 to 64 (unsigned 7 bit) |
| Smallest block | 16x8 / 8x16 | 16x16 (2Nx2N only) |
| Reference in anyIvRefPicFlag checking | All PU and all direction | 1 PU (2Nx2N only) and one direction (L0 only) |
| Bi-prediction usage | Yes | No (L0 only) |

# Highlight of proposed text

In this section, highlight of proposed text *based on JCT3V-C1005\_spec\_d0.doc [2]* is shown. Full text is available in a separate document.

*Regression improvement is highlighted with Yellow*

*Constant shift simplification are highlighted with Green*

*2Nx2N and L0 only simplification are highlighted with Light blue*

|  |  |
| --- | --- |
| coding\_unit( x0, y0, log2CbSize , ctDepth) { | **Descriptor** |
| … |  |
| if ( icEnableFlag && PartMode == 2Nx2N ) |  |
| **ic\_flag** | ae(v) |

The variable anyIvRefPicFlag specifies whether one or more prediction units of the current coding unit utilize an inter-view reference picture. The variable anyIvRefPicFlag is initially set equal to 0. When PredMode[ x0 ][ y0 ] is not equal to MODE\_INTRA the following applies for X being replaced by 0 and 1, and Y being equal to 1−X.

* If DepthFlag is equal to 0, the following applies:

anyIvRefPicFlag =   
(inter\_pred\_idc[ x0 ][ y0 ]  = Pred\_L0 && refViewIdxL0[ x0 ][ y0 ] ! = ViewIdx)   (‑)

or

anyIvRefPicFlag =   
(predFlagL0[ x0 ][ y0 ]  = 1 && ViewIdx of the RefPicListL0[ refIdxL0 ] [ x0 ][ y0 ] ! = ViewIdx)   (‑)

**Illumination compensated sample prediction process**

* If predFlagL0 is equal to 1 and predFlagL1 is equal to 0,
  + 1. clipPredVal= Clip3( 0, ( 1 << bitDepth ) − 1, ( predSamplesL0[ x ][ y ] + offset1 ) >> shift1 ) (‑184)
    2. predSamples[ x ][ y ] = !puIcFlagL0 ? clipPredVal :   
       ( Clip3( 0, ( 1 << bitDepth ) − 1, ( clipPredVal \* icWeightL0 ) >> 5 ) + icOffsetL0 ) (‑185)
* Otherwise, if predFlagL0 is equal to 0 and predFlagL1 is equal to 1,
  + 1. (‑186)
    2. predSamples[ x ][ y ] = predSamplesL1[ x ][ y ]   
        (‑187)
* Otherwise,
  + 1. predVal0 = !puIcFlagL0 ? predSamplesL0[ x ][ y ] :   
       ( ( predSamplesL0[ x ][ y ] \* icWeightL0 ) >> 5  ) + ( icOffsetL0 << shift1 ) ) (‑188)
    2. predVal1 = predSamplesL1[ x ][ y ] (‑189)
    3. predSamples[ x ][ y ] =   
       Clip3( 0, ( 1 << bitDepth ) − 1, ( predVal0 + predVal1 + offset2 ) >> shift2 ) (‑190)

**Derivation process for illumination compensation mode availability and parameters**

numerDiv= (( sumProdRefCur + sumRefSquare>>8) << avgShift ) – (sumRef \* sumCur)>> precShift (‑218)

* 1. denomDiv= (( sumRefSquare + sumRefSquare>>8) << avgShift ) – (sumRef \* sumRef)>> precShift (‑219)

~~The variable icWeight specifying a weight for illumination compensation with 7 bit precision is derived as specified in the following:~~

* + ~~If invPsIcWeight is greater than or equal to −2~~~~6~~ ~~and less than 2~~~~6~~~~, the following applies.~~
    - 1. ~~icWeight = invPsIcWeight (‑228)~~
  + ~~Otherwise, ( invPsIcWeight is less than −2~~~~6~~ ~~or greater than or equal to 2~~~~6~~ ~~), the following applies.~~
    - 1. ~~decIcShift = Max( 0, Floor(Log2( Abs( icWeight ) ) − 5 ) ) (‑229)~~
      2. ~~[Ed (GT): In software a function counting leading zero ones is utilized to derive decIcShift. Does this match with draft text?]~~
      3. ~~icWeight = invPsIcWeight >> decIcShift (‑230)~~
      4. ~~icShift −= decIcShift (‑231)~~

# Experimental results

Table 2 Regression improvement only (Test 1)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 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 | 0.00% | -1.2% | -1.3% | -0.5% | -0.5% | -0.5% | 101.0% | 98.7% | 99.4% |
| Kendo | 0.00% | -0.9% | -1.3% | -0.5% | -0.4% | -0.3% | 100.0% | 100.6% | 100.0% |
| Newspaper\_CC | 0.00% | -0.4% | -0.7% | -0.2% | -0.2% | -0.2% | 99.0% | 101.2% | 101.2% |
| GT\_Fly | 0.00% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 99.5% | 101.1% | 100.0% |
| Poznan\_Hall2 | 0.00% | -0.8% | -1.6% | -0.6% | -0.5% | -0.2% | 99.5% | 100.3% | 99.9% |
| Poznan\_Street | 0.00% | 0.1% | -0.6% | -0.1% | -0.1% | -0.1% | 100.0% | 98.7% | 100.0% |
| Undo\_Dancer | 0.00% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 99.1% | 100.1% | 100.0% |
| 1024x768 | 0.00% | -0.83% | -1.11% | -0.41% | -0.35% | -0.30% | 100.0% | 100.2% | 100.2% |
| 1920x1088 | 0.00% | -0.19% | -0.55% | -0.17% | -0.16% | -0.05% | 99.5% | 100.1% | 100.0% |
| **average** | **0.0%** | **-0.46%** | **-0.79%** | **-0.27%** | **-0.24%** | **-0.16%** | **99.7%** | **100.1%** | **100.1%** |

Table 3 Simplification only (Test 2)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 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 | 0.00% | -0.05% | -0.05% | -0.03% | -0.01% | -0.07% | 99.1% | 100.2% | 99.7% |
| Kendo | 0.00% | 0.11% | -0.03% | 0.02% | 0.04% | 0.03% | 97.0% | 99.6% | 98.4% |
| Newspaper\_CC | 0.00% | -0.18% | -0.06% | -0.04% | -0.02% | -0.14% | 98.1% | 102.5% | 101.2% |
| GT\_Fly | 0.00% | 0.00% | 0.02% | 0.00% | 0.00% | 0.00% | 99.4% | 100.8% | 99.6% |
| Poznan\_Hall2 | 0.00% | -0.20% | -0.53% | -0.15% | -0.13% | -0.04% | 98.1% | 100.8% | 100.3% |
| Poznan\_Street | 0.00% | 0.14% | 0.11% | 0.03% | 0.02% | 0.05% | 99.5% | 99.4% | 100.0% |
| Undo\_Dancer | 0.00% | 0.02% | 0.02% | 0.01% | 0.01% | -0.01% | 98.1% | 98.8% | 99.4% |
| 1024x768 | 0.00% | -0.04% | -0.05% | -0.01% | 0.00% | -0.06% | 98.1% | 100.8% | 99.8% |
| 1920x1088 | 0.00% | -0.01% | -0.09% | -0.03% | -0.03% | 0.00% | 98.8% | 100.0% | 99.8% |
| **average** | **0.0%** | **-0.02%** | **-0.07%** | **-0.02%** | **-0.01%** | **-0.02%** | **98.5%** | **100.3%** | **99.8%** |

Table 4 Regression improvement + Simplification (Test 1 + 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% | -1.3% | -1.2% | -0.5% | -0.4% | -0.4% | 99.1% | 99.7% | 99.4% |
| Kendo | 0.0% | -1.0% | -1.4% | -0.5% | -0.4% | -0.3% | 97.5% | 100.0% | 100.7% |
| Newspaper\_CC | 0.0% | -0.3% | -0.8% | -0.2% | -0.2% | -0.2% | 98.3% | 100.3% | 100.6% |
| GT\_Fly | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 99.9% | 100.4% | 99.8% |
| Poznan\_Hall2 | 0.0% | -1.2% | -2.0% | -0.7% | -0.6% | -0.4% | 99.1% | 98.6% | 99.9% |
| Poznan\_Street | 0.0% | 0.2% | -0.7% | -0.1% | -0.1% | -0.1% | 100.1% | 98.8% | 100.1% |
| Undo\_Dancer | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 99.0% | 100.5% | 100.2% |
| 1024x768 | 0.00% | -0.86% | -1.12% | -0.39% | -0.33% | -0.29% | 98.3% | 100.0% | 100.2% |
| 1920x1088 | 0.00% | -0.26% | -0.67% | -0.19% | -0.17% | -0.10% | 99.5% | 99.6% | 100.0% |
| **average** | **0.0%** | **-0.51%** | **-0.86%** | **-0.28%** | **-0.24%** | **-0.18%** | **99.0%** | **99.8%** | **100.1%** |

# Additional results

The proposed simplification consists of several parts: Constant shift operation, 2Nx2N and L0 only. Additional results show the simulation results of each change as Table 5 to Table 7.

Table 5 Constant shift operation only (Breakdown of Test 2)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 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% | 0.0% | -0.1% | 0.0% | 0.0% | -0.1% | 100.2% | 99.0% | 99.2% |
| Kendo | 0.0% | 0.0% | 0.1% | 0.0% | 0.0% | 0.0% | 99.6% | 100.4% | 99.9% |
| Newspaper\_CC | 0.0% | -0.1% | 0.1% | 0.0% | 0.0% | 0.0% | 99.0% | 100.4% | 100.4% |
| GT\_Fly | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 99.9% | 100.5% | 100.1% |
| Poznan\_Hall2 | 0.0% | -0.4% | 0.2% | -0.1% | -0.1% | 0.0% | 99.5% | 98.7% | 99.5% |
| Poznan\_Street | 0.0% | 0.2% | -0.1% | 0.0% | 0.0% | 0.0% | 100.1% | 100.6% | 100.2% |
| Undo\_Dancer | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 98.8% | 99.7% | 100.2% |
| 1024x768 | 0.00% | -0.03% | 0.05% | 0.00% | 0.00% | -0.04% | 99.6% | 99.9% | 99.8% |
| 1920x1088 | 0.00% | -0.08% | 0.04% | -0.01% | -0.02% | 0.00% | 99.6% | 99.9% | 100.0% |
| **average** | **0.0%** | **-0.06%** | **0.04%** | **-0.01%** | **-0.01%** | **-0.01%** | **99.6%** | **99.9%** | **99.9%** |

Table 6 2Nx2N simplification only (Breakdown of Test 2)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 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% | -0.2% | 0.0% | -0.1% | -0.1% | 0.0% | 98.5% | 98.6% | 99.3% |
| Kendo | 0.0% | 0.0% | 0.2% | 0.0% | 0.0% | 0.1% | 96.7% | 100.5% | 99.9% |
| Newspaper\_CC | 0.0% | -0.1% | 0.1% | 0.0% | 0.0% | 0.0% | 98.0% | 99.7% | 100.3% |
| GT\_Fly | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 99.7% | 100.7% | 100.1% |
| Poznan\_Hall2 | 0.0% | -0.1% | -0.1% | 0.0% | -0.1% | 0.0% | 98.6% | 99.7% | 99.8% |
| Poznan\_Street | 0.0% | 0.1% | 0.1% | 0.0% | 0.0% | 0.0% | 99.3% | 100.3% | 100.2% |
| Undo\_Dancer | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 98.6% | 99.4% | 100.0% |
| 1024x768 | 0.00% | -0.07% | 0.07% | 0.00% | 0.01% | 0.00% | 97.7% | 99.6% | 99.8% |
| 1920x1088 | 0.00% | 0.00% | 0.01% | -0.01% | -0.01% | 0.01% | 99.1% | 100.0% | 100.0% |
| **average** | **0.0%** | **-0.03%** | **0.04%** | **0.00%** | **0.00%** | **0.00%** | **98.5%** | **99.8%** | **99.9%** |

Table 7 L0 simplification only (Breakdown of Test 2)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 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% | -0.2% | 0.0% | 0.0% | 0.0% | 0.0% | 99.7% | 99.7% | 99.2% |
| Kendo | 0.0% | 0.2% | 0.2% | 0.1% | 0.0% | 0.0% | 99.7% | 100.5% | 99.5% |
| Newspaper\_CC | 0.0% | 0.0% | 0.1% | 0.0% | 0.0% | 0.0% | 99.1% | 100.4% | 100.6% |
| GT\_Fly | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 99.5% | 100.2% | 99.7% |
| Poznan\_Hall2 | 0.0% | -0.1% | -0.3% | -0.1% | -0.1% | 0.0% | 99.2% | 98.6% | 98.6% |
| Poznan\_Street | 0.0% | 0.0% | 0.2% | 0.0% | 0.0% | 0.0% | 99.9% | 96.8% | 98.1% |
| Undo\_Dancer | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 99.0% | 100.1% | 99.9% |
| 1024x768 | 0.00% | 0.00% | 0.09% | 0.02% | 0.02% | 0.01% | 99.5% | 100.2% | 99.8% |
| 1920x1088 | 0.00% | -0.02% | -0.05% | -0.02% | -0.02% | 0.02% | 99.4% | 98.9% | 99.1% |
| **average** | **0.0%** | **-0.01%** | **0.01%** | **0.00%** | **0.00%** | **0.01%** | **99.4%** | **99.5%** | **99.4%** |

# Conclusion

In this contribution, the parameter cost, which is optimized to IC, is added. Additionally based on a finding that the derivation is only valid in 2Nx2N case, we restrict illumination compensation to 2Nx2N. The change is also beneficial to reduce complexity. The experimental results show the proposed regression improvement with simplification shows 0.2% gain in total video. It is recommended to adopt this method in the next HTM.

# References

[1] T. Ikai “CE5-related: Removal of parsing dependency in illumination compensation”, JCT3V-Dxxxx, JCT3V 4th Meeting: Incheon, KR, 20–26 Apr. 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

[3] T. Ikai, “3D-CE5.h related: Removal of parsing dependency in illumination compensation, JCT3V-D0060, JCT3V 4th Meeting: Incheon, KR, 20–26 Apr. 2013

[4] T. Ikai, “3D-CE5.h related: Combination of JCT3V-D0060 and JCT3V-D0061,” JCT3V-D0062, 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).**