|  |  |
| --- | --- |
| **Joint Collaborative Team on 3D Video Coding Extensions**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  10th Meeting: Strasbourg, FR, 18–24 Oct. 2014 | Document: JCT3V-J0024\_r1 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Simplification of merge candidates list construction** | | |
| *Status:* | Input Document | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Tomohiro Ikai  Takeshi Tsukuba  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 contribution proposes to simplify the 3D-HEVC merge mode by

- Change Shift candidates position

- Remove B1 pruning and some A1 pruning

- A1 VSP inheritance only (Remove non A1 VSP inheritance)

Revision1 fixed the attached spreadsheet naming, which had been mixed up in the first upload.

# Introduction

The 3D-HEVC merge candidates list reuses the HEVC version 1 candidates and can be constructed without the whole HEVC version 1 candidate construction. However, in the constructing of the combined bi-prediction candidates (Combined-bi), HEVC version 1 candidates, especially spatial plus collocate candidates, may be virtually required. The basic solution can be constructing HEVC version 1 like candidates list (combined-bi input list) for Combined-Bi, by reusing constructing 3D-HEVC candidates list. But it is also not so easy because the HEVC candidate is complicatedly interleaved\*1 in the 3D HEVC candidates list and A1 and B1 candidates can be pruned\*2.

As shown table 1, the current 3D-HEVC merge candidate list consists of seven parts, from Part1 to Part7.

The combined-bi input list needs A1, B1, B0, A0, B2, Col. If we construct the input list reusing 3D-HEVC merge list, we cannot reuse Part1 and we do need to concatenate A1 + B1 + B0 + Part4 + Part6 and availability and position.of A1, B1, B0 as well as Part4 and Part6 is required.



**Figure 1** the current 3D-HEVC merge construction and its reuse for combined-bi candidates

**Table 1** Summary of HEVC 1 and the current 3D-HEVC merge construction

|  |  |  |  |
| --- | --- | --- | --- |
|  | HEVC version1 | 3D-HEVC Texture | 3D-HEVC Depth |
| Part1 |  |  | T |
|  |  | D |
|  | IvMC | IvMC |
| Part2 | A1 | A1 (!=IvMC) | A1 (!=T) |
| B1 | B1 (!=IvMC) | B1 (!=T) |
| B0 | B0 | B0 |
| Part3 |  | IvDC (!= B1) |  |
|  | VSP |  |
| Part4 | A0 | A0 | A0 |
| B2 | B2 | B2 |
| Part5 |  | IvMCShift / IvDCShift |  |
| Part6 | Col | Col | Col |
| Part7 | Combined-Bi | Combined-Bi | Combined-Bi |
| Zero | Zero | Zero |

# Proposed Method

We propose to

1) change the position of shift candidate to after VSP candidates

2) remove B1 pruning and modify A1 pruning.

Additionally, we propose to remove VSP inheritance except for A1, which has been proposed in JCT3V-I0106 [1]. By doing this, without A1, no special pruning or access is not required.



**Figure 1** the proposed 3D-HEVC merge construction and its reuse for combined-bi candidates

**Table 2** Sumary of HEVC 1 and the proposed 3D-HEVC merge construction

|  |  |  |  |
| --- | --- | --- | --- |
|  | HEVC version1 | 3D-HEVC Texture  (proposed) | 3D-HEVC Depth  (proposed) |
| Part1 |  |  | T |
|  |  | D |
|  | IvMC | IvMC |
| Part2 | A1 | A1 (!=IvMC) | A1 (!=IvMC or T) |
| B1 | B1 | B1 |
| B0 | B0 | B0 |
| Part3 |  | IvDC |  |
|  | VSP |  |
|  | IvMCShift / IvDCShift |  |
| Part4 | A0 | A0 | A0 |
| B2 | B2 | B2 |
| Col | Col | Col |
| Part5 | Combined-Bi | Combined-Bi | Combined-Bi |
| Zero | Zero | Zero |

# WD change (relative to JCT3V-I1001\_v3)

* 1. The merging candidate list, extMergeCandList, is constructed as follows:

i = 0  
 if( availableFlagT )  
 extMergeCandList[ i++ ] = T  
 if( availableFlagD )  
 extMergeCandList[ i++ ] = D  
 if( availableFlagIvMC && ( !availableFlagT  | |  differentMotion( T, IvMC ) ) )  
 extMergeCandList[ i++ ] = IvMC  
 ~~N = DepthFlag ? T : IvMC~~  
 N = !availableFlagIvMC ? T : IvMC  
 if( availableFlagA1 && ( !availableFlagN  | |  differentMotion( N, A1 ) ) )  
 extMergeCandList[ i++ ] = A1 if( availableFlagB1 ~~&& ( !availableFlagN  | |  differentMotion( N, B~~~~1~~~~) )~~ )  
 extMergeCandList[ i++ ] = B1 if( availableFlagB0 )  
 extMergeCandList[ i++ ] = B0 (I‑104) if( availableFlagIvDC && ( !availableFlagA1  | |  differentMotion( A1, IvDC ) ) ~~&&   
 ( !availableFlagB~~~~1~~~~| |  differentMotion( B~~~~1~~~~, IvDC ) )~~ && ( i < ( 5 + NumExtraMergeCand ) ) )  
 extMergeCandList[ i++ ] = IvDC  
 if( availableFlagVSP && !ic\_flag && iv\_res\_pred\_weight\_idx = = 0 &&   
 i < ( 5 + NumExtraMergeCand ) )  
 extMergeCandList[ i++ ] = VSP  
if( availableFlagIvMCShift && i < ( 5 + NumExtraMergeCand ) &&  
 ( !availableFlagIvMC  | |  differentMotion( IvMC, IvMCShift ) ) )  
 extMergeCandList[ i++ ] = IvMCShift  
 if( availableFlagIvDCShift && i < ( 5 + NumExtraMergeCand ) )  
 extMergeCandList[ i++ ] = IvDCShift  
 if( availableFlagA0 && i < ( 5 + NumExtraMergeCand ) )  
 extMergeCandList[ i++ ] = A0 if( availableFlagB2 && i < ( 5 + NumExtraMergeCand ) )  
 extMergeCandList[ i++ ] = B2~~if( availableFlagIvMCShift && i < ( 5 + NumExtraMergeCand ) &&  
 ( !availableFlagIvMC  | |  differentMotion( IvMC, IvMCShift ) ) )  
 extMergeCandList[ i++ ] = IvMCShift  
 if( availableFlagIvDCShift && i < ( 5 + NumExtraMergeCand ) )  
 extMergeCandList[ i++ ] = IvDCShift~~  
 j = 0  
 while( i < MaxNumMergeCand ) {(I‑105)  
 N = baseMergeCandList[ j++ ]  
 if( N != A1 && N != B1 && N != B0 && N != A0 && N != B2 )  
 extMergeCandList[ i++ ] = N  
 }

* 1. The variable N is derived as specified in the following:
     + If ( nOrigPbW + nOrigPbH ) is equal to 12, the following applies:
       - 1. N = baseMergeCandList[ merge\_idx[ xOrigP ][ yOrigP ] ](I‑106)
     + Otherwise, ( ( nOrigPbW + nOrigPbH ) is not equal to 12 ), the following applies:
       - 1. N = extMergeCandList[ merge\_idx[ xOrigP ][ yOrigP ] ](I‑107)
  2. ~~The derivation process for a view synthesis prediction flag as specified in subclause  is invoked with the luma location ( xCb, yCb ), the luma location ( xPb, yPb ), the variables nPbW and nPbH, the merge candidate indicator N as the inputs, and the output is the mergeCandIsVspFlag.~~
  3. The variable vspModeFlag is derived as specified in the following:
     + 1. vspModeFlag = ((N = = VSP) || ( N = = A1 && VspModeFlag[ xPb − 1 ][ yPb + nPbH − 1 ]) ) ) ~~mergeCandIsVspFlag~~  && !ic\_flag &&   
           ( iv\_res\_pred\_weight\_idx = = 0 ) && availableFlagVSP (I‑108)

~~I.8.5.3.2.18 Derivation process for a view synthesis prediction flag~~

~~Inputs to this process are:~~

* ~~a luma location ( xCb, yCb ) of the top-left sample of the current luma coding block relative to the top-left luma sample of the current picture,~~
* ~~a luma location ( xPb, yPb ) of the top-left sample of the current luma prediction block relative to the top-left luma sample of the current picture,~~
* ~~two variables nPbW and nPbH specifying the width and the height of the current prediction block,~~
* ~~a merge candidate indicator N, specifying the merge candidate.~~

~~Outputs of this process are:~~

* ~~a variable mergeCandIsVspFlag specifying, whether the merge candidate is a view synthesis prediction merge candidate.~~

1. ~~The variable mergeCandIsVspFlag is derived as specified in the following:~~

* ~~If N is equal to VSP, mergeCandIsVspFlag is set equal to 1,~~
* ~~Otherwise, if N is equal to A~~~~1~~~~, B~~~~1~~~~,~~~~B~~~~0~~~~,~~~~A~~~~0~~~~,~~~~or B~~~~2~~~~,~~~~the following applies:~~ 
  + ~~The luma position ( xN, yN ) is specified in depending on N.~~
  + ~~If one of the following conditions is true, the variable mergeCandIsVspFlag is set equal to VspModeFlag[ xN ][ yN ].~~
    - ~~N is equal to A~~~~1~~ ~~or A~~~~0~~
    - ~~N is equal to B~~~~0~~~~,~~~~B~~~~1~~~~,~~~~or B~~~~2~~ ~~and ( yN  >>  Log2CtbSizeY ) is equal to ( yCb >> Log2CtbSizeY )~~
  + ~~Otherwise, ,mergeCandIsVspFlag is set equal to 0.~~
* ~~Otherwise, mergeCandIsVspFlag is set equal to 0.~~

~~Table ‑9 – Specification of xN and yN depending on N~~

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **~~N~~** | ~~A~~~~1~~ | ~~B~~~~1~~ | ~~B~~~~0~~ | ~~A~~~~0~~ | ~~B~~~~2~~ |
| **~~xN~~** | ~~xPb − 1~~ | ~~xPb + nPbW − 1~~ | ~~xPb + nPbW~~ | ~~xPb − 1~~ | ~~xPb − 1~~ |
| **~~yN~~** | ~~yPb + nPbH − 1~~ | ~~yPb − 1~~ | ~~yPb − 1~~ | ~~yPb + nPbH~~ | ~~yPb − 1~~ |

# Experimental results

TestAll: Whole package (1)+(2)+(3)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 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% | 0.00% | -0.01% | -0.02% | 100.3% | 102.1% | 100.8% |
| Kendo | 0.0% | 0.1% | 0.1% | 0.07% | 0.06% | 0.08% | 100.0% | 99.0% | 100.1% |
| Newspaper\_CC | 0.0% | -0.1% | 0.1% | 0.00% | -0.04% | -0.05% | 100.7% | 100.5% | 100.7% |
| GT\_Fly | 0.0% | 0.0% | 0.0% | -0.01% | -0.03% | -0.05% | 100.0% | 97.3% | 98.2% |
| Poznan\_Hall2 | 0.0% | 0.3% | 0.1% | 0.10% | 0.11% | 0.08% | 100.2% | 97.7% | 98.8% |
| Poznan\_Street | 0.0% | 0.1% | 0.1% | 0.03% | 0.02% | 0.00% | 99.8% | 97.4% | 96.8% |
| Undo\_Dancer | 0.0% | 0.0% | 0.1% | 0.00% | -0.01% | -0.02% | 99.9% | 97.1% | 97.4% |
| Shark | 0.0% | 0.2% | -0.1% | 0.00% | 0.01% | 0.01% | 97.8% | 98.4% | 98.2% |
| 1024x768 | 0.0% | 0.0% | 0.0% | 0.02% | 0.01% | 0.01% | 100.3% | 100.6% | 100.6% |
| 1920x1088 | 0.0% | 0.1% | 0.0% | 0.02% | 0.02% | 0.01% | 99.5% | 97.6% | 97.9% |
| **average** | **0.0%** | **0.1%** | **0.0%** | **0.02%** | **0.01%** | **0.01%** | **99.8%** | **98.7%** | **98.9%** |

Additional individual test

Test1: Change shift candidate position

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 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% | -0.2% | -0.02% | -0.02% | -0.04% | 99.9% | 99.5% | 99.9% |
| Kendo | 0.0% | -0.1% | 0.0% | -0.01% | -0.01% | -0.05% | 100.4% | 98.5% | 101.6% |
| Newspaper\_CC | 0.0% | 0.0% | 0.0% | -0.02% | -0.04% | -0.01% | 100.7% | 100.4% | 99.6% |
| GT\_Fly | 0.0% | -0.1% | 0.0% | -0.01% | -0.01% | -0.03% | 100.8% | 99.7% | 99.8% |
| Poznan\_Hall2 | 0.0% | 0.3% | 0.2% | 0.06% | 0.05% | 0.20% | 99.9% | 99.1% | 100.7% |
| Poznan\_Street | 0.0% | 0.0% | 0.0% | 0.01% | 0.01% | 0.00% | 100.0% | 100.8% | 100.4% |
| Undo\_Dancer | 0.0% | -0.2% | -0.1% | -0.03% | -0.01% | -0.07% | 100.6% | 102.0% | 99.6% |
| Shark | 0.0% | 0.0% | 0.0% | 0.00% | 0.01% | -0.01% | 100.8% | 100.3% | 99.9% |
| 1024x768 | 0.0% | 0.0% | -0.1% | -0.02% | -0.02% | -0.03% | 100.3% | 99.5% | 100.3% |
| 1920x1088 | 0.0% | 0.0% | 0.0% | 0.01% | 0.01% | 0.02% | 100.4% | 100.4% | 100.1% |
| **average** | **0.0%** | **0.0%** | **0.0%** | **0.00%** | **0.00%** | **0.00%** | **100.4%** | **100.0%** | **100.2%** |

Test2: removing B2 pruning and modify A1 pruning

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 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% | 0.2% | 0.05% | 0.03% | 0.06% | 100.7% | 103.8% | 100.0% |
| Kendo | 0.0% | 0.0% | 0.2% | 0.06% | 0.04% | 0.06% | 100.5% | 99.4% | 102.1% |
| Newspaper\_CC | 0.0% | -0.1% | 0.2% | 0.03% | -0.02% | -0.02% | 101.2% | 100.3% | 100.7% |
| GT\_Fly | 0.0% | 0.0% | 0.0% | 0.00% | -0.01% | -0.03% | 100.3% | 99.5% | 99.4% |
| Poznan\_Hall2 | 0.0% | 0.3% | 0.2% | 0.10% | 0.09% | 0.12% | 100.2% | 99.6% | 100.3% |
| Poznan\_Street | 0.0% | 0.0% | 0.0% | 0.01% | 0.01% | -0.03% | 100.1% | 102.5% | 100.0% |
| Undo\_Dancer | 0.0% | 0.0% | -0.1% | 0.01% | 0.01% | -0.04% | 100.3% | 100.3% | 100.3% |
| Shark | 0.0% | 0.1% | 0.1% | 0.03% | 0.02% | 0.02% | 99.8% | 98.8% | 98.9% |
| 1024x768 | 0.0% | 0.0% | 0.2% | 0.04% | 0.02% | 0.03% | 100.8% | 101.2% | 100.9% |
| 1920x1088 | 0.0% | 0.1% | 0.0% | 0.03% | 0.02% | 0.01% | 100.1% | 100.1% | 99.8% |
| **average** | **0.0%** | **0.0%** | **0.1%** | **0.04%** | **0.02%** | **0.02%** | **100.4%** | **100.5%** | **100.2%** |

Test3 :removing VSP inheritance without A1

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 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% | -0.01% | -0.01% | 0.03% | 99.9% | 100.7% | 99.0% |
| Kendo | 0.0% | -0.1% | 0.0% | -0.02% | -0.04% | -0.07% | 100.1% | 100.8% | 99.7% |
| Newspaper\_CC | 0.0% | -0.1% | 0.1% | -0.01% | -0.03% | -0.14% | 100.2% | 99.9% | 100.8% |
| GT\_Fly | 0.0% | -0.2% | -0.1% | -0.03% | -0.04% | -0.07% | 100.4% | 98.8% | 100.3% |
| Poznan\_Hall2 | 0.0% | 0.1% | 0.0% | 0.02% | 0.01% | 0.00% | 99.8% | 98.2% | 100.8% |
| Poznan\_Street | 0.0% | 0.0% | 0.0% | 0.00% | 0.00% | -0.02% | 100.2% | 102.7% | 100.3% |
| Undo\_Dancer | 0.0% | 0.0% | 0.0% | -0.01% | -0.01% | 0.03% | 100.1% | 100.0% | 99.8% |
| Shark | 0.0% | 0.0% | -0.2% | -0.03% | -0.02% | -0.02% | 100.2% | 100.2% | 99.8% |
| 1024x768 | 0.0% | -0.1% | 0.0% | -0.01% | -0.02% | -0.06% | 100.1% | 100.4% | 99.8% |
| 1920x1088 | 0.0% | 0.0% | -0.1% | -0.01% | -0.01% | -0.02% | 100.1% | 100.0% | 100.2% |
| **average** | **0.0%** | **0.0%** | **0.0%** | **-0.01%** | **-0.02%** | **-0.03%** | **100.1%** | **100.2%** | **100.1%** |

# Conclusion

It is proposed to simplify the 3D-HEVC merge mode construction. It is recommended to adopt this method in the next 3D-HEVC and HTM.

# 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).**