|  |  |
| --- | --- |
| **Joint Collaborative Team on 3D Video Coding Extension Development**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  3rd Meeting: Geneva, CH, 17–23 Jan. 2013 | Document: JCT3V-C0094 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **3D-CE7.a Improved Nonlinear Depth Representation** | | |
| *Status:* | Input Document | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Ilsoon Lim  Ho-Cheon Wey  Du-Sik Park 14-1 Nongseo-dong Giheung-gu Yongin-si Gyeonggi-do 446-712 South Korea | Tel: Email: | +82 31 280 6688 ilsoon.lim@samsung.com |
| *Source:* | Samsung Electronics Co., Ltd. | | |

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

# Abstract

In this proposal, we modify the nonlinear depth representation (NDR) [1] adopted in the 3DV-ATM. The current NDR is prune to center disparity fluctuation. Because it checks just one time at the beginning of sequence and then on/off decision is fixed for the whole sequence. In this proposal, we make NDR more resilient to every frame by explicitly signalling frame by frame. The modified nonlinear depth representation was implemented in two sub-tests, respectively. and the first sub-test results shows -0.80% (decoded view) and -0.32% (rendered view) BD-rate in EHP Profile and -0.81% (decoded view) and -0.36% (rendered view) BD-rate in HP Profile when compared with 3DV-ATM ver. 6.0 under the common test conditions [2] without the encoding and decoding time increase. the second sub-test results shows -0.81% (decoded view) and -0.27% (rendered view) BD-rate in EHP Profile and -0.81% (decoded view) and -0.36% (rendered view) BD-rate in HP Profile when compared with 3DV-ATM ver. 6.0 under the common test conditions [2] without the encoding and decoding time increase.

# Introduction

We modify the nonlinear depth representation (NDR) adopted in the 3DV-ATM. Although the present NDR has good coding gain, there is still chance to improve error-resilience and coding gain. In this proposal, we address the improvement of the nonlinear depth representation.

# Algorithm description

In this section, we will overview the present NDR and then explain what we improve. NDR is a kind of the non-uniform quantization so that distant objects are quantized more roughly than the closer one. Depth sample values were defined by the following power-law expressions, similar as in the case of well known gamma correction:



For the input depth map, NDR takes the forward transform (linear to nonlinear) and then transformed depth map is used in 3DV encoder and 3DV decoder. Finally, the coded depth map is inversely transformed in order to represent the nonlinear representation to linear representation. To decide whether NDR enabling or not, the present NDR makes the analysis for the first frame of depth map and such decision is final and fixed to the whole sequence. Thus one-time on/off decision suffer from a kind of scene change that often happens and is not efficient throughout a sequence.



1. The present NDR



1. The improved NDR

Fig. 1. The comparison of the present NDR and improved NDR

In the proposal, we add the **nonlinear\_depth\_representation\_enable\_flag**, which indicates on/off decision every frame. Thus, we decide whether NDR is enabled or not and the on/off decision is explicitly signaled to the decoder every slice. In sub-proposal 1, the on/off signaling was written inside the slice header(SH) but the LUT tables related to the nonlinear representation is written SPS (EHP profile) or SEI (HP profile) as before. In sub-proposal 2, the on/off signaling was written inside the depth parameter set(DPS) but the LUT tables related to the nonlinear representation is written SPS (EHP profile) or SEI (HP profile) as before.

Compared to the present nonlinear depth representation, there are two differences as follows.

* NDR on/off flag (**nonlinear\_depth\_representation\_enable\_flag)** is signaled to the decoder every slice.
* Decision on NDR on/off is based on not only disparity center but also disparity concentration. (encoder issue only)

From this paragraph, we will describe the disparity center and disparity concentration in detail. First, disparity center has already been presented in [1] by Poznan University of Technology and is defined as follows:

where is a histogram of the first frame of depth map in center view. Second, disparity concentration is newly defined as follow:

where base histogram is defined as the histogram where the trivial tail and head are eliminated and the effective bin is defined as the bin with the more than some quantity. The disparity concentration is introduced to measure the disparity density in the histogram.



1. Sparse disparity in histogram



1. Dense disparity in histogram

Fig. 2 The comparison of the sparse and dense disparity.

It is not desirable that the dense disparity in histogram is represented in non-uniform quantization. Because there is quantization error during nonlinear mapping and nonlinear re-maping and this quantization error is more serious in the dense disparity. On the contrary, the sparse disparity can be represented in non-uniform quantization and the quantization error is negligibly small. Therefore, we propose that the disparity concentration should be also considered in deciding NDR on/off.

In the present NDR, if the disparity center is more than 100, then NDR is turned on. In the improved NDR, if disparity center is more than 50 and the disparity concentration is less than 60% , then NDR is turned on.

# Test results

## Sub-proposal 1

In sub-proposal, the **nonlinear\_depth\_representation\_enable\_flag** is placed in the slice header in EHP and the SEI in HP. The objective coding performance of the proposed method is presented compared with 3DV-ATM 6.0 anchor (both EHP profile and HP profile). We followed the common testing condition document [2]. Subjective evaluation is also available on the meeting and constant-bitrate points are satisfied.

Table 1 and Table 2 show BD-rate gain in terms of PSNR of decoded views over texture + depth bit rate and BD-rate gain in terms of PSNR of synthesized views over texture + depth bit rate.

Table 1. Test results for the sub-proposal 1 in 3 view test scenario (EHP profile)

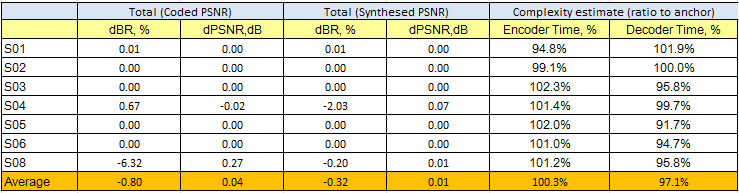
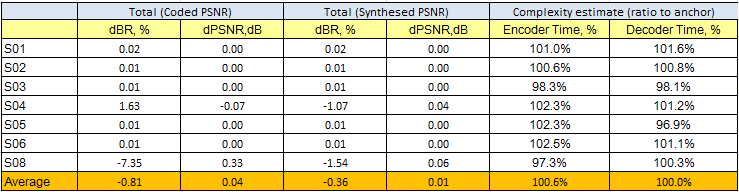


Table 2. Test results for the sub-proposal 1 in 3 view test scenario (HP profile)



# Test results in sub-proposal 2

In sub-proposal, the **nonlinear\_depth\_representation\_enable\_flag** is placed in the depth parameter set and the SEI in HP. The objective coding performance of the proposed method is presented compared with 3DV-ATM 6.0 anchor (both EHP profile and HP profile). We followed the common testing condition document [2]. Subjective evaluation is also available on the meeting and constant-bitrate points are satisfied.

Table 3 and Table 4 show BD-rate gain in terms of PSNR of decoded views over texture + depth bit rate and BD-rate gain in terms of PSNR of synthesized views over texture + depth bit rate.

Table 3. Test results for the sub-proposal 2 in 3 view test scenario (EHP profile)

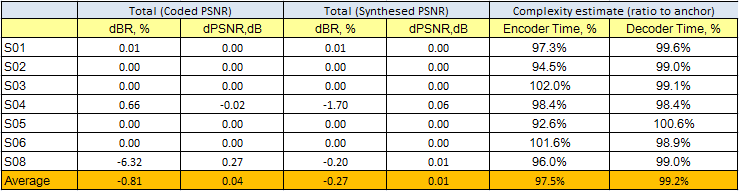
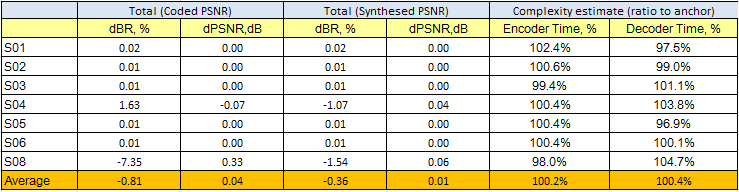


Table 4. Test results for the sub-proposal 2 in 3 view test scenario (HP profile)



# 3D-AVC draft text in sub-proposal 1

*The slice header syntax specified in subclause 7.3.3 is replaced by the following.*

|  |  |  |
| --- | --- | --- |
| slice\_header( ) { | **C** | **Descriptor** |
| **first\_mb\_in\_slice** | 2 | ue(v) |
| **slice\_type** | 2 | ue(v) |
| **pic\_parameter\_set\_id** | 2 | ue(v) |
| if( avc\_3d\_extension\_flag && slice\_header\_prediction\_flag != 0 ) { |  |  |
| **pre\_slice\_header\_src** | 2 | u(2) |
| if( slice\_type = = P | | slice\_type = = SP  | | slice\_type = = B ) { |  |  |
| **pre\_ref\_lists\_src** | 2 | u(2) |
| if ( !pre\_ref\_lists\_src ) { |  |  |
| **num\_ref\_idx\_active\_override\_flag** | 2 | u(1) |
| if( num\_ref\_idx\_active\_override\_flag ) { |  |  |
| **num\_ref\_idx\_l0\_active\_minus1** | 2 | ue(v) |
| if( slice\_type = = B ) |  |  |
| **num\_ref\_idx\_l1\_active\_minus1** | 2 | ue(v) |
| } |  |  |
| if( nal\_unit\_type = = 20 | | ( nal\_unit\_type = = 21  && DepthFlag ) ) |  |  |
| ref\_pic\_list\_mvc\_modification( ) /\* specified in Annex  \*/ | 2 |  |
| else if( nal\_unit\_type = = 21 && DepthFlag = = 0 ) |  |  |
| ref\_pic\_list\_3dvc\_modification( ) | 2 |  |
| } |  |  |
| } |  |  |
| if( ( weighted\_pred\_flag && ( slice\_type = = P | |   slice\_type = = SP ) ) | |  ( weighted\_bipred\_idc = = 1 && slice\_type = = B ) ) { |  |  |
| **pre\_pred\_weight\_table\_src** | 2 | u(2) |
| if( !pre\_pred\_weight\_table\_src ) |  |  |
| pred\_weight\_table( ) | 2 |  |
| if(nal\_ref\_idc != 0 ) { |  |  |
| **pre\_dec\_ref\_pic\_marking\_src** | 2 | u(2) |
| if( !pre\_dec\_ref\_pic\_marking\_src ) |  |  |
| dec\_ref\_pic\_marking( ) | 2 |  |
| } |  |  |
| **slice\_qp\_delta** | 2 | se(v) |
| } else { |  |  |
| if( separate\_colour\_plane\_flag = = 1 ) |  |  |
| **colour\_plane\_id** | 2 | u(2) |
| **frame\_num** | 2 | u(v) |
| if( !frame\_mbs\_only\_flag ) { |  |  |
| **field\_pic\_flag** | 2 | u(1) |
| if( field\_pic\_flag ) |  |  |
| **bottom\_field\_flag** | 2 | u(1) |
| } |  |  |
| if( IdrPicFlag ) |  |  |
| **idr\_pic\_id** | 2 | ue(v) |
| if( pic\_order\_cnt\_type = = 0 ) { |  |  |
| **pic\_order\_cnt\_lsb** | 2 | u(v) |
| if( bottom\_field\_pic\_order\_in\_frame\_present\_flag && !field\_pic\_flag ) |  |  |
| **delta\_pic\_order\_cnt\_bottom** | 2 | se(v) |
| } |  |  |
| if( pic\_order\_cnt\_type = = 1 && !delta\_pic\_order\_always\_zero\_flag ) { |  |  |
| **delta\_pic\_order\_cnt[** 0 **]** | 2 | se(v) |
| if( bottom\_field\_pic\_order\_in\_frame\_present\_flag && !field\_pic\_flag ) |  |  |
| **delta\_pic\_order\_cnt[** 1 **]** | 2 | se(v) |
| } |  |  |
| if( redundant\_pic\_cnt\_present\_flag ) |  |  |
| **redundant\_pic\_cnt** | 2 | ue(v) |
| if( slice\_type = = B ) |  |  |
| **direct\_spatial\_mv\_pred\_flag** | 2 | u(1) |
| if( slice\_type = = P | | slice\_type = = SP | | slice\_type = = B ) { |  |  |
| **num\_ref\_idx\_active\_override\_flag** | 2 | u(1) |
| if( num\_ref\_idx\_active\_override\_flag ) { |  |  |
| **num\_ref\_idx\_l0\_active\_minus1** | 2 | ue(v) |
| if( slice\_type = = B ) |  |  |
| **num\_ref\_idx\_l1\_active\_minus1** | 2 | ue(v) |
| } |  |  |
| } |  |  |
| if( nal\_unit\_type = = 20 | | ( nal\_unit\_type = = 21  && DepthFlag ) ) |  |  |
| ref\_pic\_list\_mvc\_modification( ) /\* specified in Annex  \*/ | 2 |  |
| else if( nal\_unit\_type = = 21 && DepthFlag = = 0 ) |  |  |
| ref\_pic\_list\_3dvc\_modification( ) |  |  |
| else |  |  |
| ref\_pic\_list\_modification( ) | 2 |  |
| if( ( weighted\_pred\_flag && ( slice\_type = = P | | slice\_type = = SP ) ) | |  ( weighted\_bipred\_idc = = 1 && slice\_type = = B ) ) |  |  |
| pred\_weight\_table( ) | 2 |  |
| if( nal\_ref\_idc != 0 ) |  |  |
| dec\_ref\_pic\_marking( ) | 2 |  |
| if( entropy\_coding\_mode\_flag && slice\_type != I && slice\_type != SI ) |  |  |
| **cabac\_init\_idc** | 2 | ue(v) |
| **slice\_qp\_delta** | 2 | se(v) |
| if( slice\_type = = SP | | slice\_type = = SI ) { |  |  |
| if( slice\_type = = SP ) |  |  |
| **sp\_for\_switch\_flag** | 2 | u(1) |
| **slice\_qs\_delta** | 2 | se(v) |
| } |  |  |
| if( deblocking\_filter\_control\_present\_flag ) { |  |  |
| **disable\_deblocking\_filter\_idc** | 2 | ue(v) |
| if( disable\_deblocking\_filter\_idc != 1 ) { |  |  |
| **slice\_alpha\_c0\_offset\_div2** | 2 | se(v) |
| **slice\_beta\_offset\_div2** | 2 | se(v) |
| } |  |  |
| } |  |  |
| if( num\_slice\_groups\_minus1 > 0 &&  slice\_group\_map\_type >= 3 && slice\_group\_map\_type <= 5) |  |  |
| **slice\_group\_change\_cycle** | 2 | u(v) |
| if ( nal\_unit\_type  = =  21  && ( slice\_type != I && slice\_type != SI )) { |  |  |
| if( DepthFlag ) |  |  |
| **depth\_weighted\_pred\_flag** | 2 | u(1) |
| else if( avc\_3d\_extension\_flag &&   ViewCompOrder( 0, view\_idx ) > ViewCompOrder( 1, view\_idx ) ) |  |  |
| **dmvp\_flag** | 2 | u(1) |
| if ( 3dv\_acquisition\_idc != 1 &&   (depth\_weighted\_pred\_flag = = 1 || dmvp\_flag = = 1) ) |  |  |
| **dps\_id** | 2 | ue(v) |
| } |  |  |
| } |  |  |
| if( DepthFlag && profile\_idc = = 139 ) { |  |  |
| **nonlinear\_depth\_representation\_enable \_flag** | 2 | u(1) |
| } |  |  |
| } |  |  |

*In 7.4.4, Slice data semantics, add the following text*

**nonlinear\_depth\_representation\_enable\_flag** equalto 0 specifies that the current slice is not represented in nonlinear form. Nonlinear\_depth\_representation\_flag equal to 1 specifies that the current slice is represented in nonlinear form.

# 3D-AVC draft text in sub-proposal 2

*The syntax table is specified in subclause I.7.3.2.13 is replaced by the following.*

|  |  |  |
| --- | --- | --- |
| depth\_parameter\_set\_rbsp( ) { | C | Descriptor |
| **depth\_parameter\_set\_id** | 11 | ue(v) |
| **pred\_direction** | 11 | ue(v) |
| if( pred\_direction  = =  0  | |  pred\_direction  = =  1 ) { |  |  |
| **ref\_dps\_id0** | 11 | ue(v) |
| predWeight0 = 64 |  |  |
| } |  |  |
| if( pred\_direction  = =  0 ) { |  |  |
| **ref\_dps\_id1** | 11 | ue(v) |
| **pred\_weight0** | 11 | u(6) |
| predWeight0 = pred\_weight0 |  |  |
| } |  |  |
| **num\_views\_minus1** | 11 | ue(v) |
| depth\_ranges( pred\_direction, depth\_parameter\_set\_id ) |  |  |
| **vsp\_param\_flag** | 11 | u(1) |
| if( vsp\_param\_flag ) |  |  |
| vsp\_param( pred\_direction, depth\_parameter\_set\_id ) |  |  |
| **depth\_param\_additional\_extension\_flag** | 11 | u(1) |
| if(depth\_param\_additional\_extension\_flag = = 1 ) |  |  |
| while( more\_rbsp\_data( ) ) |  |  |
| **depth\_param\_additional\_extension\_data\_flag** | 11 | u(1) |
| If(profile\_idc = = 139 ) { |  |  |
| for( i = 0; i <= num\_views\_minus1; i++ ) { |  |  |
| **nonlinear\_depth\_representation\_enable flag** | 11 | u(1) |
| } |  |  |
| } |  |  |
| rbsp\_trailing\_bits( ) |  |  |
| } |  |  |

*In 7.4.2.13, Depth parameter set RBSP semantics, add the following text*

**nonlinear\_depth\_representation\_enable\_flag** equalto 0 specifies that the current slice is not represented in nonlinear form. Nonlinear\_depth\_representation\_flag equal to 1 specifies that the current slice is represented in nonlinear form.

# Conclusion

In this document, improved NDR was proposed. Experimental results in sub-proposal 1 shows -0.80% (decoded view) and -0.32% (rendered view) BD-rate in EHP Profile and -0.81% (decoded view) and -0.36% (rendered view) BD-rate in HP Profile for 3 view case, compared to anchor and the encoding time and decoding time remains the similar to the anchor.

Experimental results in sub-proposal 2 shows -0.81% (decoded view) and -0.27% (rendered view) BD-rate in EHP Profile and -0.81% (decoded view) and -0.36% (rendered view) BD-rate in HP Profile for 3 view case, compared to anchor and the encoding time and decoding time remains the similar to the anchor.

We recommend that either the sub-proposal 1or the sub-proposal 2 is integrated into 3DV-ATM

# Patent rights declaration(s)

**Samsung Electronics Co., Ltd. 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. ISO/IEC JTC1/SC29/WG11“3D-CE2a results on Nonlinear Depth Representation,” M25017, May 2012.
2. “Common Test Conditions for 3DV experimentation,” JCT2-B1100, JCT-3V 2nd Meeting , Shanghai, CH, 13-19 Oct. 2012.