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-D0080

Title: **Comparison of half resolution depth map coding versus full resolution depth map coding in 3D-ATM**

Status: Input Document

Purpose: Report

Author(s) or Contact(s): Krzysztof Wegner, Olgierd Stankiewicz, Marek Domański
Tel: +48 61 6653893
Email: kwegner@multimedia.edu.pl

Source: Poznan University of Technology

Abstract

In this document, evaluation of impact of coded depth resolution on performance of the current AVC-based 3D video coding technology implemented in 3D-ATM [ATM] is presented. The evaluation has been done with respect to Common Test Conditions [CTC] in HP and EHP profiles. Two depth configurations has been tested - full depth resolution and half depth resolution:

The results show that coding with half resolution of depth (currently used in CTC) provides better results in comparison to coding with full resolution of depth. It is also more computationally efficient. One of observations of the document is that comparing schemes of coding with different depth resolution require common reference in order to obtain meaningful results.

Coding with full resolution of depth requires 64-bit version of 3D-ATM to run correctly in case of Full HD sequences.

1 Introduction

In this document, we present an evaluation of depth resolution impact on performance of the current AVC-based 3D video coding technology implemented in 3D-ATM [ATM] with respect to Common Test Conditions [CTC] in HP and EHP profiles. We have tested two configurations regarding resolution of the depth maps:

- Full res – where coded depths have the same resolution as the coded texture views,
- Half res – where coded depths have been decimated by the factor of 2, resulting in two times less resolution compared to coded texture views (default in CTC).

2 Methodology

We have encoded 3 videos along with 3 correspondent depth maps with 3D-ATM in version 7.0 according to the Common Test Condition. Based on the decoded videos and depth maps we have synthesized 6 view in between positions of input views according to table 1. Synthesized views has been then compared via PSNR with views synthesized at the same spatial positions but basing on the original data.

Table 1. Input view positions and synthesized views positions for 3 view case.

	Test Sequence	Input views positions	Synthesized views positions
S01	Poznan_Hall2	7-6-5	6.750 6.500 6.250 5.750 5.500 5.250
S02	Poznan_Street	5-4-3	4.750 4.500 4.250 3.750 3.500 3.250
S03	Undo_Dancer	1-5-9	2.000 3.000 4.000 6.000 7.000 8.000
S04	GT_Fly	9-5-1	8.000 7.000 6.000 4.000 3.000 2.000
S05	Kendo	1-3-5	1.500 2.000 2.500 3.500 4.000 4.500
S06	Balloons	1-3-5	1.500 2.000 2.500 3.500 4.000 4.500
S08	Newspaper1	2-4-6	2.500 3.000 3.500 4.500 5.000 5.500

Unfortunately, excel sheet, commonly used for reporting of the coding result used in CEs, for depth maps reports PNSRs calculated with respect to the image at the input of the encoder. In case of half resolution depth maps coding, such PNSR is calculated with respect to decimated depth maps, which makes comparison with case of full resolution depth map coding unfair.

Comparison of different resolutions of video and depth maps in coding require a common ground. Therefore, we have also calculated PNSR with respect to the original full resolution depth maps.

Complexity assessment was evaluated as execution time of the encoder and the decoder.

3 Simulation results

The simulations results were generated on a ~80 core cluster system. The cluster platform's processing units have the following specifications:

- Processor: Intel Xeon X5675
- Clock Speed: 3.06 GHz
- Memory: approx. 4 GB per Core
- OS: 64-bit Windows Server 2008
- Compiler: Microsoft Visual Studio 2008 (64 bit)

Overview of the results for EHP and HP profile of the Full resolution depth map coding versus half resolution is shown in Tables 2-5. We have provided two sets of results. First, based on depth PSNR values taken form the encoder log files Table 2 and 4, and the second, where PSNR of the coded depth maps was calculated against original full resolution depth maps Table 3 and 5.

PSNR values calculated against original full resolution depth maps are always lower than those taken form encoder log files:

- In case of half resolution depth maps coding, depth maps PSNR values given in encoder log files don't take into account quality loss caused by depth map decimation.
- In case of full resolution depth maps coding, PSNR value of the depth maps given in the encoder log also don't match those calculated from decoded data. Probably there is a unknown bug in the ATM encoder PSNR reporting routines.
  NOTE: values of PSNR for texture reported in encoder log are in perfect match with those calculated outside encoder. Bug is related only to depth maps PSNR value.

Coding the depth maps with full resolution increase depth bitstream by 150% (table 3) in HP and 176% (table 5) in case of EHP.

On average this depth bitstream increase results in 23% total bitstream increase in HP and 14,5 % total bitstram increase in EHP when quality of synthesized view are evaluated (Total – Synthesized PSNR column of table 2-5).

Full resolution depth map coding increase encoder complexity by 16% (HP) and 13% (EHP). Results of decoder complexity are inconclusive as for some sequences it increase for some decrease resulting on average in decoder comprexity increase by 2% in HP and decrease by 3% in EHP.

All simulation results are attached to this document in excel sheets.

Table 2. FullRes vs HalfRes (depth PSNRs taken from encoder log files) - HP profile

	Texture Coding		Depth Coding		Total (Coded PSNR)		Total (Synthesed PSNR)	
	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB
Poznan Hall 2	0,00	0,00	136,85	-3,83	23,51	-0,78	17,77	-0,60
Poznan Street	0,00	0,00	119,35	-3,01	24,28	-0,69	21,67	-0,64
Dancer	0,00	0,00	149,83	-7,37	12,95	-0,47	1,97	-0,06
GT Fly	0,00	0,00	199,10	-6,15	22,00	-0,82	16,75	-0,59
Kendo	0,00	0,00	103,37	-4,24	42,70	-1,82	38,11	-1,53
Balloons	0,00	0,00	103,90	-3,20	33,14	-1,46	29,67	-1,24
Newspaper CC	0,00	0,00	95,75	-2,68	40,91	-1,44	34,47	-1,07
Average	0,00	0,00	129,74	-4,35	28,50	-1,07	22,92	-0,82

Table 3. FullRes vs HalfRes (depth PSNRs calculated with respect do original full resolution depth maps) - HP profile

	Texture Coding		Depth Coding		Total (Coded PSNR)		Total (Synthesed PSNR)	
	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB
Poznan Hall 2	0,00	0,00	140,13	-2,71	23,51	-0,78	17,77	-0,60
Poznan Street	0,00	0,00	154,82	-1,32	24,28	-0,69	21,67	-0,64
Dancer	0,00	0,00	-100,00	-2,33	12,95	-0,47	1,97	-0,06
GT Fly	0,00	0,00	229,76	-3,17	22,00	-0,82	16,75	-0,59
Kendo	0,00	0,00	128,75	-2,43	42,70	-1,82	38,11	-1,53
Balloons	0,00	0,00	151,48	-1,48	33,14	-1,46	29,67	-1,24
Newspaper CC	0,00	0,00	356,15	-1,11	40,91	-1,44	34,47	-1,07
Average	0,00	0,00	151,59	-2,08	28,50	-1,07	22,92	-0,82

Table 4. FullRes vs HalfRes (depth PSNRs taken from encoder log files) - EHP profile

	Texture Coding		Depth Coding		Total (Coded PSNR)		Total (Synthesed PSNR)	
	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB
Poznan Hall 2	-0,08	0,00	123,62	-4,65	15,59	-0,46	10,60	-0,32

Poznan Street	-0,79	0,02	136,33	-3,90	17,51	-0,48	15,28	-0,44
Dancer	-1,49	0,05	207,66	-10,19	18,58	-0,60	7,47	-0,21
GT Fly	-1,55	0,06	213,38	-6,87	21,22	-0,69	15,49	-0,50
Kendo	-0,50	0,02	86,51	-3,01	24,95	-1,00	20,48	-0,77
Balloons	-0,42	0,02	99,15	-3,12	22,50	-0,96	18,88	-0,76
Newspaper CC	-0,18	0,01	98,42	-3,15	20,42	-0,77	13,75	-0,45
Average	-0,72	0,03	137,87	-4,98	20,11	-0,71	14,56	-0,49

Table 5. FullRes vs HalfRes (depth PSNRs calculated with respect do original full resolution depth maps) - EHP profile

	Texture Coding		Depth Coding		Total (Coded PSNR)		Total (Synthesed PSNR)	
	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB
Poznan Hall 2	-0,08	0,00	125,58	-3,93	15,59	-0,46	10,60	-0,32
Poznan Street	-0,79	0,02	148,58	-2,64	17,51	-0,48	15,28	-0,44
Dancer	-1,49	0,05	337,68	-4,52	18,58	-0,60	7,47	-0,21
GT Fly	-1,55	0,06	252,18	-4,34	21,22	-0,69	15,49	-0,50
Kendo	-0,50	0,02	100,21	-2,31	24,95	-1,00	20,48	-0,77
Balloons	-0,42	0,02	118,71	-2,01	22,50	-0,96	18,88	-0,76
Newspaper CC	-0,18	0,01	151,55	-1,98	20,42	-0,77	13,75	-0,45
Average	-0,72	0,03	176,36	-3,10	20,11	-0,71	14,56	-0,49

Table 6. Complexity assessment of FullRes vs HalfRes HP profile

	Complexity estimate (ratio to anchor)		
	Encoder Time, %	Decoder Time, %	Rendering Time, %
Poznan Hall 2	116,5%	91,9%	100,0%
Poznan Street	107,0%	112,8%	100,0%
Dancer	108,8%	101,8%	100,0%
GT Fly	103,6%	109,1%	100,0%
Kendo	120,9%	99,3%	100,0%
Balloons	125,2%	92,2%	100,0%
Newspaper CC	130,5%	109,1%	100,0%
Average	116,0%	102,3%	100,0%

Table 7. Complexity assessment of FullRes vs HalfRes EHP profile

	Complexity estimate (ratio to anchor)		
	Encoder Time, %	Decoder Time, %	Rendering Time, %
Poznan Hall 2	106,6%	84,3%	100,0%
Poznan Street	111,1%	105,9%	100,0%
Dancer	102,9%	114,7%	100,0%
GT Fly	106,1%	98,2%	100,0%
Kendo	124,6%	94,0%	100,0%
Balloons	127,9%	93,0%	100,0%

Newspaper CC	113,0%	88,7%	100,0%
Average	113,2%	97,0%	100,0%

ff

4 Conclusions

Coding with half resolution of depth, (as currently specified in CTC) provides better results in comparison with full resolution depth map coding. Half resolution depth map coding provide 17,9% total bitstream reduction in HP and 12,6% total bitstream reduction in EHP when quality of the synthesized view are evaluated (Table 8).

Table 8. Summary of the coding results Half res vs Full res

HP		EHP	
Coded	Synthesized	Coded	Synthesized
-21,7 %	-17,9 %	-16,7%	-12,6%

Coding with haft resolution depth maps provide 13,8% encoder complexity reduction in HP and 11,6% in EHP, while decoder complexity in is decreased by xx % in HP and increase by xx% in EHP.

Comparing different coding schemes with various resolution of depth maps require common reference in order to obtain meaningful results.

Coding with full resolution of depth requires 64-bit version of 3D-ATM to run correctly in case of Full HD sequences.

5 Recommendations

Fix bug in reporting full resolution depth maps PSNR values.

6 Patent rights declaration(s)

Poznan University of Technology 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).