

# CE7: Removal of texture-to-depth resolution ratio restrictions



JCT3V-E0035

Payman Aflaki, Miska M. Hannuksela

# Introduction

- Current 3D-AVC supports two texture to depth resolutions
  - 1:1 (texture and depth are coded at equal spatial resolution)
  - 1:0.5 (depth at half spatial resolution in horizontally and vertically)
- In-loop resampling has been removed  
→ no reason for restricting to only these two depth resolutions
- The contribution shows that
  - A lower depth resolution can provide better RD performance for synthesized views
  - Depth maps at very low resolution can be coded as side information to improve the performance of texture coding

# Simulations

ATM v8.1 software with a modification to support different ratios between spatial resolution of texture and depth

## Tested resolutions

- 1/2, 1/4, 1/8, and 1/16 in both directions
- GDV: Single depth value per each depth view component

## Considered use cases

1. Best depth resolution for RD performance of synthesized views (CTC)
  - Results against MVC+D, i.e. HP anchor bitstreams according to the CTC
2. Best depth resolution for multiview texture coding
  - Coded depth maps are used only as side information to enhance the RD performance of texture views
  - Results against MVC, i.e. extracted texture views of HP anchor

# Simulations: tested resolutions

Experiment	Depth/Texture resolution ratio (vertical x horizontal)	How to derive the lower resolution depth map
Anchor	0.5x0.5	JSVM Downsampled
1/2	0.5x0.5	JSVM Downsampled
1/4	0.25x0.25	JSVM Downsampled
1/8	0.125x0.125	Max Downsampled*
1/16	0.0625x0.0625	Max Downsampled
GDV	Only one value	Mean value of frame selected

\* Max value of each block with size NxN is used as the downsampled respective value of that block

# Results, Use Case 1: Best depth resolution for RD performance of synthesized views (CTC), VSO on

	Texture Coding		Depth Coding		Total (Coded PSNR)		Total Vs. anchor w/o depth		Total (Synthesed PSNR)	
	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB
S01	-30.96	1.15	-0.24	0.00	-28.52	1.06	-20.62	0.72	-26.22	0.97
S02	-12.23	0.38	-42.37	2.20	-16.89	0.54	-8.12	0.25	-11.30	0.35
S03	-18.76	0.71	6.55	-0.51	-17.06	0.66	-10.48	0.39	-14.76	0.51
S04	-23.69	0.98	-24.31	1.70	-21.46	0.89	-14.86	0.59	-21.12	0.83
S05	-22.86	1.25	-38.97	2.25	-29.23	1.63	-13.74	0.68	-22.75	1.07
S06	-27.26	1.61	-35.94	2.05	-31.20	1.88	-21.53	1.22	-24.92	1.29
S08	-15.07	0.69	12.93	-0.55	-18.74	0.87	-1.69	0.07	-12.72	0.48
Average	-21.55	0.97	-17.48	1.02	-23.30	1.08	-13.01	0.56	-19.11	0.79

Difference (dBR, %-unit) = proposed - 3D-AVC (EHP with 1/2 resolution depth)

	Texture Coding	Depth Coding	Total (Coded PSNR)	Total Vs. anchor w/o depth	Total (Synthesed PSNR)
S01	0.00	0.00	0.00	0.00	0.00
S02	0.54	-42.34	-5.08	-5.61	-1.14
S03	0.00	0.00	0.00	0.00	0.00
S04	0.00	0.00	0.00	0.00	0.00
S05	0.53	-38.62	-9.44	-11.50	-4.04
S06	0.36	-35.16	-6.16	-7.03	-2.15
S08	0.00	0.00	0.00	0.00	0.00
Average	0.21	-16.59	-2.95	-3.45	-1.05

	Best resolution
S01	1/2
S02	1/4
S03	1/2
S04	1/2
S05	1/4
S06	1/4
S08	1/2

# Results, Use Case 1: Best depth resolution for RD performance of synthesized views (CTC), VSO off

	Texture Coding		Depth Coding		Total (Coded PSNR)		Total Vs. anchor without depth		Total (Synthesed PSNR)	
	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB
S01	-30.96	1.15	-53.75	4.01	-32.27	1.22	-24.78	0.89	-22.95	0.73
S02	-12.49	0.38	-47.44	2.60	-16.13	0.52	-7.28	0.22	-9.66	0.29
S03	-19.14	0.73	-29.43	2.78	-18.91	0.74	-12.48	0.47	-11.22	0.38
S04	-19.24	0.76	-76.59	8.21	-23.47	0.97	-17.03	0.67	-21.03	0.74
S05	-23.17	1.26	-49.40	3.31	-25.37	1.38	-9.03	0.43	-19.34	0.88
S06	-27.51	1.62	-44.51	2.70	-29.16	1.72	-19.19	1.07	-23.49	1.20
S08	-15.09	0.69	-2.50	0.10	-15.73	0.72	1.95	-0.08	-7.80	0.28
Average	-21.09	0.94	-43.37	3.39	-23.00	1.04	-12.55	0.52	-16.50	0.64

	Difference = Proposed - 3D-AVC (EHP with 1/2 resolution depth)				
	Texture Coding	Depth Coding	Total (Coded PSNR)	Total Vs. anchor without depth	Total (Synthesed PSNR)
	dBR, %	dBR, %	dBR, %	dBR, %	dBR, %
S01	0.13	-35.53	-6.93	-7.70	-1.81
S02	0.53	-33.33	-6.83	-7.55	-2.78
S03	0.00	0.00	0.00	0.00	0.00
S04	5.21	-36.49	-3.04	-3.29	-1.41
S05	0.46	-28.08	-15.90	-19.38	-10.43
S06	0.30	-26.31	-10.29	-11.73	-6.50
S08	0.00	0.00	0.00	0.00	0.00
Average	0.95	-22.82	-6.14	-7.09	-3.28

	Best resolution
S01	1/4
S02	1/4
S03	1/2
S04	1/8
S05	1/4
S06	1/4
S08	1/2

# Results, Use Case 2: Best depth resolution for multiview texture coding, VSO on

	Texture Coding		Depth Coding		Total (Coded PSNR)		Total Vs. anchor w/o depth	
	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB
S01	-30.22	1.12	-60.94	7.57	-34.68	1.36	-27.45	1.02
S02	-9.41	0.28	-50.50	2.96	-16.03	0.51	-7.17	0.21
S03	-18.14	0.69	-50.26	5.54	-21.38	0.85	-15.15	0.57
S04	-22.93	0.95	-65.95	6.49	-26.01	1.11	-19.79	0.81
S05	-22.48	1.23	-48.64	2.82	-32.59	1.90	-17.83	0.94
S06	-26.94	1.59	-40.20	2.32	-33.06	2.04	-23.65	1.38
S08	-14.18	0.66	-47.63	28.65	-26.89	1.36	-11.56	0.54
Average	-20.61	0.93	-52.02	8.05	-27.23	1.30	-17.52	0.78

Difference (dBR, %-unit) = proposed - 3D-AVC (EHP with 1/2 resolution depth)

	Texture Coding	Depth Coding	Total (Coded PSNR)	Total Vs. anchor w/o depth
	dBR, %	dBR, %	dBR, %	dBR, %
S01	0.74	-60.70	-6.15	-6.83
S02	3.36	-50.47	-4.22	-4.66
S03	0.61	-56.81	-4.32	-4.67
S04	0.76	-41.64	-4.54	-4.93
S05	0.92	-48.30	-12.80	-15.60
S06	0.68	-39.42	-8.02	-9.15
S08	0.89	-60.56	-8.16	-9.87
Average	1.14	-51.13	-6.89	-7.96

	Best resolution
S01	1/8
S02	1/8
S03	1/4
S04	1/4
S05	1/8
S06	1/8
S08	1/16

# Results, Use Case 2: Best depth resolution for multiview texture coding, VSO off

	Texture Coding		Depth Coding		Total (Coded PSNR)		Total Vs. anchor without depth	
	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB	dBR, %	dPSNR,dB
S01	-30.36	1.13	-63.55	6.01	-34.19	1.33	-26.91	0.99
S02	-12.49	0.38	-47.44	2.60	-16.13	0.52	-7.28	0.22
S03	-18.62	0.71	-60.73	7.46	-21.99	0.87	-15.80	0.60
S04	-23.58	0.98	-71.19	7.29	-25.83	1.10	-19.60	0.80
S05	-20.67	1.14	-63.90	4.82	-31.97	1.89	-17.08	0.92
S06	-25.86	1.53	-53.39	4.28	-32.65	2.03	-23.18	1.36
S08	-14.20	0.66	-48.75	8.19	-26.70	1.35	-11.33	0.53
Average	-20.82	0.93	-58.42	5.81	-27.07	1.30	-17.31	0.77

	Difference = Proposed - 3D-AVC (EHP with 1/2 resolution depth)			
	Texture Coding	Depth Coding	Total (Coded PSNR)	total Vs. anchor without dept
	dBR, %	dBR, %	dBR, %	dBR, %
S01	0.73	-45.34	-8.85	-9.83
S02	0.53	-33.33	-6.83	-7.55
S03	0.52	-31.30	-3.08	-3.33
S04	0.87	-31.10	-5.40	-5.86
S05	2.96	-42.59	-22.50	-27.43
S06	1.95	-35.19	-13.78	-15.72
S08	0.89	-46.24	-10.98	-13.28
Average	1.21	-37.87	-10.20	-11.86

	Best resolution
S01	1/8
S02	1/4
S03	1/4
S04	1/4
S05	1/16
S06	1/16
S08	1/16



# Results, Use Case 2: Global disparity vector (GDV) for multiview texture coding

- Depth view components contain samples with a constant sample value equal to GDV
- Coded conventionally according to CTC (depth resolution 1/2 vertically and horizontally)

	Texture Coding		Total (Coded PSNR)		Total Vs. anchor without depth	
	dB <sub>R</sub> , %	dPSNR,dB	dB <sub>R</sub> , %	dPSNR,dB	dB <sub>R</sub> , %	dPSNR,dB
S01	-26.70	0.99	-30.84	1.23	-23.19	0.87
S02	-7.08	0.21	-13.77	0.45	-4.68	0.14
S03	-13.25	0.50	-18.72	0.74	-12.26	0.46
S04	-14.12	0.55	-18.94	0.77	-12.11	0.47
S05	-17.35	0.95	-30.06	1.80	-14.80	0.82
S06	-20.63	1.19	-28.43	1.75	-18.39	1.07
S08	-11.73	0.54	-24.71	1.24	-8.93	0.41
Average	-15.84	0.70	-23.64	1.14	-13.48	0.61

# Specification text changes, background

In current 3D-AVC depth view components used only for disparity derivation:

Max out of 4 corners:

A: top-left, B: top-right, C: bottom-left and D: bottom-right

$$A = d(Cb)_{x,y}$$

$$B = d(Cb)_{x+size\_h>>mixed\_res,y}$$

$$C = d(Cb)_{x,y+size\_v>>mixed\_res}$$

$$D = d(Cb)_{x+size\_h>>mixed\_res,y+size\_h>>mixed\_res}$$

$$d = \max(A,B,C,D)$$

# Specification text changes, overview

Proposal: in SPS integer multiplier and an integer divisor for horizontal and vertical resolution ratios: dHM, dHD, dVM, and dVD

$$A = d(Cb)_{dL,dT}$$

$$B = d(Cb)_{dR,dT}$$

$$A = d(Cb)_{dL,dB}$$

$$B = d(Cb)_{dR,dB}$$

$$d = \max(A,B,C,D)$$

Options:

1) Unconstrained depth resolution

$$dL = x * dHM / dHD$$

$$dT = y * dVM / dVD$$

$$dR = (x + size\_h) * dHM / dHD$$

$$dB = (y + size\_v) * dVM / dVD$$

2) Power-of-2 divisor

$$dL = x * dHM \gg \log2(dHD)$$

$$dT = y * dVM \gg \log2(dVD)$$

$$dR = (x + size\_h) * dHM \gg \log2(dHD)$$

$$dB = (y + size\_v) * dVM \gg \log2(dVD)$$

3) Dyadic depth-to-texture resolution ratio

$$dL = x \gg \log2(dHD)$$

$$dT = y \gg \log2(dVD)$$

$$dR = (x + size\_h) \gg \log2(dHD)$$

$$dB = (y + size\_v) \gg \log2(dVD)$$

# Conclusions

Depth is used for view synthesis as post-processing for decoding (i.e. the conventional use case as considered in the CTC)

Anchor	Tested	Texture coding	Total (coded PSNR)	Total (synthesized PSNR)
MVC+D	3D-AVC	-21.75%	-20.35%	-18.07%
MVC+D	JCT3V-E0035	-21.55%	-23.30%	-19.11%

Depth is used only as side information for multiview texture coding and is not intended for view synthesis

Anchor	Tested	Total (coded PSNR)
MVC	3D-AVC	-9.56%
MVC	JCT3V-E0035	-17.52%
MVC	JCT3V-E0035 (GDV)	-13.48%

We thank MERL / Dong Tian for cross-checking (JCT3V-E0103)

**Thank you**

**NOKIA**  
Connecting People