

**INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC 1/SC 29/WG 4
MPEG VIDEO CODING**

ISO/IEC JTC 1/SC 29/WG 4 m 57349
July 2021, Online

Title: The comparisons of encoder-side and decoder-side depth estimation
Source: Dawid Mieloch, Adrian Dziembowski
(Poznań University of Technology)

Abstract & Recommendations

The document presents the comparisons of encoder-side and decoder-side depth estimation in multiple configurations. It is recommended to discuss the possibility of updating the configuration of IVDE in the G17 anchor with one of the proposed configurations.

1 Introduction

The document presents the multiple comparisons of encoder-side and decoder-side depth estimation that were evaluated using TMIV 9.0. The focus is put not only on the quality of the final synthesized views but also on the runtimes of encoding and decoding processes.

2 Experiments

The performed experiments were conducted in a reduced-frame configuration of TMIV 9.0. Section 2.1 describes the comparison of A17 and G17 experiments, while Section 2.2 includes a comparison of G17 with the modified configuration of IVDE 4.0.

2.1 A17 vs G17

2.1.1 Comparison of anchors:

Mandatory content - Proposal vs. Low/High-bitrate Anchors							Runtime ratio (%)					
Sequence		High-BR BD rate Y-PSNR	Low-BR BD rate Y-PSNR	Max delta Y-PSNR	High-BR BD rate IV-PSNR	Low-BR BD rate IV-PSNR	Pixel rate [%]	Pixel rate [GP/s]	Frame rate [Hz]	Atlas encoding	Video encoding	Decoding & Rendering
ClassroomVideo	A	---	38.8%	5.66	105.8%	-15.7%	100%	1.07	30	0.4%	110.4%	995.7%
Museum	B	287.0%	37.2%	9.70	73.5%	-6.5%	#VALUE!	#VALUE!	30	0.3%	72.0%	935.8%
Fan	O	---	-89.6%	10.81	-66.8%	-77.3%	100%	1.07	30	1.1%	91.4%	3235.1%
Kitchen	J	-38.2%	-33.7%	11.74	24.1%	3.4%	#VALUE!	#VALUE!	30	0.8%	87.2%	3384.4%
Painter	D	-55.3%	-60.2%	8.99	-30.3%	-48.2%	100%	1.07	30	1.2%	87.3%	4038.2%
Frog	E	-72.1%	-64.2%	7.61	-56.4%	-57.0%	100%	1.07	30	1.4%	99.8%	8033.9%
Carpark	P	43.1%	4.4%	11.01	42.3%	7.7%	#VALUE!	#VALUE!	25	1.7%	138.7%	2791.8%
Chess	N	---	---	32.01	---	---	100%	1.07	30	0.7%	75.3%	2080.3%
Group	R	---	---	22.62	---	---	100%	1.07	30	0.7%	85.5%	3347.2%
MIV		---	---	13.35	---	---	#VALUE!	#VALUE!		0.9%	94.2%	3204.7%

Optional content - Proposal vs. Low/High-bitrate Anchors							Runtime ratio (%)					
Sequence		High-BR BD rate Y-PSNR	Low-BR BD rate Y-PSNR	Max delta Y-PSNR	High-BR BD rate IV-PSNR	Low-BR BD rate IV-PSNR	Pixel rate [%]	Pixel rate [GP/s]	Frame rate [Hz]	Atlas encoding	Video encoding	Decoding & Rendering
Fencing	L	72.9%	-3.6%	13.35	24.3%	-27.3%	83%	0.89	25	1.2%	150.0%	2695.7%
Hall	T	1009.3%	322.6%	18.55	630.9%	220.3%	83%	0.89	25	1.9%	186.3%	1915.7%
Street	U	-65.2%	-46.8%	7.02	-33.1%	-30.6%	83%	0.89	25	2.0%	127.1%	3102.5%
ChessPieces	Q	---	---	34.86	---	---	#VALUE!	#VALUE!	30	0.5%	62.0%	2007.0%
Hijack	C	---	---	21.92	---	---	#VALUE!	#VALUE!	30	0.8%	83.4%	526.9%
Mirror	I	-14.4%	-42.1%	13.51	18.2%	-32.9%	100%	1.07	30	1.5%	111.0%	3218.3%
MIV		---	---	18.20	---	---	#VALUE!	#VALUE!		1.3%	119.9%	2244.4%

Comments:

- In general, the encoding is 100 times faster in G17 than in A17, while decoding (that includes depth estimation for all views) is ~32 times longer.
- The objective quality is better for most perspective sequences and sometimes for omnidirectional content for low bitrates.

2.1.2 EE1 vs G17 (encoder-side depth estimation vs decoder-side depth estimation):

Mandatory content - Proposal vs. Low/High-bitrate Anchors							Runtime ratio (%)					
Sequence		High-BR BD rate Y-PSNR	Low-BR BD rate Y-PSNR	Max delta Y-PSNR	High-BR BD rate IV-PSNR	Low-BR BD rate IV-PSNR	Pixel rate [%]	Pixel rate [GP/s]	Frame rate [Hz]	Atlas encoding	Video encoding	Decoding & Rendering
ClassroomVideo	A	-73.4%	-82.3%	5.66	-67.0%	-82.2%	100%	1.07	30	0.4%	53.4%	955.5%
Museum	B	---	---	9.70	-87.5%	-85.6%	#VALUE!	#VALUE!	30	0.1%	42.7%	918.1%
Fan	O	-53.9%	-67.4%	10.81	-42.2%	-57.2%	100%	1.07	30	1.0%	80.6%	2929.3%
Kitchen	J	-61.4%	-55.5%	11.74	-11.4%	-22.0%	#VALUE!	#VALUE!	30	0.6%	60.8%	2687.8%
Painter	D	-73.8%	-73.1%	8.99	-60.7%	-66.2%	100%	1.07	30	1.2%	74.8%	4322.0%
Frog	E	-63.7%	-60.7%	7.61	-51.2%	-55.9%	100%	1.07	30	1.2%	85.6%	7297.2%
Carpark	P	-0.0%	-28.3%	11.01	-2.9%	-27.5%	#VALUE!	#VALUE!	25	1.7%	83.2%	2689.7%
Chess	N	---	---	32.01	---	---	100%	1.07	30	0.3%	62.4%	1837.0%
Group	R	---	---	22.62	---	---	100%	1.07	30	0.3%	73.1%	3266.1%
MIV		---	---	13.35	---	---	#VALUE!	#VALUE!		0.8%	68.5%	2989.2%

Optional content - Proposal vs. Low/High-bitrate Anchors							Runtime ratio (%)					
Sequence		High-BR BD rate Y-PSNR	Low-BR BD rate Y-PSNR	Max delta Y-PSNR	High-BR BD rate IV-PSNR	Low-BR BD rate IV-PSNR	Pixel rate [%]	Pixel rate [GP/s]	Frame rate [Hz]	Atlas encoding	Video encoding	Decoding & Rendering
Fencing	L	39.6%	-22.8%	13.35	18.1%	-39.2%	83%	0.89	25	1.6%	107.1%	2529.2%
Hall	T	1934.4%	739.9%	18.55	1249.8%	485.7%	83%	0.89	25	1.6%	204.3%	1956.8%
Street	U	-68.3%	-55.6%	7.02	-38.6%	-42.0%	83%	0.89	25	2.0%	85.7%	2967.2%
ChessPieces	Q	---	---	34.86	---	---	#VALUE!	#VALUE!	30	0.4%	53.0%	1931.4%
Hijack	C	---	---	21.92	---	---	#VALUE!	#VALUE!	30	0.3%	67.8%	506.5%
Mirror	I	-14.4%	-42.1%	13.51	18.2%	-32.9%	100%	1.07	30	1.0%	80.8%	2926.0%
MIV		---	---	18.20	---	---	#VALUE!	#VALUE!		1.2%	99.8%	2136.2%

Comments:

- The runtimes differences for TMIV encoding and decoding are very similar to the A17 vs G17 comparison. The VVC encoding is 40% faster.
- The objective quality is better for all mandatory content with an exception for SN (probably due to a bug in TMIV reported with anchors) and SR (wrong calculation of z far value).

2.1.3 EE1 vs G17, both with modified configuration of IVDE:

"NumOfThreads": 4 (up from 2),

"NumberOfCycles": 2 (up from 1),

"TemporalEnhancementThresh": 1.5 (up from 0.5).

Mandatory content - Proposal vs. Low/High-bitrate Anchors							Runtime ratio (%)					
Sequence		High-BR	Low-BR	Max delta	High-BR	Low-BR	Pixel rate [%]	Pixel rate [GPs/s]	Frame rate [Hz]	Atlas encoding	Video encoding	Decoding & Rendering
		BD rate Y-PSNR	BD rate Y-PSNR		BD rate IV-PSNR	BD rate IV-PSNR						
ClassroomVideo	A	-87.6%	-88.5%	5.76	-75.4%	-85.3%	100%	1.07	30	0.3%	59.3%	975.9%
Museum	B	---	---	9.29	-77.2%	-75.3%	#VALUE!	#VALUE!	30	0.1%	63.1%	1622.7%
Fan	O	-59.0%	-69.3%	10.50	-45.4%	-59.7%	100%	1.07	30	1.5%	104.1%	4091.0%
Kitchen	J	-65.2%	-59.5%	13.08	8.6%	-19.3%	#VALUE!	#VALUE!	30	0.6%	70.6%	5257.7%
Painter	D	-70.7%	-70.7%	8.45	-50.8%	-60.3%	100%	1.07	30	1.7%	87.7%	6999.0%
Frog	E	-65.2%	-61.4%	7.60	-52.0%	-56.1%	100%	1.07	30	1.7%	112.6%	#####
Carpark	P	-23.7%	-42.1%	10.59	-13.5%	-37.6%	#VALUE!	#VALUE!	25	1.7%	135.8%	4391.3%
Chess	N	---	---	31.64	---	---	100%	1.07	30	0.5%	79.3%	2604.8%
Group	R	---	---	22.64	---	---	100%	1.07	30	0.2%	88.2%	5162.7%
MIV				13.28	---	---	#VALUE!	#VALUE!		0.9%	89.0%	4852.1%

Optional content - Proposal vs. Low/High-bitrate Anchors							Runtime ratio (%)					
Sequence		High-BR	Low-BR	Max delta	High-BR	Low-BR	Pixel rate [%]	Pixel rate [GPs/s]	Frame rate [Hz]	Atlas encoding	Video encoding	Decoding & Rendering
		BD rate Y-PSNR	BD rate Y-PSNR		BD rate IV-PSNR	BD rate IV-PSNR						
Fencing	L	20.0%	-31.2%	13.30	13.7%	-42.0%	83%	0.89	25	1.9%	137.0%	3994.8%
Hall	T	2124.5%	794.6%	17.64	913.7%	155.6%	83%	0.89	25	2.3%	320.6%	4058.4%
Street	U	-65.4%	-51.2%	7.03	-30.6%	-35.4%	83%	0.89	25	2.0%	134.2%	5318.7%
ChessPieces	Q	---	---	34.87	---	---	#VALUE!	#VALUE!	30	0.6%	55.2%	3785.9%
Hijack	C	---	---	21.32	---	---	#VALUE!	#VALUE!	30	0.6%	90.3%	744.7%
Mirror	I	-30.4%	-43.8%	12.58	-6.2%	-30.6%	100%	1.07	30	1.6%	110.8%	4200.0%
MIV				17.79	---	---	#VALUE!	#VALUE!		1.5%	141.3%	3683.7%

Comments:

- The runtimes for decoding are longer than in previous experiments, but twice as many threads are used by IVDE, so wall time is similar.
- The differences in objective quality are similar despite the better quality of depth maps (see Section 2.2.1). The comparison of DSDE vs ESDE seems to be not very dependent on the used configuration of the depth estimator.

2.2 G17 vs modified G17

2.2.1 Higher quality:

"NumOfThreads": 4 (up from 2),

"NumberOfCycles": 2 (up from 1),

"TemporalEnhancementThresh": 1.5 (up from 0.5).

Mandatory content - Proposal vs. Low/High-bitrate Anchors						Runtime ratio (%)						
Sequence		High-BR BD rate Y-PSNR	Low-BR BD rate Y-PSNR	Max delta Y-PSNR	High-BR BD rate IV-PSNR	Low-BR BD rate IV-PSNR	Pixel rate [%]	Pixel rate [GP/s]	Frame rate [Hz]	Atlas encoding	Video encoding	Decoding & Rendering
ClassroomVideo	A	-36.2%	-23.1%	5.76	-10.4%	-8.1%	100%	1.07	30	100.0%	100.0%	88.5%
Museum	B	-17.8%	-14.1%	9.29	-7.5%	-6.6%	#VALUE!	#VALUE!	30	100.0%	100.0%	188.4%
Fan	O	-8.4%	-3.4%	10.50	-10.5%	-5.6%	100%	1.07	30	100.0%	100.0%	160.9%
Kitchen	J	11.6%	2.0%	13.08	17.3%	5.8%	#VALUE!	#VALUE!	30	100.0%	100.0%	186.9%
Painter	D	-26.1%	-15.3%	8.45	-23.7%	-14.6%	100%	1.07	30	100.0%	100.0%	165.2%
Frog	E	5.5%	2.6%	7.60	3.3%	1.3%	100%	1.07	30	100.0%	100.0%	140.7%
Carpark	P	-64.9%	-48.5%	10.59	-48.1%	-38.4%	#VALUE!	#VALUE!	25	100.0%	100.0%	152.9%
Chess	N	---	---	31.64	145.9%	189.7%	100%	1.07	30	100.0%	100.0%	140.3%
Group	R	-35.2%	-34.9%	22.64	-17.2%	-9.6%	100%	1.07	30	100.0%	100.0%	171.9%
MIV		---	---	13.28	5.5%	12.6%	#VALUE!	#VALUE!		100.0%	100.0%	155.1%

Optional content - Proposal vs. Low/High-bitrate Anchors						Runtime ratio (%)						
Sequence		High-BR BD rate Y-PSNR	Low-BR BD rate Y-PSNR	Max delta Y-PSNR	High-BR BD rate IV-PSNR	Low-BR BD rate IV-PSNR	Pixel rate [%]	Pixel rate [GP/s]	Frame rate [Hz]	Atlas encoding	Video encoding	Decoding & Rendering
Fencing	L	-16.0%	-7.3%	13.30	1.5%	1.3%	83%	0.89	25	100.0%	100.0%	152.5%
Hall	T	-23.1%	-87.8%	17.64	-32.2%	-88.0%	83%	0.89	25	100.0%	100.0%	176.0%
Street	U	-3.4%	-3.6%	7.03	-1.8%	-2.2%	83%	0.89	25	100.0%	100.0%	155.3%
ChessPieces	Q	-21.7%	-20.1%	34.87	10.9%	-7.6%	#VALUE!	#VALUE!	30	100.0%	100.0%	178.7%
Hijack	C	---	---	21.32	---	---	#VALUE!	#VALUE!	30	100.0%	100.0%	148.6%
Mirror	I	-23.2%	-12.9%	12.58	-27.7%	-12.8%	100%	1.07	30	100.0%	100.0%	159.6%
MIV		---	---	17.79	---	---	#VALUE!	#VALUE!		100.0%	100.0%	161.8%

Comments:

- The bitrates and encoding time are always the same in all G17 experiments, as only the configuration of IVDE is changed and depth estimation is performed always on the same input.
- The runtimes for decoding are longer than in previous experiments, but twice as many threads are used by IVDE, so wall time is similar.
- The objective quality is better for most sequences.

2.2.2 Similar quality, faster depth estimation:

"NumOfThreads": 4 (up from 2),
 "NumberOfSuperpixels": 25000 (down from 100000),
 "NumberOfZSteps": 128 (down from 256),
 "NumberOfCycles": 2 (up from 1),
 "TemporalEnhancementThresh": 1.5 (up from 0.5).

Mandatory content - Proposal vs. Low/High-bitrate Anchors							Runtime ratio (%)					
Sequence		High-BR BD rate Y-PSNR	Low-BR BD rate Y-PSNR	Max delta Y-PSNR	High-BR BD rate IV-PSNR	Low-BR BD rate IV-PSNR	Pixel rate [%]	Pixel rate [GP/s]	Frame rate [Hz]	Atlas encoding	Video encoding	Decoding & Rendering
ClassroomVideo	A	45.3%	21.1%	6.01	18.9%	10.3%	100%	1.07	30	100.0%	100.0%	26.1%
Museum	B	3.3%	-2.2%	9.40	-0.8%	-1.3%	100%	1.07	30	100.0%	100.0%	40.8%
Fan	O	8.0%	6.6%	10.75	4.6%	5.4%	100%	1.07	30	100.0%	100.0%	23.3%
Kitchen	J	53.0%	34.1%	12.13	28.0%	20.4%	#VALUE!	#VALUE!	30	100.0%	100.0%	27.0%
Painter	D	-17.2%	-10.0%	8.86	-17.3%	-10.9%	100%	1.07	30	100.0%	100.0%	22.4%
Frog	E	11.7%	6.7%	7.76	15.3%	8.4%	100%	1.07	30	100.0%	100.0%	17.9%
Carpark	P	-28.9%	-21.1%	10.95	-27.4%	-21.7%	#VALUE!	#VALUE!	25	100.0%	100.0%	27.1%
Chess	N	---	---	30.95	215.4%	150.4%	100%	1.07	30	100.0%	100.0%	34.3%
Group	R	-37.7%	-37.5%	22.63	-5.7%	-5.4%	100%	1.07	30	100.0%	100.0%	29.5%
MIV		---	---	13.27	25.7%	17.3%	#VALUE!	#VALUE!		100.0%	100.0%	27.6%

Optional content - Proposal vs. Low/High-bitrate Anchors							Runtime ratio (%)					
Sequence		High-BR BD rate Y-PSNR	Low-BR BD rate Y-PSNR	Max delta Y-PSNR	High-BR BD rate IV-PSNR	Low-BR BD rate IV-PSNR	Pixel rate [%]	Pixel rate [GP/s]	Frame rate [Hz]	Atlas encoding	Video encoding	Decoding & Rendering
Fencing	L	---	---	13.80	370.9%	165.0%	83%	0.89	25	100.0%	100.0%	24.6%
Hall	T	---	---	19.85	---	---	83%	0.89	25	100.0%	100.0%	28.5%
Street	U	294.9%	87.0%	7.87	100.1%	34.4%	83%	0.89	25	100.0%	100.0%	23.4%
ChessPieces	Q	82.9%	43.1%	32.80	56.1%	27.1%	#VALUE!	#VALUE!	30	100.0%	100.0%	37.6%
Hijack	C	19.2%	-75.4%	21.67	-24.6%	-8.0%	#VALUE!	#VALUE!	30	100.0%	100.0%	55.9%
Mirror	I	0.4%	-2.6%	13.04	-4.4%	-3.7%	100%	1.07	30	100.0%	100.0%	30.5%
MIV		---	---	18.17	---	---	#VALUE!	#VALUE!		100.0%	100.0%	33.4%

Comments:

- The runtimes for decoding are much lower (4 times faster than in the G17 anchor) due to decreased number of depth levels and superpixels in each view during the depth estimation. Moreover, twice as many threads are used by IVDE.
- The objective quality differences vary for different sequences and are on average the same as in G17 for mandatory sequences (when N is not taken into account).
- When compared to the A17 anchor (below), the decoding time difference A17 and G17 becomes much smaller (down from 32 times slower to 8 times slower):

Mandatory content - Proposal vs. Low/High-bitrate Anchors							Runtime ratio (%)					
Sequence		High-BR BD rate Y-PSNR	Low-BR BD rate Y-PSNR	Max delta Y-PSNR	High-BR BD rate IV-PSNR	Low-BR BD rate IV-PSNR	Pixel rate [%]	Pixel rate [GP/s]	Frame rate [Hz]	Atlas encoding	Video encoding	Decoding & Rendering
ClassroomVideo	A	---	33.0%	6.01	138.4%	-15.6%	100%	1.07	30	0.4%	110.4%	260.3%
Museum	B	328.6%	37.6%	9.40	72.3%	-7.3%	#VALUE!	#VALUE!	30	0.3%	72.0%	381.7%
Fan	O	---	-89.1%	10.75	-65.9%	-76.4%	100%	1.07	30	1.1%	91.4%	752.5%
Kitchen	J	-4.4%	-9.2%	12.13	55.4%	25.4%	#VALUE!	#VALUE!	30	0.8%	87.2%	913.6%
Painter	D	-62.1%	-63.5%	8.86	-37.2%	-51.4%	100%	1.07	30	1.2%	87.3%	905.0%
Frog	E	-69.4%	-62.7%	7.76	-50.8%	-54.3%	100%	1.07	30	1.4%	99.8%	1437.2%
Carpark	P	29.7%	-2.8%	10.95	31.9%	-1.3%	#VALUE!	#VALUE!	25	1.7%	138.7%	756.2%
Chess	N	---	---	30.95	---	---	100%	1.07	30	0.7%	75.3%	713.4%
Group	R	---	---	22.63	---	---	100%	1.07	30	0.7%	85.5%	986.2%
MIV		---	---	13.27	---	---	#VALUE!	#VALUE!		0.9%	94.2%	789.6%

Optional content - Proposal vs. Low/High-bitrate Anchors							Runtime ratio (%)					
Sequence		High-BR BD rate Y-PSNR	Low-BR BD rate Y-PSNR	Max delta Y-PSNR	High-BR BD rate IV-PSNR	Low-BR BD rate IV-PSNR	Pixel rate [%]	Pixel rate [GP/s]	Frame rate [Hz]	Atlas encoding	Video encoding	Decoding & Rendering
Fencing	L	125.3%	20.4%	13.80	83.5%	-9.8%	83%	0.89	25	1.2%	150.0%	664.1%
Hall	T	---	762.2%	19.85	---	388.5%	83%	0.89	25	1.9%	186.3%	545.3%
Street	U	30.5%	-3.4%	7.87	24.7%	-7.8%	83%	0.89	25	2.0%	127.1%	724.7%
ChessPieces	Q	---	---	32.80	---	---	#VALUE!	#VALUE!	30	0.5%	62.0%	754.7%
Hijack	C	---	---	21.67	---	---	#VALUE!	#VALUE!	30	0.8%	83.4%	294.5%
Mirror	I	-12.5%	-43.9%	13.04	21.2%	-35.8%	100%	1.07	30	1.5%	111.0%	980.1%
MIV		---	---	18.17	---	---	#VALUE!	#VALUE!		1.3%	119.9%	660.6%

2.2.3 Lower quality, much faster depth estimation:

"NumOfThreads": 4 (up from 2),
 "NumberOfSuperpixels": 12500 (down from 100000),
 "NumberOfZSteps": 64 (down from 256),
 "NumberOfCycles": 2 (up from 1),
 "TemporalEnhancementThresh": 1.5 (up from 0.5).

Mandatory content - Proposal vs. Low/High-bitrate Anchors							Runtime ratio (%)					
Sequence		High-BR BD rate Y-PSNR	Low-BR BD rate Y-PSNR	Max delta Y-PSNR	High-BR BD rate IV-PSNR	Low-BR BD rate IV-PSNR	Pixel rate [%]	Pixel rate [GP/s]	Frame rate [Hz]	Atlas encoding	Video encoding	Decoding & Rendering
ClassroomVideo	A	550.2%	138.3%	6.41	54.5%	28.5%	100%	1.07	30	100.0%	100.0%	16.7%
Museum	B	54.0%	21.7%	9.55	10.0%	4.4%	#VALUE!	#VALUE!	30	100.0%	100.0%	23.8%
Fan	O	52.4%	27.8%	11.28	19.7%	13.4%	100%	1.07	30	100.0%	100.0%	14.6%
Kitchen	J	379.9%	113.8%	14.42	90.8%	46.5%	#VALUE!	#VALUE!	30	100.0%	100.0%	17.8%
Painter	D	19.3%	9.7%	9.81	-8.4%	-6.9%	100%	1.07	30	100.0%	100.0%	14.1%
Frog	E	16.8%	9.6%	7.89	20.8%	11.8%	100%	1.07	30	100.0%	100.0%	7.6%
Carpark	P	466.8%	173.3%	11.75	43.6%	33.0%	#VALUE!	#VALUE!	25	100.0%	100.0%	13.9%
Chess	N	---	---	31.07	---	---	100%	1.07	30	100.0%	100.0%	20.1%
Group	R	-10.1%	-21.4%	22.46	27.4%	8.3%	100%	1.07	30	100.0%	100.0%	17.8%
MIV		---	---	13.85	---	---	#VALUE!	#VALUE!		100.0%	100.0%	16.3%

Optional content - Proposal vs. Low/High-bitrate Anchors							Runtime ratio (%)					
Sequence		High-BR BD rate Y-PSNR	Low-BR BD rate Y-PSNR	Max delta Y-PSNR	High-BR BD rate IV-PSNR	Low-BR BD rate IV-PSNR	Pixel rate [%]	Pixel rate [GP/s]	Frame rate [Hz]	Atlas encoding	Video encoding	Decoding & Rendering
Fencing	L	---	---	15.45	---	---	83%	0.89	25	100.0%	100.0%	13.2%
Hall	T	---	814.9%	17.72	-46.7%	-43.1%	83%	0.89	25	100.0%	100.0%	17.1%
Street	U	---	491.0%	9.57	267.6%	79.6%	83%	0.89	25	100.0%	100.0%	11.6%
ChessPieces	Q	---	---	31.94	---	---	#VALUE!	#VALUE!	30	100.0%	100.0%	22.7%
Hijack	C	---	---	22.34	---	---	#VALUE!	#VALUE!	30	100.0%	100.0%	46.0%
Mirror	I	50.8%	19.4%	14.38	28.4%	9.5%	100%	1.07	30	100.0%	100.0%	16.9%
MIV		---	---	18.57	---	---	#VALUE!	#VALUE!		100.0%	100.0%	21.2%

Comments:

- The runtimes for decoding are much lower (8 times faster than in the G17 anchor) due to even more decreased number of depth levels and superpixels in each view during the depth estimation. Moreover, twice as many threads are used by IVDE.
- The objective quality is lower in most sequences. However, when compared with A17, such configuration is still competitive for many sequences and the decoding time becomes even closer to the MIV anchor (4.5 times longer decoding with 100 times faster encoding):

Mandatory content - Proposal vs. Low/High-bitrate Anchors							Runtime ratio (%)					
Sequence		High-BR	Low-BR	Max delta	High-BR	Low-BR	Pixel rate [%]	Pixel rate [GP/s]	Frame rate [Hz]	Atlas	Video	Decoding & Rendering
		BD rate Y-PSNR	BD rate Y-PSNR		BD rate IV-PSNR	BD rate IV-PSNR				encoding	encoding	
ClassroomVideo	A	---	150.4%	6.41	196.8%	-14.0%	100%	1.07	30	0.4%	110.4%	166.2%
Museum	B	---	---	9.55	86.1%	-4.0%	#VALUE!	#VALUE!	30	0.3%	72.0%	222.8%
Fan	O	---	-88.3%	11.28	-62.7%	-75.3%	100%	1.07	30	1.1%	91.4%	471.3%
Kitchen	J	---	43.4%	14.42	---	51.3%	#VALUE!	#VALUE!	30	0.8%	87.2%	601.8%
Painter	D	-47.6%	-56.9%	9.81	-36.4%	-50.5%	100%	1.07	30	1.2%	87.3%	571.2%
Frog	E	-68.4%	-61.9%	7.89	-49.0%	-53.1%	100%	1.07	30	1.4%	99.8%	606.9%
Carpark	P	88.5%	35.0%	11.75	54.8%	16.5%	#VALUE!	#VALUE!	25	1.7%	138.7%	388.5%
Chess	N	---	---	31.07	---	---	100%	1.07	30	0.7%	75.3%	418.1%
Group	R	---	---	22.46	---	---	100%	1.07	30	0.7%	85.5%	594.4%
MIV		---	---	13.85	---	---	#VALUE!	#VALUE!		0.9%	94.2%	449.0%

Optional content - Proposal vs. Low/High-bitrate Anchors							Runtime ratio (%)					
Sequence		High-BR	Low-BR	Max delta	High-BR	Low-BR	Pixel rate [%]	Pixel rate [GP/s]	Frame rate [Hz]	Atlas	Video	Decoding & Rendering
		BD rate Y-PSNR	BD rate Y-PSNR		BD rate IV-PSNR	BD rate IV-PSNR				encoding	encoding	
Fencing	L	---	24.0%	15.45	227.7%	8.4%	83%	0.89	25	1.2%	150.0%	356.7%
Hall	T	---	585.5%	17.72	---	610.1%	83%	0.89	25	1.9%	186.3%	327.5%
Street	U	306.9%	70.5%	9.57	66.1%	13.8%	83%	0.89	25	2.0%	127.1%	358.5%
ChessPieces	Q	---	---	31.94	---	---	#VALUE!	#VALUE!	30	0.5%	62.0%	455.0%
Hijack	C	---	---	22.34	---	---	#VALUE!	#VALUE!	30	0.8%	83.4%	242.4%
Mirror	I	-0.2%	-39.8%	14.38	28.2%	-34.8%	100%	1.07	30	1.5%	111.0%	543.4%
MIV		---	---	18.57	---	---	#VALUE!	#VALUE!		1.3%	119.9%	380.6%

3 Posetraces

P01 posetraces for QP1 and QP5 were uploaded to MPEG content server. For all sequences we provide:

- A17 anchor,
- G17 anchor,
- G17 from section 2.2.1 (2_cycles in the filenames),
- G17 from section 2.2.2 (25k_128 in the filenames),
- G17 from section 2.2.3 (12k_64 in the filenames).

Other posetraces are available upon request.

4 Recommendations

It is recommended to discuss the possibility of updating the configuration of IVDE in G17 anchor with one of the proposed configurations.

Acknowledgement

The research was supported by the Ministry of Science and Higher Education of Republic of Poland.