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 60668
October 2022, Mainz, DE

Title: **[MIV] Inhomogeneous view selection for omnidirectional content**

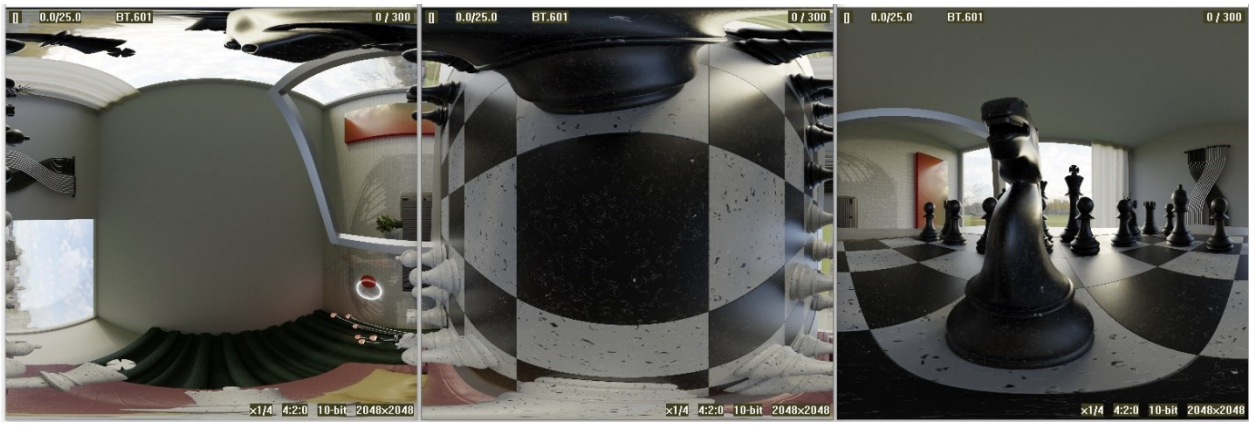
Source: **Adrian Dziembowski, Dominika Klóska (PUT)**
Jun Young Jeong, Gwangsoon Lee (ETRI)

Abstract

The view selector in TMIV14 analyzes camera arrangement in order to select basic and additional views. Such a selection performs well for perspective content with cameras arranged in a line or an array. However, for omnidirectional content, the efficiency of current view selector is worse than expected. We propose to modify the view selection algorithm for omnidirectional sequences, in order to better handle typical movements of the final viewer. The recommendation is to include proposed modification into TMIV15.

1 Introduction

The current view selector tries to find views, which are most distant to each other. In some cases, it chooses views, which are not important to the viewer (i.e., he or she does not watch these fragments for most of the time, e.g., ceiling or floor). An example is presented in Fig. below, where two first views are transmitted as basic ones (SN, G17), while the last one – containing the most crucial information – is omitted.



2 Proposal

The proposed approach takes into account the typical user movements trying to increase the quality of most probable virtual trajectories of the user. Therefore, camera distances calculated in the view selection process are not homogenous, and the vertical direction is treated differently from horizontal directions.

In proposed approach, the view selection algorithm depends on the repulsion and attraction cost:

$$J = 2 \sum_{\substack{1 \le i < k \\ i < j \le k}} r_{c_i,c_j}^{-2}$$

$$J = - \sum_{1 \le i < c, c < i \le n} r_{c,i}^{-2}$$

but the distance between cameras $r_{i,j}$ is calculated as:

$$r_{i,j} = \sqrt{\left(r_{i,j}^{x}\right)^2 + \left(r_{i,j}^{y}\right)^2 + \left(w \cdot r_{i,j}^{z}\right)^2}$$

where r^x, r^y, r^z indicate a distance between two cameras among three axes, and w is the inhomogeneity coefficient. In the experiments, w was set to 0.4.

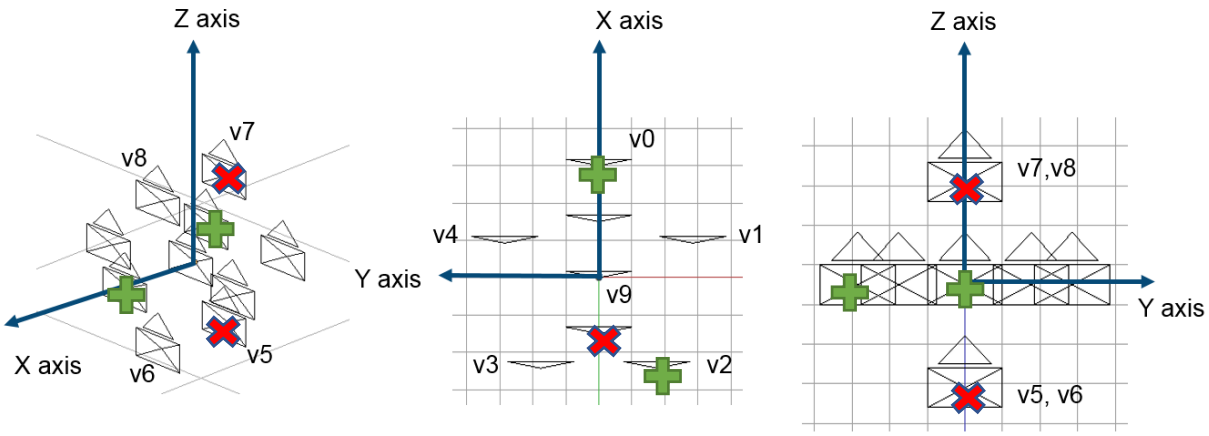
Chosen views (black – views selected for anchor and proposed, red – views selected for anchor only, green – views selected for proposed only)

Anchor	SA	SB	SC	SN	SQ
A17	0	2, 19	9	1, 5	1, 5
V17	9, 14	1, 3, 12, 15	0, 2	1, 3, 5, 7	1, 3, 5, 7
G17	9, 10, 13, 14	0, 1, 6, 10, 11, 17, 19, 22	1, 4, 5, 7	0, 2, 4, 5, 7, 8, 9	0, 2, 4, 5, 7, 8, 9

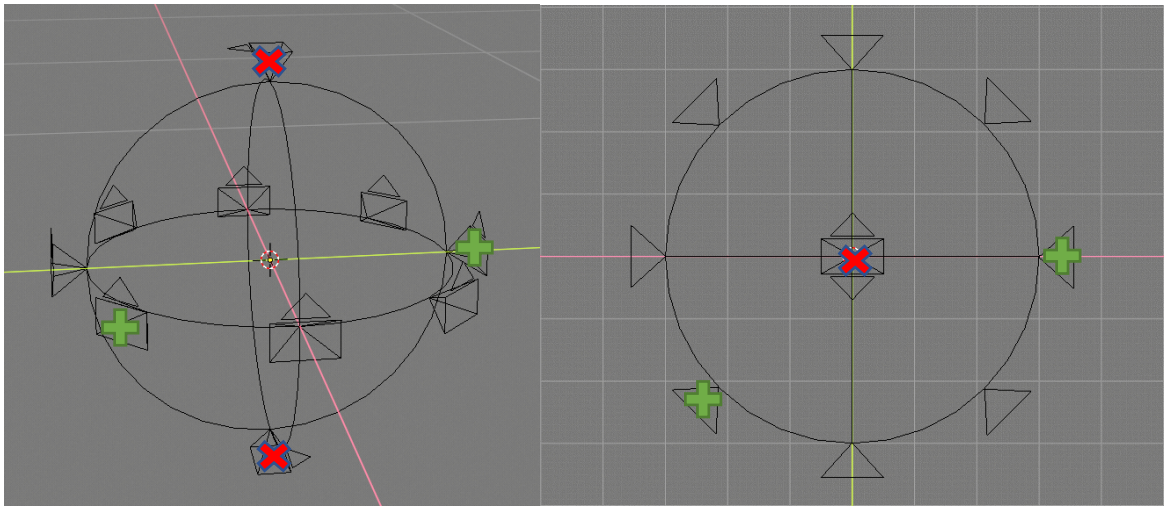
Proposed	SA	SB	SC	SN	SQ
A17	0	8, 19	9	1, 5	1, 5
V17	9, 14	3, 8, 17, 19	0, 2	1, 3, 5, 7	1, 3, 5, 7

G17	9, 10, 13, 14	1, 2, 3, 5, 13, 16, 17, 19	0, 1, 2, 4	2, 3, 4, 5, 6, 7, 8	2, 3, 4, 5, 6, 7, 8

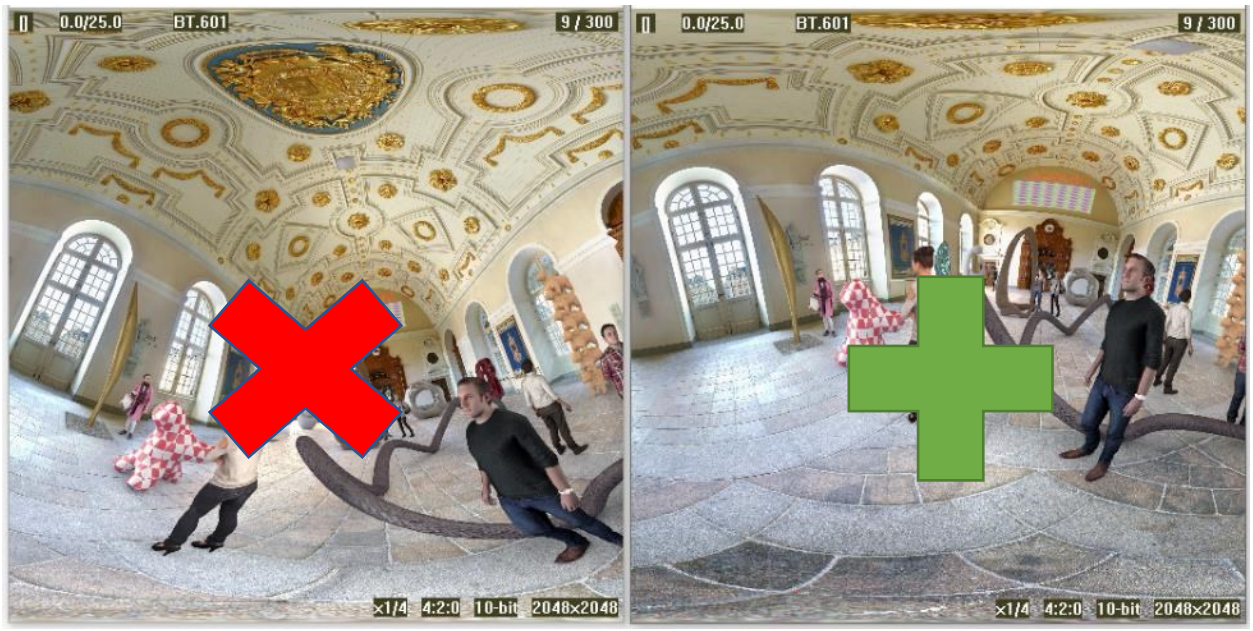
SC – chosen views change (G17):



SN – chosen views change (G17):



SB – chosen views change (A17):



3 Experimental results

3.1 A17

Change only for SB:

Mandatory content - Proposal vs. Low/High-bitrate Anchors					
Sequence		High-BR BD rate Y-PSNR	Low-BR BD rate Y-PSNR	High-BR BD rate IV-PSNR	Low-BR BD rate IV-PSNR
ClassroomVideo	A	0.0%	0.0%	0.0%	0.0%
Museum	B	-3.4%	0.2%	0.9%	2.3%
Chess	N	0.0%	0.0%	0.0%	0.0%
ChessPieces	Q	0.0%	0.0%	0.0%	0.0%
Hijack	C	0.0%	0.0%	0.0%	0.0%
MIV		**-0.7%**	**0.0%**	**0.2%**	**0.5%**

Runtime ratio (%)		
Atlas encoding	Video encoding	Decoding & Rendering
92.6%	83.0%	104.1%
104.2%	93.1%	77.0%
88.0%	109.0%	81.4%
111.5%	82.2%	65.7%
101.8%	81.0%	87.0%
99.6%	**89.7%**	**83.1%**

Max delta Y-PSNR [dB]		
MIV	#######	Difference [%]
1.10	1.10	0.0%
9.28	8.62	-7.1%
13.04	13.04	0.0%
13.85	13.85	0.0%
8.17	8.17	0.0%
9.09	**8.96**	**-1.4%**

Max delta IV-PSNR [dB]		
MIV	#######	Difference [%]
0.76	0.76	0.0%
5.32	4.88	-8.2%
12.35	12.35	0.0%
14.49	14.49	0.0%
6.93	6.93	0.0%
7.97	**7.88**	-1.6%

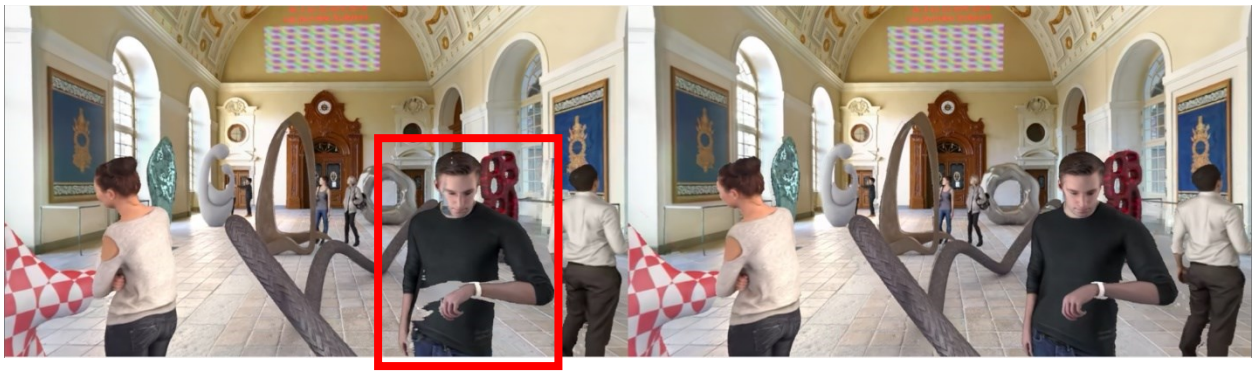
SB, QP1, p01:

3.2 V17

Change only for SB:

Mandatory content - Proposal vs. Low/High-bitrate Anchors						Runtime ratio (%)			Max delta Y-PSNR [dB]			Max delta IV-PSNR [dB]		
Sequence		High-BR BD rate Y-PSNR	Low-BR BD rate Y-PSNR	High-BR BD rate IV-PSNR	Low-BR BD rate IV-PSNR	Atlas encoding	Video encoding	Decoding & Rendering	MIV view	#######	Difference [%]	MIV view	#######	Difference [%]
ClassroomVideo	A	0.0%	0.0%	0.0%	0.0%	101.3%	79.0%	101.6%	1.53	1.53	0.0%	1.24	1.24	0.0%
Museum	B	-4.6%	-5.5%	-7.7%	-3.2%	102.1%	83.1%	74.6%	9.28	7.20	-22.4%	8.38	6.08	-27.4%
Chess	N	0.0%	0.0%	0.0%	0.0%	86.8%	83.5%	105.8%	22.05	22.05	0.0%	22.35	22.35	0.0%
ChessPieces	Q	0.0%	0.0%	0.0%	0.0%	107.1%	86.2%	66.4%	23.58	23.58	0.0%	24.48	24.48	0.0%
Hijack	C	0.0%	0.0%	0.0%	0.0%	103.9%	72.7%	116.9%	15.27	15.27	0.0%	15.31	15.31	0.0%
MIV		**-0.9%**	**-1.1%**	**-1.5%**	**-0.6%**	**100.2%**	**80.9%**	**93.1%**	**14.34**	**13.92**	**-4.5%**	**14.35**	**13.89**	-5.5%

SB, QP1, p02:



3.3 G17

Change for SB, SC, SN, and SQ:

Mandatory content - Proposal vs. Low/High-bitrate Anchors						Runtime ratio (%)			Max delta Y-PSNR [dB]			Max delta IV-PSNR [dB]		
Sequence		High-BR BD rate Y-PSNR	Low-BR BD rate Y-PSNR	High-BR BD rate IV-PSNR	Low-BR BD rate IV-PSNR	Atlas encoding	Video encoding	Decoding & Rendering	MIV DSDE	#######	Difference [%]	MIV DSDE	#######	Difference [%]
ClassroomVideo	A	0.0%	0.0%	0.0%	0.0%	68.6%	110.6%	98.8%	5.69	5.69	0.0%	4.05	4.05	0.0%
Museum	B	57.2%	32.5%	13.1%	10.2%	110.8%	111.2%	117.4%	9.18	9.70	5.7%	6.34	7.04	11.1%
Hijack	C	---	---	---	---	116.8%	114.9%	105.9%	22.33	26.93	20.6%	21.03	25.98	23.5%
Chess	N	---	---	---	---	71.5%	88.2%	78.0%	25.19	27.97	11.0%	23.89	27.11	13.5%
ChessPieces	Q	---	---	---	---	87.7%	78.6%	82.9%	27.71	33.01	19.2%	25.79	31.25	21.2%
MIV		---	---	---	---	**91.1%**	**100.7%**	**96.6%**	**18.02**	**20.66**	**11.3%**	**16.22**	**19.09**	13.9%

SB, QP5, p03:

SC, QP1, p01:

SN, QP1, p01:

SQ, QP3, p03:

Why objective results indicate that anchor is better? And why there is a discrepancy between objective and subjective results?

Two reasons:

1. We decreased the quality of views which are not important to the viewer (v0 and v9, looking to the top and bottom),
2. The (negative) difference of IV-PSNR and WS-PSNR for new basic views is lower, than (positive) difference for basic views which are not basic in proposed approach (results for SN):

View Synthesis [MIV DSDE]								
View		Y-MSE	Y-PSNR	U-PSNR	V-PSNR	IV-MSE	IV-PSNR	Total T [s]
v0		2.7E-05	45.64	55.11	57.60	2.0E-06	57.00	17201
		4.3E-05	43.69	52.76	55.31	3.7E-06	54.36	17354
		6.5E-05	41.85	50.84	53.37	6.3E-06	52.03	24897
		1.0E-04	39.83	48.53	50.91	1.2E-05	49.09	20607
		1.7E-04	37.62	46.88	49.65	2.1E-05	46.72	22325
v1		2.8E-03	25.46	48.64	53.47	1.6E-04	37.84	17201
		2.9E-03	25.43	48.15	52.81	1.7E-04	37.62	17354
		2.9E-03	25.32	47.59	51.83	1.9E-04	37.30	24897
		3.0E-03	25.16	46.53	50.20	2.1E-04	36.74	20607
		3.0E-03	25.16	45.48	49.01	2.3E-04	36.35	22325
v2		3.0E-05	45.28	54.26	57.80	2.1E-06	56.78	17201
		5.0E-05	42.98	51.88	55.63	4.0E-06	53.94	17354
		7.9E-05	41.05	49.82	53.74	7.1E-06	51.48	24897
		1.3E-04	38.86	47.40	51.09	1.5E-05	48.34	20607
		2.2E-04	36.57	45.82	49.58	2.6E-05	45.82	22325
v3		7.5E-03	21.27	44.31	51.37	4.8E-04	33.20	17201
		7.7E-03	21.15	44.09	50.94	5.0E-04	33.01	17354
		7.3E-03	21.36	43.99	50.60	4.9E-04	33.13	24897
		7.2E-03	21.43	44.11	49.80	4.9E-04	33.09	20607
		7.8E-03	21.07	43.29	48.74	5.6E-04	32.54	22325
v4		2.7E-05	45.65	54.75	57.76	2.1E-06	56.73	17201
		5.0E-05	43.05	52.33	55.45	4.3E-06	53.62	17354
		7.8E-05	41.06	50.21	53.56	8.2E-06	50.87	24897
		1.4E-04	38.47	48.02	51.13	1.8E-05	47.52	20607
		2.3E-04	36.30	46.31	49.46	3.1E-05	45.02	22325
v5		2.8E-05	45.47	54.27	56.82	2.5E-06	56.01	17201
		5.6E-05	42.55	51.88	54.52	5.2E-06	52.87	17354
		9.2E-05	40.38	49.78	52.63	9.7E-06	50.13	24897
		1.7E-04	37.77	47.81	50.53	2.0E-05	46.90	20607
		2.9E-04	35.43	46.10	48.97	3.8E-05	44.19	22325
v6		2.1E-03	26.86	49.18	48.34	1.3E-04	38.85	17201
		2.0E-03	26.96	48.85	47.82	1.3E-04	38.74	17354
		2.1E-03	26.87	48.23	47.57	1.4E-04	38.53	24897
		2.2E-03	26.63	47.23	46.90	1.6E-04	38.04	20607
		2.2E-03	26.67	46.13	46.28	1.7E-04	37.76	22325
v7		2.9E-05	45.40	54.34	56.91	2.5E-06	56.02	17201
		5.1E-05	42.96	51.85	54.47	4.8E-06	53.19	17354
		8.3E-05	40.80	50.06	52.39	8.9E-06	50.49	24897
		1.5E-04	38.18	47.91	50.14	1.9E-05	47.26	20607
		2.7E-04	35.72	46.47	48.71	3.6E-05	44.49	22325
v8		2.3E-05	46.47	55.09	57.58	2.0E-06	57.08	17201
		4.0E-05	43.98	52.75	55.26	3.7E-06	54.37	17354
		6.4E-05	41.94	50.78	53.25	6.6E-06	51.83	24897
		1.2E-04	39.31	48.47	50.80	1.4E-05	48.62	20607
		2.1E-04	36.73	47.05	49.40	2.7E-05	45.76	22325
v9		3.6E-05	44.49	55.62	58.75	2.3E-06	56.34	17201
		6.0E-05	42.24	53.36	56.63	4.1E-06	53.87	17354
		8.8E-05	40.55	51.43	54.72	7.4E-06	51.28	24897
		1.5E-04	38.11	49.04	52.11	1.7E-05	47.81	20607
		2.7E-04	35.71	47.46	50.85	3.4E-05	44.71	22325

View Synthesis [ ]								
View		Y-MSE	Y-PSNR	U-PSNR	V-PSNR	IV-MSE	IV-PSNR	Total T [s]
v0		3.1E-03	25.16	46.30	47.91	2.4E-04	36.17	16830
		3.4E-03	24.64	45.57	47.97	2.8E-04	35.58	16612
		3.3E-03	24.84	45.47	47.93	2.6E-04	35.79	15432
		3.2E-03	24.96	44.94	46.72	2.8E-04	35.58	15446
		3.3E-03	24.87	44.36	46.09	3.0E-04	35.25	15496
v1		2.6E-03	25.86	47.77	52.43	1.6E-04	37.99	16830
		2.7E-03	25.68	47.34	51.84	1.7E-04	37.70	16612
		2.7E-03	25.66	47.05	51.34	1.7E-04	37.60	15432
		2.8E-03	25.60	46.07	49.79	2.0E-04	37.09	15446
		2.8E-03	25.58	45.31	48.76	2.1E-04	36.75	15496
v2		2.5E-05	46.02	54.34	57.86	2.1E-06	56.76	16830
		4.3E-05	43.63	51.93	55.54	4.0E-06	53.98	16612
		7.0E-05	41.55	50.02	53.65	7.1E-06	51.51	15432
		1.2E-04	39.14	47.46	51.19	1.5E-05	48.37	15446
		2.2E-04	36.67	45.93	49.49	2.6E-05	45.82	15496
v3		3.4E-05	44.70	54.69	58.18	2.4E-06	56.28	16830
		5.4E-05	42.70	52.44	56.11	4.3E-06	53.63	16612
		8.1E-05	40.89	50.57	54.15	7.6E-06	51.17	15432
		1.4E-04	38.70	47.91	51.40	1.6E-05	47.96	15446
		2.2E-04	36.49	46.44	49.98	2.8E-05	45.47	15496
v4		2.9E-05	45.43	54.74	57.76	2.3E-06	56.44	16830
		5.2E-05	42.86	52.36	55.46	4.5E-06	53.42	16612
		8.0E-05	40.98	50.24	53.59	8.3E-06	50.80	15432
		1.4E-04	38.53	48.06	51.16	1.8E-05	47.51	15446
		2.3E-04	36.31	46.35	49.50	3.1E-05	45.02	15496
v5		2.8E-05	45.48	54.28	56.82	2.6E-06	55.93	16830
		5.5E-05	42.59	51.89	54.53	5.3E-06	52.76	16612
		9.0E-05	40.45	49.79	52.64	9.8E-06	50.11	15432
		1.7E-04	37.81	47.82	50.54	2.0E-05	46.89	15446
		2.9E-04	35.43	46.12	48.99	3.8E-05	44.19	15496
v6		2.8E-05	45.48	54.74	57.06	2.5E-06	56.11	16830
		5.3E-05	42.77	52.20	54.65	5.0E-06	52.98	16612
		9.0E-05	40.47	50.58	52.83	9.2E-06	50.37	15432
		1.6E-04	37.86	48.62	50.38	2.0E-05	47.04	15446
		2.9E-04	35.44	47.09	49.05	3.8E-05	44.26	15496
v7		2.9E-05	45.44	54.32	56.85	2.5E-06	56.04	16830
		5.0E-05	42.98	51.80	54.53	4.8E-06	53.18	16612
		8.5E-05	40.73	49.88	52.56	9.0E-06	50.47	15432
		1.5E-04	38.16	47.81	50.14	1.9E-05	47.25	15446
		2.7E-04	35.70	46.27	48.57	3.7E-05	44.36	15496
v8		2.1E-05	46.79	55.04	57.56	1.9E-06	57.18	16830
		3.8E-05	44.17	52.67	55.27	3.6E-06	54.39	16612
		6.3E-05	41.98	50.75	53.40	6.5E-06	51.85	15432
		1.2E-04	39.35	48.46	50.64	1.4E-05	48.62	15446
		2.1E-04	36.74	47.07	49.36	2.7E-05	45.74	15496
v9		1.3E-02	18.82	40.80	46.37	9.8E-04	30.08	16830
		1.3E-02	18.77	40.88	46.42	9.9E-04	30.06	16612
		1.3E-02	18.89	40.84	46.52	9.6E-04	30.16	15432
		1.3E-02	18.79	40.98	46.36	9.9E-04	30.05	15446
		1.3E-02	18.81	40.82	46.12	1.0E-03	30.01	15496

Deltas				
Y-PSNR	U-PSNR	V-PSNR	IV-PSNR	
-20.48	-8.80	-9.69	-20.84	
-19.05	-7.19	-7.34	-18.78	
-17.01	-5.37	-5.44	-16.24	
-14.87	-3.59	-4.19	-13.51	
-12.74	-2.52	-3.56	-11.47	
0.40	-0.87	-1.03	0.15	
0.25	-0.82	-0.97	0.08	
0.34	-0.54	-0.49	0.30	
0.43	-0.46	-0.41	0.35	
0.41	-0.17	-0.25	0.40	
0.75	0.08	0.06	-0.02	
0.66	0.05	-0.10	0.04	
0.50	0.20	-0.10	0.03	
0.29	0.06	0.10	0.03	
0.10	0.10	-0.09	0.00	
23.42	10.38	6.80	23.09	
21.56	8.35	5.16	20.61	
19.54	6.58	3.55	18.04	
17.27	3.80	1.60	14.86	
15.43	3.14	1.24	12.94	
-0.22	-0.01	0.00	-0.29	
-0.19	0.02	0.01	-0.19	
-0.08	0.03	0.03	-0.07	
0.06	0.04	0.03	-0.01	
0.01	0.05	0.04	0.00	
0.00	0.01	0.01	-0.08	
0.04	0.01	0.01	-0.12	
0.07	0.01	0.01	-0.02	
0.03	0.01	0.01	-0.01	
0.00	0.02	0.02	0.00	
18.62	5.56	8.72	17.26	
15.81	3.35	6.83	14.23	
13.60	2.35	5.26	11.85	
11.23	1.40	3.48	9.00	
8.77	0.95	2.77	6.49	
0.05	-0.02	-0.06	0.03	
0.03	-0.05	0.07	-0.01	
-0.07	-0.18	0.17	-0.01	
-0.02	-0.10	0.00	-0.01	
-0.02	-0.20	-0.14	-0.13	
0.33	-0.05	-0.02	0.10	
0.20	-0.08	0.02	0.02	
0.04	-0.04	0.15	0.03	
0.04	0.00	-0.16	0.00	
0.01	0.02	-0.04	-0.03	
-25.67	-14.82	-12.38	-26.26	
-23.47	-12.47	-10.21	-23.82	
-21.67	-10.59	-8.20	-21.13	
-19.33	-8.06	-5.75	-17.76	
-16.89	-6.64	-4.73	-14.71	

On request, all posetraces will be available on the MPEG content server.

4 Recommendation

We recommend including proposed modification into TMIV15.

5 Acknowledgement

This work was supported by Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2018-0-00207, Immersive Media Research Laboratory).