INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC 1/SC 29/WG 04
MPEG VIDEO CODING

ISO/IEC JTC 1/SC 29/WG 04 m 58002
October 2021, Online

Title: **Outcomes of exploration experiments – summary report**

Source: **Joel Jung (Tencent),**
Vinod Kumar Malamal Vadakital (Nokia),
Dawid Mieloch (PUT)

Abstract

This contribution is a summary of outcomes of all experiments listed in N0115. A total of 10 organizations participated in one or more of the listed experiments. Five main experiments were agreed upon, with all except EE-1 having additional sub-experiments. Significant participation and engagement from experimenters were observed, and several useful recommendations are provided from participating organizations.

1 Introduction

Five main exploration experiments, most having additional sub-experiments, were agreed upon in MPEG-135. These experiments, along with their sub-experiments, are listed below:

[EE-1]: IVDE depth-map generation

[EE-2]: Verification test preparation

- [EE-2.1]: MV-HEVC anchor generation
- [EE-2.2]: 3D-HEVC anchor generation
- [EE-2.3]: TMIV anchor with HM generation

[EE-3]: Coding and rendering of non-Lambertian content

- [EE-3.1]: Pruner improvements
- [EE-3.2]: Pruner, Packer, and Renderer improvements
- [EE-3.3]: View synthesis with RVS 4.0

[EE-4]: Multi-layer VVC coding

- [EE-4.1]: Use of interlayer prediction for texture only
- [EE-4.2]: Use of interlayer prediction for texture and depth of CG sequences

[EE-4.3]: Multilayer VVC encoding with a modified set of selected views

[EE-5]: Decoder side depth estimation

[EE-5.1]: Feature extractor threshold refinements

[EE-5.2]: Study of block-sizes, shapes, and recursive splits

[EE-5.3]: Interaction between IVDE Superpixel size and use of features

[EE-5.4]: Usage of chroma information for pixel selection

The following member organizations agreed to be part in one or more of the conducted experiments: ETRI-Immersive Media, ETRI-Media Codec, Nokia, Orange, Philips, PUT, Tencent, ULB, Interdigital

The summary in this contribution is collated from detailed reports from experimenters produced in documents listed in Table 1.

Table 1: Input document from experimenters

m57748	Tencent results for Exploration Experiments on Coding for Future MPEG Immersive Video
m57752	Interdigital results for EE2 Exploration Experiments on Future MPEG Immersive Video
m57830	ETRI-IM results for Exploration Experiments on Future MIV
m57832	Nokia results for Future MIV Exploration Experiments
m57833	Exploration Experiments on Future MIV: PUT results
m57917	Result of EE3 and EE4 for Future MIV
m57918	Orange results for EE5 Exploration Experiments on Future MPEG Immersive Video

2 Outcomes of exploration experiments

2.1 EE1: IVDE anchor depth generation

This experiment generates a MIV anchor based on depth maps obtained with IVDE 4.1 with features extracted from source textures.

Participants: ETRI-IM, PUT, Tencent

Cross-check : only minor mismatch were observed between PUT and ETRI results. The cross-check is considered as successful.

Results:

Mandatory content - Proposal vs. Low/High-bitrate Anchors

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
ClassroomVideo	A	---	727.6%	4.17	667.6%	500.1%
Museum	B	---	---	18.22	---	---
Fan	O	-71.9%	-68.9%	6.20	-46.0%	-44.4%
Kitchen	J	150.2%	80.0%	14.69	73.5%	50.5%
Painter	D	3.4%	8.0%	7.93	5.9%	9.8%
Frog	E	-10.8%	-1.8%	6.19	-1.9%	3.4%
Carpark	P	23.8%	29.3%	6.98	31.8%	33.7%
Chess	N	---	---	26.35	---	---
Group	R	---	---	22.32	---	322.9%
MIV		---	---	**12.56**	---	---

Optional content - Proposal vs. Low/High-bitrate Anchors

Fencing	L	-1.5%	25.5%	8.79	0.4%	28.4%
Hall	T	-61.0%	-49.7%	10.63	-43.4%	-41.1%
Street	U	34.2%	41.6%	8.29	39.7%	45.7%
ChessPieces	Q	---	---	32.34	---	---
Hijack	C	---	---	22.05	---	---
Mirror	I	0.3%	-1.3%	8.72	0.1%	-1.5%
Cadillac	G	41.2%	4.3%	11.84	64.3%	14.9%
MIV		---	---	**14.67**	---	---

According to ETRI-IM, the results derived with EE1 depths are better than the A17 anchor for sequence O, E and T. In case of O, doubts are expressed that the objective gain would be visible in pose traces. In case of E, the gain seems too small to replace CTC depth maps. Probably, T can be a candidate for replacement under appropriate viewing test.

According to PUT, the high quality in SO, as previously, is the result of much higher redundancy in atlases when estimated depth maps are used (more information from input views is transmitted, resulting in the increased quality of synthesized views). There are also fewer high-frequency edges in depth maps (fewer details on a fan), which decreased the bitrate of encoded geometry atlases. SE and ST are slightly better, probably the subjective difference is unnoticeable.

Recommendations:

ETRI-IM

- Based on objective result, Hall sequence may be a possible candidate for replacing CTC depth maps under proper viewing test

PUT

- EE1 should be continued to test the performance of the new TMIV 11.0 only if considerable changes will be made.

2.2 EE2.1 and EE2.3: MV-HEVC and TMIV with HM anchors generation

With a view of producing anchors for the verification tests, the goal of this experiment was to refine simulation pipeline from the previous meeting cycle and have an initial performance evaluation of using the Multi-View High Efficiency Video Codec (MV-HEVC). For this experiment,

only sequences that never used for the MIV development were evaluated. The allocation of sequence to each organization is shown in Table 2.

Table 2: Sequence allocation

Tester	Barn (Z)	Breakfast (Y)	Breaktime (K)	Dancing (W)	ChocoFountain
1	PUT	PUT	ETRI-IM	InterDigital	Philips
2	InterDigital	InterDigital	Philips	ETRI-IM	PUT

Almost whole configuration was included in EE description. Philips confirmed the configuration parameters for the use of the new RVS 4.0 for rendering in decoded source bitstreams.

MV-HEVC should not be compared with the default TMIV anchor that uses VVEnc. The EE2.3 was used to provide a meaningful comparison in which TMIV utilizes HM as an internal codec

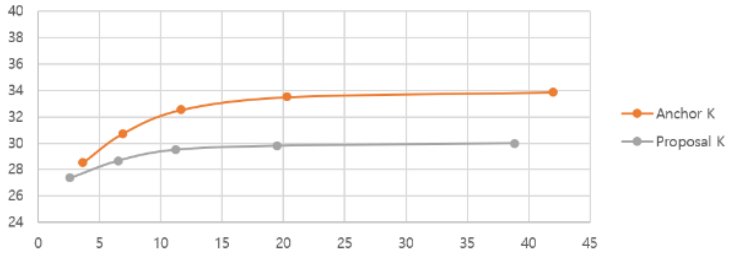
2.2.1 Cross-check

- For Barn, Breakfast, Breaktime, and Dancing the crosscheck was successful, no differences between testers were reported.
- For ChocoFountain the cross-check was not performed, as observed quality this sequence led to suggestion that it should be not considered for verification tests.

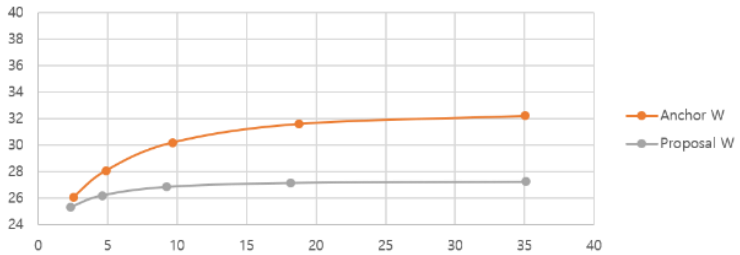
2.2.2 Results

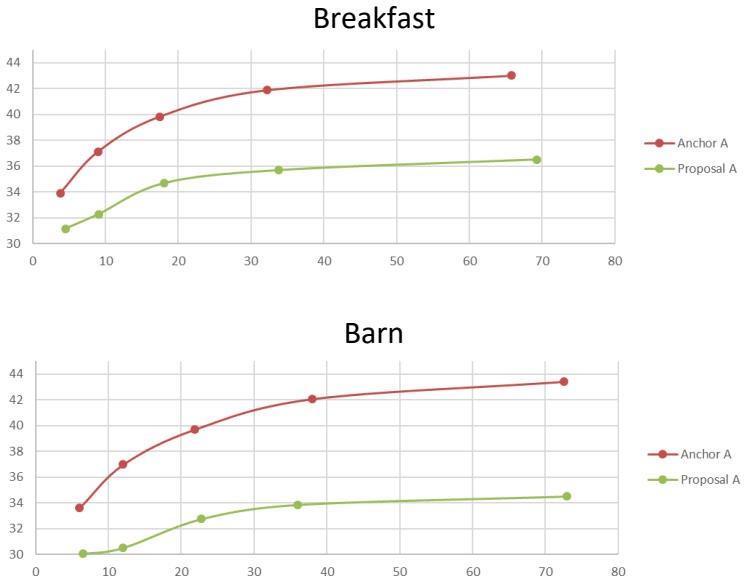
According to objective results (figures below), MV-HEVC provides worse quality than the MIV anchor (with HM – obtained in EE2.3) in A97 configuration for all tested sequences.

Breaktime



Dancing





Comments from Interdigital:

- Viewing session comments on W:
  - A first conclusion is that whatever the QP and the pose trace, images are quite blurry (cf. objects on the walls). Consequently, at high bitrate, EE2.1 pose traces are worse than EE2.3 ones, which don't suffer this artefact.
  - A second conclusion is observed at low bitrate, where jittering artefacts are visible and really annoying in EE2.3 near the dancing character. This kind of artefacts is absent from EE2.1 pose traces, which only suffer from blurriness, hence remain more stable.
- Viewing session comments on Y:
  - neither EE2.1 nor EE2.3 gives nice watchable pose traces. Indeed, some artefacts due to depth estimation are clearly visible for both, whatever the QP. Nevertheless, at very low bitrate, some patch seams and jitter near characters appear on EE2.3 which are not present in EE2.1.
- Viewing session comments on Z:
  - again, neither EE2.1 and EE2.3 are gracefully rendered, with a lot of artefacts near the contours of objects. At high bitrate, there are more artefacts near the ladder on EE2.1 than in EE2.3. At low bitrate, there is more noise on EE2.3 and some patch seams are visible, while EE2.1 is blurrier which gives a less annoying experience.

2.2.3 Recommendations

ETRI:

- For sequence K and W, the results of EE 2.3 (TMIV + HEVC) unquestionably outperforms the EE 2.1 (MV-HEVC + RVS) in the aspect of objective results.

PUT:

- Recommended Barn and Breakfast to be used for MIV verification tests.

Interdigital:

- Recommended using obtained results for verification tests.

2.3 EE2.2: 3D-HEVC anchor generation

This experiment explored the possibility of using 3D-HEVC to generate verification test anchors. In the previous EE, there was no configuration for 3D-HEVC provided by participants. In this cycle, the configuration for encoding of 7 views was acquired by PUT.

2.3.1 Cross-check

Experiment was not crosschecked during the current meeting cycle.

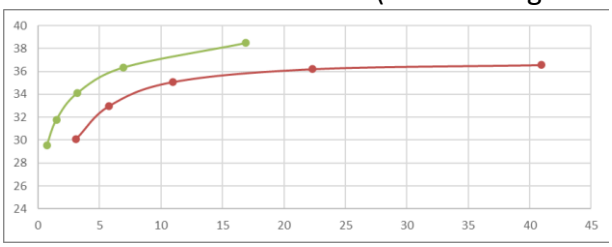
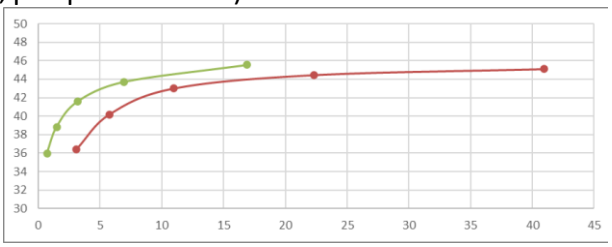
2.3.2 Results

The results for three types of sequences (linear, planar, and spherical arrangement) were acquired by PUT and shown in m57753.

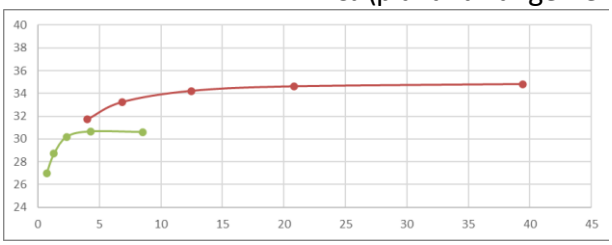
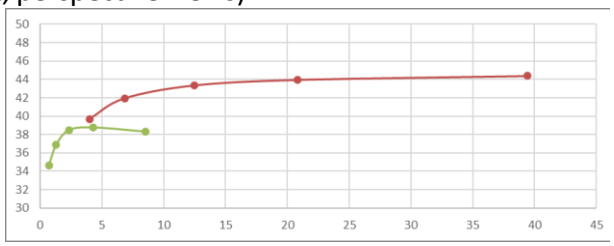
MIV (red curve) vs. 3D-HEVC (green curve) for three tested sequences:

WS-PSNR [dB] IV-PSNR [dB]

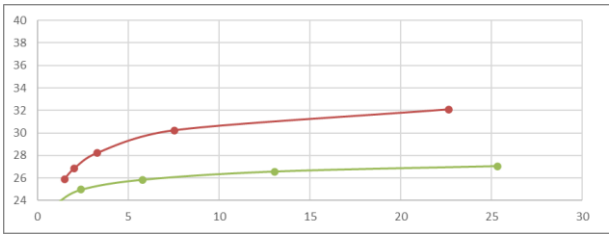
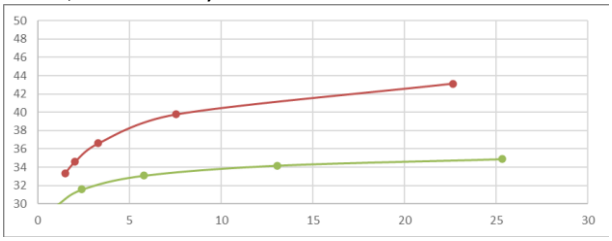
SP (linear arrangement, perspective views)

SJ (planar arrangement, perspective views)

SB (spherical arrangement, ERP views)

3D-HEVC performs well for linear multicamera systems, but it cannot provide reasonable results for other camera arrangements.

2.3.3 Recommendations

PUT:

- Recommended using provided configuration files for the crosscheck purposes,
- Recommended not using the 3D-HEVC encoder in MIV verification tests.

2.4 EE3.1: Coding and rendering of non-Lambertian content – Pruner improvements

The specular hierarchical pruner, described in m56338, was shown to provide objective quality gains. The software, was provided as branch m56338-9.0, is was based on TMIV-9.0. The goal of this experiment was to first rebase the implementation to TMIV-10.0, and then evaluate performance comparing it to the A17 anchors generated using TMIV-10.0.

Table 3: Sequence allocation

Seq	SA	SB	SO	SJ	SD	SE	SP	SN	SR
Tester	Nokia	Nokia	Nokia	Nokia	Nokia	Nokia	Nokia	Nokia	Nokia
Tester	ULB	ETRI-MC	ULB	ETRI-MC	ULB	ETRI-MC	ULB	ETRI-MC	ULB

2.4.1 Cross-check

The cross-check between ETRI-MC and Nokia has been successful.

2.4.2 Results

Mandatory content - Proposal vs. Low/High-bitrate Anchors

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]
ClassroomVideo	A	-3.4%	1.3%	0.85	5.3%	4.4%	0%	0.00	30
Museum	B	4.9%	6.6%	8.50	8.7%	8.7%	0%	0.00	30
Fan	O	6.0%	12.5%	7.05	11.5%	16.5%	0%	0.00	30
Kitchen	J	-38.7%	-11.0%	12.69	-26.2%	-1.5%	0%	0.00	30
Painter	D	8.7%	12.3%	7.65	11.7%	12.6%	0%	0.00	30
Frog	E	-10.1%	0.4%	6.48	5.0%	10.0%	0%	0.00	30
Carpark	P	4.0%	8.2%	6.57	19.4%	14.5%	0%	0.00	25
Chess	N	-5.9%	0.1%	12.93	1.5%	2.2%	0%	0.00	30
Group	R	-17.6%	-9.9%	12.72	-5.4%	-3.2%	0%	0.00	30
MIV		**-5.8%**	**2.3%**	**8.38**	**3.5%**	**7.1%**	0%	0.00	

Optional content - Proposal vs. Low/High-bitrate Anchors

Fencing	L	-70.4%	-26.4%	7.83	-14.3%	-0.7%	0%	0.00	25
Hall	T	-52.7%	-16.2%	9.19	-7.3%	4.0%	0%	0.00	25
Street	U	8.4%	12.2%	7.67	50.3%	31.2%	0%	0.00	25
ChessPieces	Q	-4.5%	-1.2%	15.62	-14.1%	-7.3%	0%	0.00	30
Hijack	C	-14.8%	-4.8%	7.77	44.8%	22.1%	0%	0.00	30
Mirror	I	-4.8%	0.2%	8.31	-1.0%	2.4%	0%	0.00	30
Cadillac	G	-19.5%	-16.2%	9.19	-15.5%	-14.3%	0%	0.00	30
MIV		**-22.6%**	**-7.5%**	**9.37**	**6.1%**	**5.3%**	0%	0.00	

2.4.3 Recommendations

ETRI-MC:

- In experiment result, the gain has been verified and It is recommended that EE3.1 is discontinued.

Nokia:

- We recommend including the Specular Hierarchical Pruner as an alternative Pruner method in the TMIV encoder as it shows benefits for many sequences that contain non-Lambertian content, so that TMIV users can test different pruning approaches.

2.5 EE3.2: Coding and rendering of non-Lambertian content – Pruner, Packer and Renderer improvements

The goal was to evaluate the performance of packing, coding, and rendering of non-Lambertian surfaces, for which multiple texture patches are encoded, without coding redundant geometry.

Table 4: Sequence allocation

Seq	I	J	Cadillac	Magritte-M	Magritte-T
Tester	Nokia	Nokia	Nokia	Nokia	ULB
Tester	ETRI-IM	ETRI-IM	ETRI-MC	ULB	Nokia

2.5.1 Cross-check

ETRI-MC and Nokia reported a successful crosscheck.

2.5.2 Results

Mandatory content - Proposal vs. Low/High-bitrate Anchors

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]
Kitchen	J	9.5%	5.6%	13.02	4.9%	4.3%	0%	0.00	30
MIV		**9.5%**	**5.6%**	**13.02**	**4.9%**	**4.3%**	0%	0.00	

Optional content - Proposal vs. Low/High-bitrate Anchors

Mirror	I	12.7%	8.1%	9.38	9.2%	6.4%	0%	0.00	30
Cadillac	G	3.0%	-2.1%	10.65	6.5%	-1.1%	0%	0.00	30
Magritte-Mirror	Y	26.6%	33.6%	9.88	23.3%	31.6%			
Magritte-Transp.	Z	26.0%	39.9%	8.06	21.4%	38.4%			
MIV		**7.9%**	**3.0%**	**10.02**	**7.9%**	**2.7%**	0%	0.00	

2.5.3 Recommendations

ETRI-MC:

- It is recommended that EE3.2 is further study if new improvement would be developed.

Nokia:

- The experiment has shown potentially encouraging results to enable bitrate reduction. However, further experimentation will be required to improve synthesis quality.

2.6 EE3.3: Coding and rendering of non-Lambertian content – View synthesis with RVS 4.0

The goal of this experiment was to evaluate the performance of RVS4.0 multi-depth, described in m57104 which provides a way to render reflections and refractions in non-Lambertian objects.

Table 5: Sequence allocation

Seq	Magritte-M	Magritte-T	Mirror	Cadillac
Tester	ULB	ULB	ULB	ULB
Tester	Nokia	Nokia	Nokia	Nokia

2.6.1 Cross-check

Nokia has confirmed the experimental results of ULB.

2.6.2 Results

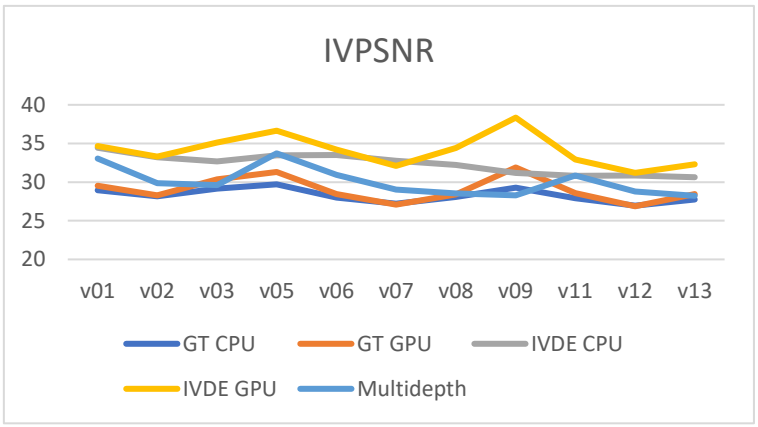


Figure 1: Mirror

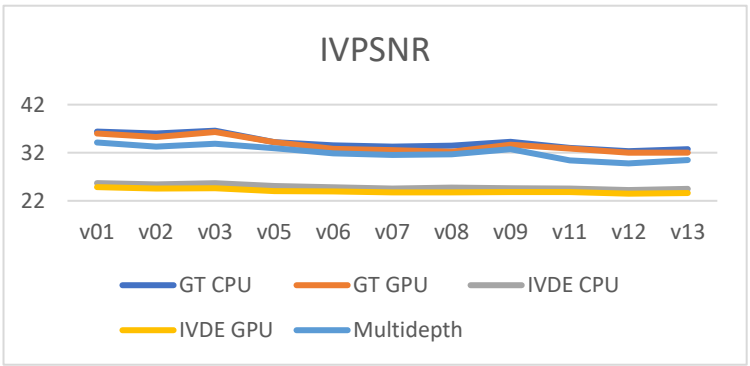


Figure 2: Cadillac

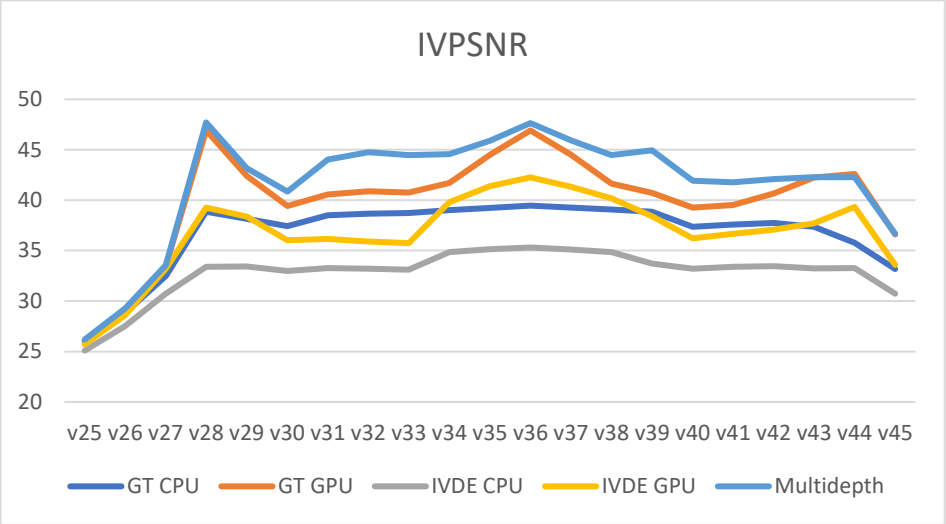


Figure 3: Magritte-M

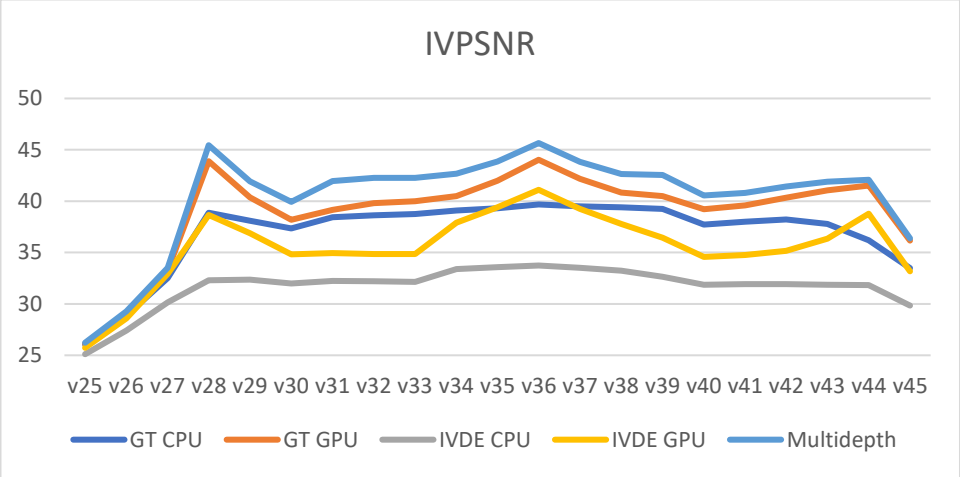


Figure 4: Magritte-T

2.6.3 Recommendations

Nokia:

- From the observed results, the proposed rendering approach shows improved visual quality. It would be interesting to also consider how compression artifacts of the views and depth maps influence the rendering results.
- We recommend continuing exploring this research topic.

2.7 EE4.1: Multi-layer VVC coding – Use of interlayer prediction for texture only

Table 6: Sequence allocation

Seq	SA	SO	SD	SE	SP	SN
Tester 1	Nokia	Nokia	Nokia	Nokia	Nokia	Nokia
Tester 2	Tencent	Tencent	Tencent	Tencent	Tencent	Tencent
Tester 3			ETRI-MC			ETRI-MC

2.7.1 Cross-check

ETRI-MC did the crosscheck for all sequences. However due to different process used by Nokia and ETRI-MC the cross-check did not match.

2.7.2 Results

Mandatory content - Proposal vs. Low/High-bitrate Anchors

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]
ClassroomVideo	A	396,5%	63,3%	4,00	-8,4%	-15,1%	0 %	0,00	30
Fan	O	-26,0%	-43,4%	12,48	-9,6%	-34,9%	#####	####	30
Painter	D	63,7%	58,4%	8,57	52,4%	51,1%	#####	####	30
Frog	E	4,5%	3,5%	7,05	-1,3%	-1,6%	#####	####	30
Carpark	P	52,6%	45,7%	10,38	21,5%	25,6%	#####	####	25
Chess	N	---	---	28,85	268,5%	61,6%	#####	####	30
MIV		**---**	**---**	**11,89**	**53,9%**	**14,5%**	#####	####	

Runtime ratio (%)

Atlas encoding	Video encoding	Decoding & Rendering
0,0%	112,3%	66,4%
0,0%	118,3%	51,2%
0,0%	138,6%	60,6%
0,0%	113,4%	57,1%
0,0%	123,1%	51,4%
0,0%	130,9%	52,9%
0,0%	122,8%	56,6%

2.7.3 Recommendations

ETRI-MC:

- MIV has been explored to outperform Multi-Layer VVC. Therefore, It is recommended that EE4 is discontinued.

Nokia:

- The experiments have shown encouraging results in terms of bitrate reduction. However, further studies will be required to improve the redundancy reduction and overall quality using inter-layer prediction.

2.8 EE4.2: Multi-layer VVC coding – Use of interlayer prediction for texture and depth of CG sequences

Table 7: Sequence allocation

Seq	SA	SO	SN
Tester 1	Nokia	Nokia	Nokia
Tester 2	KAU	KAU	ETRI-MC

2.8.1 Cross-check

ETRI-MC did the crosscheck for all sequences. However due to different process used by Nokia and ETRI-MC the cross-check did not match.

2.8.2 Results

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	394,1%	55,1%	3,99	-11,6%	-19,5%	0 %	0,00	30	0,0%	114,0%	66,5%
Fan	O	-27,6%	-45,6%	12,51	-12,0%	-37,8%	#####	####	30	0,0%	122,5%	51,4%
Chess	N	---	---	28,84	249,5%	61,1%	#####	####	30	0,0%	132,5%	53,1%
MIV		**---**	**---**	**15,11**	**75,3%**	**1,3%**	#####	####		0,0%	123,0%	57,0%

2.8.3 Recommendations

ETRI-MC:

- MIV has been explored to outperform Multi-Layer VVC. Therefore, It is recommended that EE4 is discontinued.

Nokia:

- The experiments have shown encouraging results in terms of bitrate reduction. However, further studies will be required to improve the redundancy reduction and overall quality using inter-layer prediction.

2.9 EE4.3: Multi-layer VVC coding – Multilayer VVC encoding with a modified set of selected views

Table 8: Sequence allocation

Seq	SO	SD	SE	SP	SN
Tester 1	Nokia	Nokia	Nokia	Nokia	Nokia
Tester 2	KAU	ETRI-MC	KAU	KAU	ETRI-MC

2.9.1 Cross-check

ETRI-MC did the crosscheck for all sequences. However due to different process used by Nokia and ETRI-MC the cross-check did not match.

2.9.2 Results

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
Painter	D	51,0%	45,2%	9,39	42,4%	39,7%	#####	####	30	0,0%	127,6%	60,5%
Frog	E	40,4%	7,5%	8,81	12,1%	-5,0%	#####	####	30	0,0%	75,5%	43,0%
Carpark	P	177,9%	73,0%	11,74	68,7%	35,5%	#####	####	25	0,0%	89,3%	43,8%
Chess	N	-57,3%	-37,5%	27,75	-59,2%	-49,2%	#####	####	30	0,0%	112,6%	53,0%
MIV		**53,0%**	**22,0%**	**14,42**	**16,0%**	**5,3%**	#####	####		0,0%	101,2%	50,1%

The results for O are not reported, since there was an issue during the decoding of the bitstream. This needs to be further investigated.

2.9.3 Recommendations

ETRI-MC:

- MIV has been explored to outperform Multi-Layer VVC. Therefore, it is recommended that EE4 is discontinued.

Nokia:

- The experiments have shown encouraging results in terms of bitrate reduction. However, further studies will be required to improve the redundancy reduction and overall quality using inter-layer prediction.

2.10 EE5.1: DSDE: feature extractor threshold refinement

This experiment tests different values of the feature extractor threshold (default value = 0.7) to check if better trade-off between quality and runtime can be found.

Participants: Tencent, Orange

2.10.1 Cross-check

Threshold 0.5 cross-checked by Orange for Frog and Fan: perfect match for Frog, one slight mismatch for Fan on RP5, ranging from -0.01 through -0.18 db depending on the view.

2.10.2 Results

th = 0.3

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	IVDE
Fan	O	-1.3%	-0.8%	10.82	-0.4%	-0.2%	0%	0.00	30	#DIV/0!	100.0%	109.0%
Kitchen	J	0.1%	0.0%	13.09	-0.0%	-0.0%	0%	0.00	30	#DIV/0!	100.0%	101.4%
Painter	D	-14.8%	-6.5%	5.90	-16.1%	-6.9%	0%	0.00	30	#DIV/0!	100.0%	124.5%
Frog	E	-1.5%	-0.9%	7.47	-1.4%	-0.8%	0%	0.00	30	#DIV/0!	100.0%	114.9%
Carpark	P	-6.8%	-4.4%	10.35	-2.9%	-2.3%	0%	0.00	25	#DIV/0!	100.0%	118.7%
Group	R	-2.8%	-2.4%	22.52	-1.0%	-0.7%	0%	0.00	30	#DIV/0!	100.0%	115.2%
MIV		**-4.5%**	**-2.5%**	**11.69**	**-3.6%**	**-1.8%**	0%	0.00		#DIV/0!	100.0%	114.0%

th = 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	IVDE
Fan	O	-0.8%	1.0%	10.82	-0.3%	1.1%	0%	0.00	30	#DIV/0!	100.0%	104.4%
Kitchen	J	0.1%	0.0%	13.09	0.0%	-0.0%	0%	0.00	30	#DIV/0!	100.0%	103.0%
Painter	D	-8.9%	-3.2%	6.51	-9.1%	-3.4%	0%	0.00	30	#DIV/0!	100.0%	117.0%
Frog	E	-0.8%	-0.4%	7.46	-0.7%	-0.3%	0%	0.00	30	#DIV/0!	100.0%	106.8%
Carpark	P	-4.4%	-3.1%	10.33	-2.1%	-1.9%	0%	0.00	25	#DIV/0!	100.0%	106.1%
Group	R	-0.6%	-0.7%	22.54	-0.2%	-0.2%	0%	0.00	30	#DIV/0!	100.0%	111.2%
MIV		**-2.6%**	**-1.1%**	**11.79**	**-2.1%**	**-0.8%**	0%	0.00		#DIV/0!	100.0%	108.1%

th = 0.9

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	IVDE
Fan	O	0.5%	0.4%	10.85	-0.1%	0.0%	0%	0.00	30	#DIV/0!	100.0%	100.7%
Kitchen	J	-0.0%	-0.0%	13.09	0.0%	0.0%	0%	0.00	30	#DIV/0!	100.0%	100.0%
Painter	D	4.3%	1.4%	7.30	2.8%	0.5%	0%	0.00	30	#DIV/0!	100.0%	105.8%
Frog	E	1.2%	0.5%	7.49	1.2%	0.5%	0%	0.00	30	#DIV/0!	100.0%	98.0%
Carpark	P	6.3%	5.0%	10.42	2.7%	2.6%	0%	0.00	25	#DIV/0!	100.0%	98.2%
Group	R	0.5%	0.3%	22.52	0.1%	-0.0%	0%	0.00	30	#DIV/0!	100.0%	101.4%
MIV		**2.1%**	**1.3%**	**11.94**	**1.1%**	**0.6%**	0%	0.00		#DIV/0!	100.0%	100.7%

th = 1.1

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	IVDE
Fan	O	1.2%	0.7%	10.85	0.3%	0.1%	0%	0.00	30	#DIV/0!	100.0%	94.4%
Kitchen	J	0.0%	0.0%	13.09	0.1%	0.0%	0%	0.00	30	#DIV/0!	100.0%	99.4%
Painter	D	7.0%	2.5%	7.52	4.8%	1.3%	0%	0.00	30	#DIV/0!	100.0%	104.2%
Frog	E	2.5%	1.2%	7.53	2.2%	1.0%	0%	0.00	30	#DIV/0!	100.0%	95.4%
Carpark	P	17.4%	11.9%	10.53	8.6%	6.7%	0%	0.00	25	#DIV/0!	100.0%	97.6%
Group	R	1.1%	1.1%	22.51	0.1%	0.1%	0%	0.00	30	#DIV/0!	100.0%	98.8%
MIV		**4.9%**	**2.9%**	**12.01**	**2.7%**	**1.5%**	0%	0.00		#DIV/0!	100.0%	98.3%

th = 1.3

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	IVDE
Fan	O	2.2%	1.5%	10.86	1.1%	0.7%	0%	0.00	30	#DIV/0!	100.0%	95.3%
Kitchen	J	0.1%	0.0%	13.09	0.1%	0.1%	0%	0.00	30	#DIV/0!	100.0%	97.7%
Painter	D	9.5%	4.5%	7.69	6.3%	2.8%	0%	0.00	30	#DIV/0!	100.0%	99.5%
Frog	E	3.3%	1.9%	7.56	2.9%	1.7%	0%	0.00	30	#DIV/0!	100.0%	91.7%
Carpark	P	30.0%	20.9%	10.57	16.5%	12.9%	0%	0.00	25	#DIV/0!	100.0%	88.7%
Group	R	2.2%	1.8%	22.51	0.5%	0.3%	0%	0.00	30	#DIV/0!	100.0%	101.7%
MIV		**7.9%**	**5.1%**	**12.05**	**4.6%**	**3.1%**	0%	0.00		#DIV/0!	100.0%	95.8%

Thresholds 0.9, 1.1, 1.3 have high coding losses. The threshold 0.3 has a high runtime increase. The threshold 0.5 increases the coding efficiency with limited runtime increase.

2.10.3 Recommendations

Tencent

- The threshold 0.5 increases the coding efficiency with limited runtime increase. We recommend adopting this threshold.

2.11 EE5.2: DSDE study of block sizes, shapes and recursive split

Participants: PUT, Orange, Tencent

Test 1: studying 128x128 sized blocks with 64x64 minimum size, no recursive splitting.

2.11.1 Cross-check

Results supplied by Orange, PUT performed a successful partial check.

2.11.2 Results

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
Fan	O	-8.3%	-6.0%	10.54	-9.5%	-6.2%	0%	0.00	30	98.4%	100.9%	167.0%
Kitchen	J	-0.1%	0.1%	13.05	-0.2%	-0.0%	0%	0.00	30	101.4%	110.7%	117.0%
Painter	D	-11.7%	-4.5%	5.91	-11.5%	-4.2%	0%	0.00	30	144.6%	100.0%	169.1%
Frog	E	-1.7%	-1.0%	7.46	-1.4%	-0.8%	0%	0.00	30	95.0%	79.6%	118.5%
Carpark	P	-7.6%	-4.2%	10.34	-4.1%	-2.6%	0%	0.00	25	72.8%	91.4%	183.7%
Group	R	-6.3%	-6.3%	22.54	-3.3%	-3.0%	0%	0.00	30	99.8%	117.1%	144.3%
MIV		**-6.0%**	**-3.6%**	**11.64**	**-5.0%**	**-2.8%**	0%	0.00		102.0%	100.0%	150.0%

The quality is sometimes improved, but the decoding + rendering time has consistently increased.

Test 2: the experiment tests the possible reduction of decoding time using a different feature extractor configuration (Block size 32x32: initial grid size: 32x32,min size of the block: 16, no recursive splitting).

2.11.3 Cross-check

Results supplied by Tencent, PUT performed a partial check. The partial check is not successful.

2.11.4 Results

		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	IVDE
Fan	O	-5.1%	-3.2%	10.62	-5.2%	-2.8%	0%	0.00	30	#DIV/0!	99.3%	109.0%
Kitchen	J	-0.5%	-0.0%	13.07	-0.6%	-0.1%	0%	0.00	30	#DIV/0!	98.8%	101.4%
Painter	D	-15.9%	-7.1%	5.99	-16.3%	-7.1%	0%	0.00	30	#DIV/0!	99.0%	124.5%
Frog	E	-1.5%	-1.0%	7.47	-1.2%	-0.8%	0%	0.00	30	#DIV/0!	99.0%	114.9%
Carpark	P	-5.8%	-3.8%	10.40	-2.1%	-1.6%	0%	0.00	25	#DIV/0!	99.3%	118.7%
Group	R	-2.2%	-4.2%	22.55	-0.7%	-1.7%	0%	0.00	30	#DIV/0!	98.9%	115.2%
MIV		**-5.2%**	**-3.2%**	**11.68**	**-4.3%**	**-2.3%**	0%	0.00		#DIV/0!	99.1%	114.0%

Test 3: the experiment tests the possible reduction of decoding time using a different feature extractor configuration (Recursive splitting: initial grid size: 128x128, min size of the block: 16, allowance of square splitting, symmetrical rectangular, and asymmetrical rectangular splitting).

2.11.5 Cross-check

Results supplied by PUT, Tencent performed a partial check. The partial check is not successful.

2.11.6 Results

Mandatory content - Proposal vs. Low/High-bitrate Anchors

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	Runtime ratio (%) Atlas encoding	Video encoding	Decoding & Rendering
Fan	O	-1.9%	-1.5%	10.83	-0.6%	-0.7%	100.0%	100.0%	133.7%
Kitchen	J	-0.3%	0.0%	13.05	-0.2%	-0.0%	100.0%	100.0%	149.6%
Painter	D	-5.6%	-1.1%	6.38	-5.4%	-1.1%	100.0%	100.0%	234.7%
Frog	E	-1.3%	-0.9%	7.44	-1.3%	-0.8%	100.0%	100.0%	160.5%
Carpark	P	-0.1%	-0.7%	10.35	1.6%	0.4%	100.0%	100.0%	191.2%
Group	R	-6.0%	-6.1%	22.52	-3.0%	-2.5%	100.0%	100.0%	163.8%
MIV		**-2.5%**	**-1.7%**	**11.76**	**-1.5%**	**-0.8%**	100.0%	100.0%	172.2%

Optional content - Proposal vs. Low/High-bitrate Anchors

Fencing	L	-0.5%	-0.1%	13.19	1.8%	1.1%	100.0%	100.0%	206.2%
Hall	T	-19.7%	-87.6%	17.26	-3.9%	-65.8%	100.0%	100.0%	300.2%
Street	U	-5.5%	-3.0%	6.99	-2.6%	-1.3%	100.0%	100.0%	240.3%
ChessPieces	Q	-13.5%	-12.8%	26.86	-6.1%	-3.8%	100.0%	100.0%	134.7%
Hijack	C	-2.8%	3.4%	21.50	2.4%	6.0%	100.0%	100.0%	137.0%
Mirror	I	0.0%	0.1%	12.63	-0.2%	0.1%	100.0%	100.0%	168.1%
Cadillac	G	-1.0%	-0.8%	14.28	-0.6%	-0.6%	100.0%	100.0%	171.7%
ClassroomVideo	A	-6.0%	-7.9%	5.80	-1.4%	-2.6%	100.0%	100.0%	399.2%
Museum	B	-3.1%	-3.7%	9.16	-0.9%	-1.6%	100.0%	100.0%	133.4%
Chess	N	-76.9%	-28.7%	23.69	-5.2%	-4.5%	100.0%	100.0%	127.9%
MIV		**-12.9%**	**-14.1%**	**15.14**	**-1.7%**	**-7.3%**	100.0%	100.0%	201.9%

2.11.7 Recommendations

PUT

- We recommend using 64x64 block without splitting for skip flag derived from the decoded textures (no change to CTC).

Tencent

- Investigate further test 2.

2.12 EE5.3: DSDE interaction between super pixel size and usage of features

The experiment tests the performance of DSDE anchor when 150000 superpixels per view are used in IVDE (default value is 100000). Using 150000 segments per view means that the superpixels will be smaller. The rest of configuration of IVDE and TMIV is the same as in DSDE anchor (G17).

Participants: Tencent, PUT

2.12.1 Cross-check

Results supplied by Tencent, PUT performed a partial check. The partial check is successful.

2.12.2 Results

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	IVDE
Fan	O	-0.0%	-0.4%	10.76	-2.2%	-1.4%	0%	0.00	30	#DIV/0!	100.0%	173.1%
Kitchen	J	-2.4%	-2.5%	12.99	0.4%	-1.5%	0%	0.00	30	#DIV/0!	100.0%	180.8%
Painter	D	4.7%	1.4%	7.63	5.9%	2.2%	0%	0.00	30	#DIV/0!	100.0%	181.3%
Frog	E	0.3%	0.2%	7.45	0.0%	0.2%	0%	0.00	30	#DIV/0!	100.0%	185.7%
Carpark	P	-4.6%	-2.4%	10.21	-0.0%	0.8%	0%	0.00	25	#DIV/0!	100.0%	222.8%
Group	R	13.9%	5.9%	22.49	7.7%	-0.8%	0%	0.00	30	#DIV/0!	100.0%	182.0%
MIV		**2.0%**	**0.4%**	**11.92**	**2.0%**	**-0.1%**	0%	0.00		#DIV/0!	100.0%	187.6%

2.12.3 Recommendations

Tencent:

- The anchor configuration with 100000 superpixels provides better results than the tested one with 150000 superpixels. No CTC change.

PUT:

- We recommend keeping the default value of 100000 superpixels.

2.13 EE5.4: DSDE usage of chroma for pixels selection

This experiment tests the usage of chroma values as well as luma, when deciding whether or not to 'skip' a block.

Participants: Orange, Tencent

2.13.1 Cross-check

Results supplied by Orange, Tencent performed a partial check. The partial check is successful.

2.13.2 Results

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
Fan	O	0.9%	0.6%	10.85	0.3%	0.1%	0%	0.00	30	97.3%	98.4%	103.3%
Kitchen	J	0.0%	0.0%	13.09	0.0%	0.0%	0%	0.00	30	99.9%	118.8%	86.8%
Painter	D	4.5%	1.4%	7.33	3.0%	0.5%	0%	0.00	30	99.4%	99.6%	92.5%
Frog	E	1.6%	0.7%	7.51	1.4%	0.6%	0%	0.00	30	99.3%	77.9%	92.3%
Carpark	P	14.1%	9.5%	10.48	7.6%	5.5%	0%	0.00	25	75.0%	96.8%	99.2%
Group	R	0.8%	0.6%	22.52	0.1%	0.0%	0%	0.00	30	99.5%	111.7%	105.0%
MIV		**3.6%**	**2.1%**	**11.96**	**2.1%**	**1.1%**	0%	0.00		95.1%	100.5%	96.5%

2.13.3 Recommendations

Orange:

- Chroma inclusion does not provide a benefit at least with current parameters.