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 m57833
October 2021, Online

Title: **Exploration Experiments on Future MIV: PUT results**

Source: **Dawid Mieloch, Adrian Dziembowski**
(Poznań University of Technology)

Abstract & Recommendations

The document presents the results of exploration experiments that were conducted by PUT. The results include a description of conducted experiments and their results or crosschecks. The recommendations are:

- EE1 should be continued to test the performance of the new TMIV 11.0 only if considerable changes will be made,
- (EE2.1) Barn and Breakfast sequences are recommended to be used for MIV verification tests,
- (EE5.2) We recommend using 64x64 block without splitting for skip flag derived from the decoded textures (no change to CTC).
- (EE5.3) We recommend keeping the default value of 100000 superpixels.

1 Introduction

The document presents the results of EE-related experiments that were conducted by PUT. The results include full results for EE1, EE2.2, and EE5.2 (part III), partial results for EE2.1, EE5.2 (parts I and II), and EE5.3.

2 Experiments

For EE1, GCC 10.2.0 was used to compile all used software, while for EE2.1, VS10 was used.

2.1 EE1

The experiment tested the performance of depth maps estimated by IVDE in comparison with the current CTC depth maps. The table below compares the performance of the A17 anchor against the new depth maps (estimated at the TMIV encoder side).

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**	**---**	**---**

Comments:

- As expected, the quality of depth maps generated in the experiment is lower than for CTC depth maps. The depth maps in this experiment are generated using the same estimation parameters for all sequences, while for CTC depth maps (even if they were generated earlier using IVDE), the parameters were fine-tuned to give the best possible quality.
- 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.

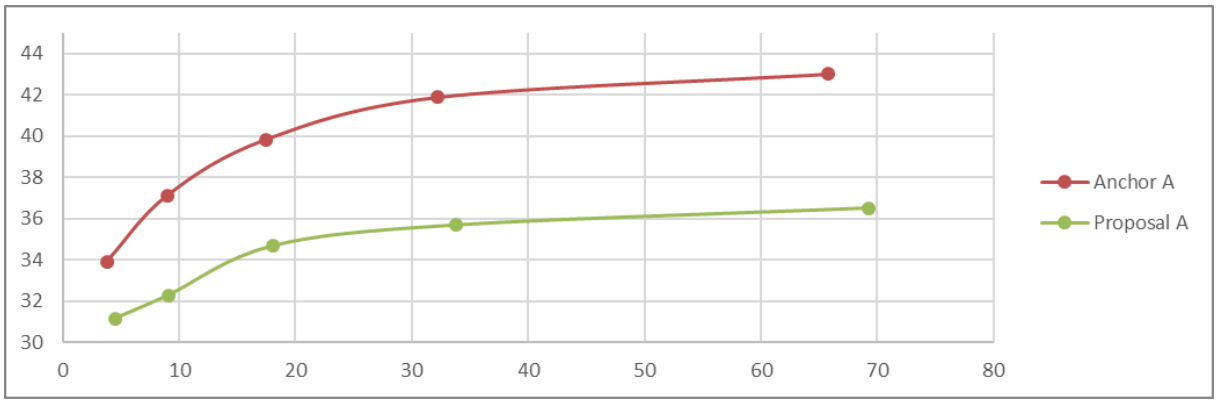
Recommendations:

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

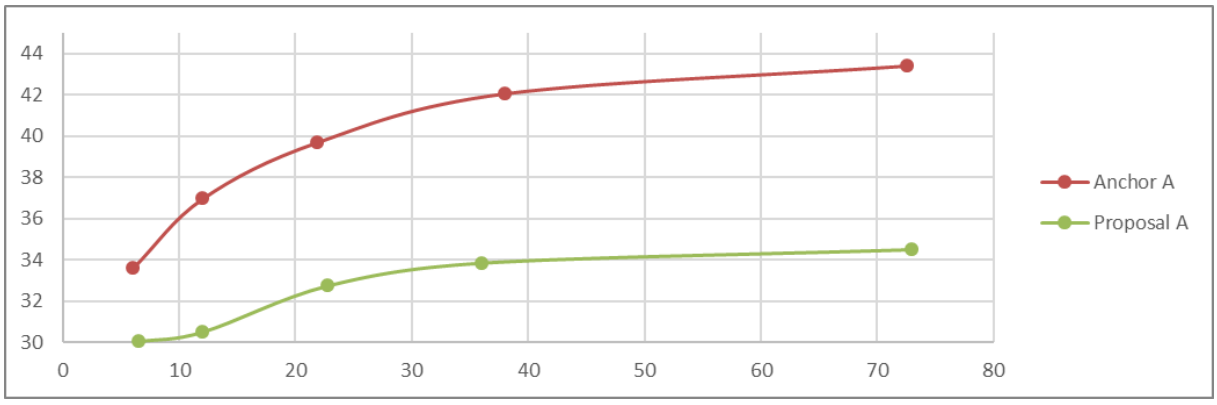
2.2 EE2.1

The experiment tested the performance of MV-HEVC for 97 frames. PUT was testing Barn and Breakfast sequences. The comparison of MIV Anchor that uses HEVC (Anchor, provided by Interdigital in EE2.3) and MV-HEVC (Proposal) is shown below.

Breakfast



Barn



Comments:

- MV-HEVC is showing worse performance than TMIV for both sequences (objectively).

Recommendations:

- Barn and Breakfast sequences are recommended to be used for MIV verification tests.

2.3 EE2.2

This experiment explored the possibility of using the 3D-HEVC to generate verification test anchors. The results and their summary were provided in m57753.

2.4 EE5.2

The experiment was testing the possible reduction of decoding time using different initial grid size, and the recursive block splitting.

During testing, we saw that in the G17 anchor the size of blocks coming from the extractor is always 64 (they are newer divided into smaller ones). The reason for that is the value of *partitioning_threshold* which was set to 1024. Unfortunately, such value was set for 16bps input, so it should be equal to 16, as our input is 10bps.

Therefore, we proposed to change *partitioning_threshold* to 16 in EE5.2. Without this change, this experiment does not make any sense, as blocks are never split, so nothing would have been tested.

To show the impact of changing *partitioning_threshold* to 16 in normal CTC conditions, we also include such a comparison below (CTC G17 configuration with changed partitioning).

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	-2.2%	-1.8%	10.73	-1.4%	-1.4%	100.0%	100.0%	187.7%
Kitchen	J	-0.2%	0.1%	13.07	-0.6%	-0.1%	100.0%	100.0%	120.3%
Painter	D	-4.0%	-2.0%	6.12	-2.5%	-1.3%	100.0%	100.0%	210.7%
Frog	E	-1.5%	-0.9%	7.47	-1.4%	-0.9%	100.0%	100.0%	172.4%
Carpark	P	-4.7%	-3.2%	10.35	-1.8%	-1.6%	100.0%	100.0%	217.6%
Group	R	-7.2%	-6.4%	22.52	-3.6%	-2.7%	100.0%	100.0%	177.2%
MIV		**-3.3%**	**-2.4%**	**11.71**	**-1.9%**	**-1.3%**	100.0%	100.0%	181.0%

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

Fencing	L	2.8%	0.0%	13.19	3.4%	1.9%	100.0%	100.0%	265.9%
Hall	T	-27.5%	-90.3%	17.24	1.3%	-51.4%	100.0%	100.0%	369.4%
Street	U	-5.2%	-2.2%	6.99	-2.2%	-0.9%	100.0%	100.0%	260.5%
ChessPieces	Q	-10.1%	-19.5%	26.76	-1.8%	-4.4%	100.0%	100.0%	135.0%
Hijack	C	-6.0%	-2.4%	21.55	-1.6%	-1.1%	100.0%	100.0%	127.2%
Mirror	I	-0.6%	-0.4%	12.63	-0.7%	-0.2%	100.0%	100.0%	171.8%
Cadillac	G	-1.8%	-1.3%	14.25	-1.6%	-1.2%	100.0%	100.0%	174.9%
ClassroomVideo	A	-3.0%	-6.9%	5.81	-0.5%	-2.2%	100.0%	100.0%	394.5%
Museum	B	-2.1%	-3.1%	9.17	-0.2%	-1.2%	100.0%	100.0%	131.1%
Chess	N	-98.1%	-36.8%	23.50	-3.8%	-2.1%	100.0%	100.0%	128.2%
MIV		**-15.2%**	**-16.3%**	**15.11**	**-0.8%**	**-6.3%**	100.0%	100.0%	215.9%

2.4.1 Part I

Block size 128x128: initial grid size: 128x128, min size of the block: 64

Comments:

- PUT successfully cross-checked SJ and SP.
- Results from Orange indicate constant decrease of BD-rate over mandatory sequences with 50% increase in the estimation time.

2.4.2 Part II

Block size 32x32: initial grid size: 32x32, min size of the block: 16

Comments:

- PUT was cross-checking SO and SR, at the time of preparing this document the cross-check was not successful.

2.4.3 Part III

Recursive splitting: initial grid size: 128x128, min size of the block: 16 (enabled recursive splitting), allowance of square splitting, symmetrical rectangular, and asymmetrical rectangular splitting

PUT was providing full results for all sequences.

Comparison of G17 anchor and EE5.3 (part III)

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%

Comparison of G17 anchor with changed partitioning and EE5.3 (part III)

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	0.3%	0.4%	10.83	0.8%	0.7%	100.0%	100.0%	71.2%
Kitchen	J	-0.1%	-0.0%	13.05	0.4%	0.1%	100.0%	100.0%	124.3%
Painter	D	-1.4%	0.8%	6.38	-2.8%	0.2%	100.0%	100.0%	111.4%
Frog	E	0.2%	0.0%	7.44	0.1%	0.0%	100.0%	100.0%	93.1%
Carpark	P	4.8%	2.6%	10.35	3.5%	2.0%	100.0%	100.0%	87.9%
Group	R	1.0%	0.3%	22.52	0.5%	0.2%	100.0%	100.0%	92.5%
MIV		**0.8%**	**0.7%**	**11.76**	**0.4%**	**0.5%**	100.0%	100.0%	96.7%

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

Fencing	L	-3.2%	-0.2%	13.19	-1.6%	-0.8%	100.0%	100.0%	77.5%
Hall	T	5.4%	-31.5%	17.26	-3.5%	-31.6%	100.0%	100.0%	81.2%
Street	U	-0.2%	-0.8%	6.99	-0.4%	-0.4%	100.0%	100.0%	92.2%
ChessPieces	Q	-5.5%	6.1%	26.86	-3.5%	0.4%	100.0%	100.0%	99.7%
Hijack	C	1.5%	5.9%	21.50	3.2%	7.3%	100.0%	100.0%	107.7%
Mirror	I	0.7%	0.5%	12.63	0.5%	0.3%	100.0%	100.0%	97.9%
Cadillac	G	0.8%	0.5%	14.28	1.0%	0.5%	100.0%	100.0%	98.2%
ClassroomVideo	A	-3.3%	-1.0%	5.80	-0.9%	-0.4%	100.0%	100.0%	101.2%
Museum	B	-1.0%	-0.6%	9.16	-0.7%	-0.4%	100.0%	100.0%	101.8%
Chess	N	2145.2%	10.3%	23.69	-0.7%	-2.3%	100.0%	100.0%	99.7%
MIV		**214.0%**	**-1.1%**	**15.14**	**-0.7%**	**-2.7%**	100.0%	100.0%	95.7%

Comments:

- When compared to the current G17 anchor, the recursive splitting provides better objective quality in 8 sequences.
- The runtime of the estimator is almost twice longer.

Recommendations for all parts of EE5.2:

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

2.5 EE5.3

The experiment tested the performance of DSDE anchor when 150000 superpixels per view are used in IVDE (default value is 100000).

Comments:

- PUT successfully cross-checked SD and SJ.
- Results indicate that the default value provides similar quality and depth estimation is much faster.

Recommendations:

- We recommend keeping the default value of 100000 superpixels.

Acknowledgement

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