INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11 MPEG2019/M48090
July 2019, Göteborg, Sweden

Source	**Poznań University of Technology (PUT), Poznań, Poland** **Electronics and Telecommunications Research Institute (ETRI), Daejeon, Republic of Korea**
Status	**Input**
Title	**PUT/ETRI Response to Immersive Video CE-2: Temporal and spatial aggregation**
Author	Dawid Mieloch*, Adrian Dziembowski*, Olgierd Stankiewicz*, Marek Domański*, Gwangsoon Lee** * – Poznań University of Technology, Poland ** – Electronics and Telecommunications Research Institute (ETRI), Korea

1 Introduction

This document presents a technical description of PUT/ETRI experiment on the pruning, view selection, and temporal consistency of patches (Immersive Video CE-2 [1]).

The idea of the experiments is to study these issues from the point of view of diverse techniques that may be merged into Test Model for Immersive Video [2]. In particular, the work was inspired by the response of Poznań University of Technology and Electronics and Telecommunications Research Institute (ETRI) to the Call for Proposals on 3DoF+ Visual" [3].

2 Overview of the proposed technique

There are many edges between patches in an atlas, what result in different colors of patches and the existence of empty spaces in atlases. If a respective edge is horizontal/vertical and is located on a boundary of the Coding Units (CUs), it may be encoded more efficiently than, if it has random shape and it is located inside the CU. Therefore, we try to reduce the number of edges inside CUs by enlarging patches into the CU grid.

In the proposal, patches in atlases are not enlarged to bounding boxes, as in TMIV, but only to a 64x64 grid. The resulting shape of patches is not changed through the whole intra period.

Described enlargement is performed only in texture atlases, in depth atlases the patches are left in their original shape (without enlargement to bounding box), as depth maps tend to be more smooth than corresponding views.

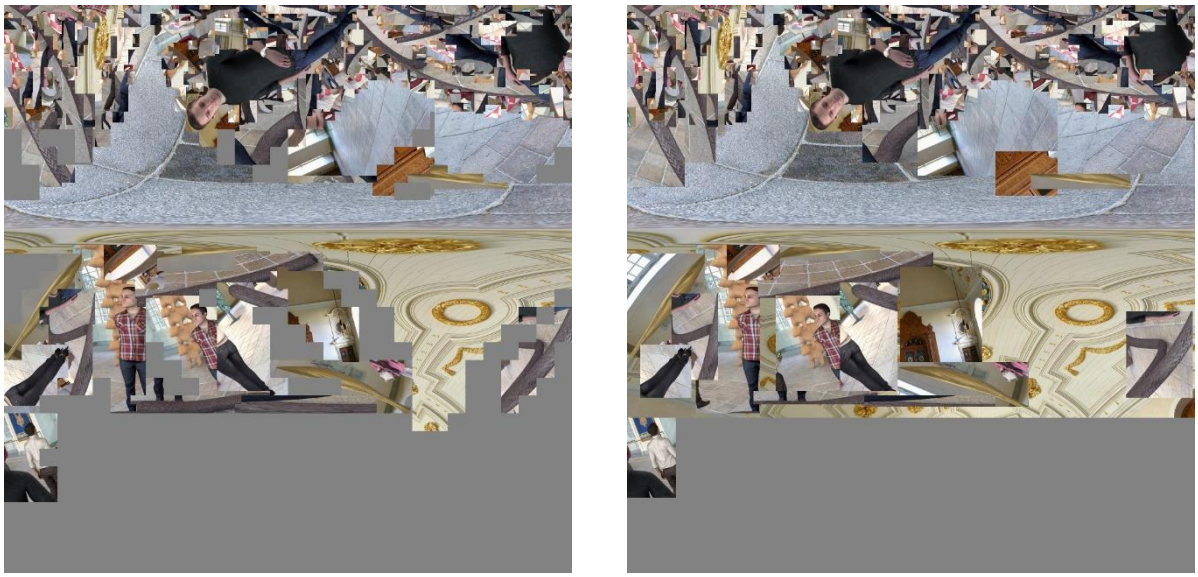

Fig. 1. The atlas from TechnicolorMuseum sequence using
the proposed enlargement of patches (left) and using the anchor (right).

3 Experimental results

Test class	Sequence	High-bitrate BD rate Y-WSPSNR	Low-bitrate BD rate Y-WSPSNR	High-bitrate BD rate VMAF	Low-bitrate BD rate VMAF	High-bitrate BD rate MM-SSIM	Low-bitrate BD rate MS-SSIM	Pixel rate ratio
CG1	ClassroomVideo	0.7%	-2.4%	-1.0%	-3.2%	0.5%	-2.3%	0.00%
	TechnicolorMuseum	4.2%	6.8%	3.7%	5.2%	4.2%	3.5%	0.00%
	TechnicolorHijack	32.3%	25.4%	11.6%	7.2%	43.9%	31.2%	0.00%
		12.4%	9.9%	4.8%	3.1%	16.2%	10.8%	0.00%
NC1	TechnicolorPainter	17.3%	13.5%	13.3%	7.1%	12.0%	8.7%	0.00%
	IntelFrog	279.3%	78.6%	20.2%	5.3%	41.7%	15.0%	0.00%
		148.3%	46.0%	16.8%	6.2%	26.9%	11.8%	0.00%
All		66.8%	24.4%	9.6%	4.3%	20.5%	11.2%	0.00%

The overall quality of final synthesized views was lower than in TMIV. However, the BD rate of encoded atlases was significantly higher. The proposal is providing almost the same quality of synthesis when no HEVC encoding is used. Therefore, we presume that the encoding of depth atlases caused errors in view synthesis (such as ghost edges). We plan to continue the research on the CE2 in order to test this presumptions.

4 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).

We would like to thank Bin Wang from ZJU for thorough crosscheck of the presented results.

5 Recommendations

As it was stated in section 3, we recommend to continue the Core Experiment 2.

6 References

[1] V. Vadakital, “Description of Immersive Video Core Experiment 2”, ISO/IEC JTC1/SC29/WG11 MPEG/ N18466, Mar. 2019, Geneva, Switzerland.

[2] "Call for Proposals on 3DoF+ Visual" ISO/IEC JTC1/SC29/WG11 MPEG/N18145, January 2019, Marrakesh, Morocco.

[2] M. Domański, A. Dziembowski, D. Mieloch, O. Stankiewicz, J. Stankowski, A. Grzelka, G. Lee, J. Seo, “Technical description of proposal for Call for Proposals on 3DoF+ Visual prepared by Poznań University of Technology (PUT) and Electronics and Telecommunications Research Institute (ETRI)”, ISO/IEC JTC1/SC29/WG11 MPEG/ M47407, Mar. 2019, Geneva, Switzerland.