INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG4
MPEG VIDEO CODING

ISO/IEC JTC1/SC29/WG4 MPEG/M54967
June 2020, Online

Source	**Poznań University of Technology (PUT), Poznań, Poland; Electronics and Telecommunications Research Institute (ETRI), Daejeon, Republic of Korea**
Status	**Input**
Title	**MIV CE1.2-related: Geometry scaling**
Authors	Dawid Mieloch*, Adrian Dziembowski*, Jakub Stankowski*, Marek Domański*, Gwangsoon Lee**, Jun Young Jeong** * – Poznań University of Technology ** – Electronics and Telecommunications Research Institute

1 Introduction

This document presents a technical description of the PUT/ETRI experiment on geometry upscaling (MPEG Immersive Video CE1.2). In the proposed solution, decoder-side geometry atlas upscaling uses superpixel-based segmentation.

2 Proposed technique

In the proposed method the depth map is refined using superpixel segmentation [1] that is performed on atlas. The segmentation is performed only for occupied pixels in an atlas.

The idea of the proposal is based on the high probability depth values inside very small segments. Segments are well-fitted to edges of objects, making them a potentially good tool for the removal of errors caused both by downscaling and encoding of geometry.

We look for outliers in depth maps for all segments, i.e., depths which do not fit into the range between 3 standard deviations from mean depth in a segment. Such depth values are corrected by changing them to the nearest value from the mentioned range.

Fig. 1. The comparison of the fragment of the atlas in the anchor (left), the corresponding fragment of the upsampled depth map (middle), and the proposal (right).

3 Experimental results

Table 1. Objective quality evaluation.

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 VMAF	Low-BR BD rate VMAF	High-BR BD rate IV-PSNR	Low-BR BD rate IV-PSNR	Pixel rate [%]	Pixel rate [GP/s]	Frame rate [Hz]	TMIV encoding	HM encoding	HM decoding	TMIV decoding
ClassroomVideo	SA	1.8%	0.7%	1.97	-3.2%	-1.2%	-0.1%	-0.1%	63%	0.67	30	100.0%	100.0%	#####	131.2%
Museum	SB	-1.4%	-0.7%	16.68	-0.7%	-0.4%	0.0%	0.0%	63%	0.67	30	100.0%	100.0%	#####	147.7%
Hijack	SC	0.5%	0.4%	9.63	0.2%	0.2%	-0.0%	-0.1%	63%	0.67	30	100.0%	100.0%	#####	133.5%
Chess	SN	-0.4%	0.3%	16.68	-3.0%	-1.1%	-0.1%	-0.1%	63%	0.67	30	100.0%	100.0%	#####	201.4%
Kitchen	SJ	3.0%	1.2%	16.89	3.0%	1.0%	0.1%	-0.1%	62%	0.67	30	100.0%	100.0%	#####	200.3%
Painter	SD	-0.2%	-0.1%	8.16	-0.3%	-0.1%	-0.1%	-0.1%	63%	0.67	30	100.0%	100.0%	#####	213.7%
Frog	SE	0.4%	0.2%	5.96	0.1%	0.0%	-0.2%	-0.1%	62%	0.67	30	100.0%	100.0%	#####	180.4%
Carpark	SP	2.2%	0.8%	7.40	1.2%	0.4%	-0.1%	-0.1%	52%	0.56	25	100.0%	100.0%	#####	199.0%
MIV		**0.7%**	**0.4%**	**10.42**	**-0.3%**	**-0.1%**	**-0.1%**	**-0.1%**	61%	0.65		100.0%	100.0%	#####	175.9%
Optional content - Proposal vs. Low/High-bitrate Anchors															
Fencing	SL	-0.9%	-0.4%	12.90	-0.3%	-0.1%	-0.6%	-0.3%	52%	0.56	25	100.0%	100.0%	#####	222.3%
Hall	ST	6.1%	2.7%	12.77	2.8%	1.3%	0.1%	-0.2%	52%	0.56	25	100.0%	100.0%	#####	234.7%
Street	SU	2.0%	0.8%	10.57	1.3%	0.5%	-0.1%	-0.0%	52%	0.56	25	100.0%	100.0%	#####	189.1%
Group	SR	-5.7%	-3.6%	11.91	-5.8%	-3.2%	-3.6%	-2.3%	62%	0.67	30	100.0%	100.0%	#####	211.0%
Fan	SO	0.1%	-0.0%	8.82	0.3%	0.2%	-0.1%	0.0%	62%	0.67	30	100.0%	100.0%	#####	199.5%
MIV		**0.3%**	**-0.1%**	**11.40**	**-0.3%**	**-0.3%**	**-0.9%**	**-0.6%**	56%	0.60		100.0%	100.0%	#####	211.3%

The differences are not high, but the proposal achieves a steady increase in the quality for VMAF and IV-PSNR. As can be seen in the example shown below, differences are not large, but the final synthesized image has a smaller amount of errors.

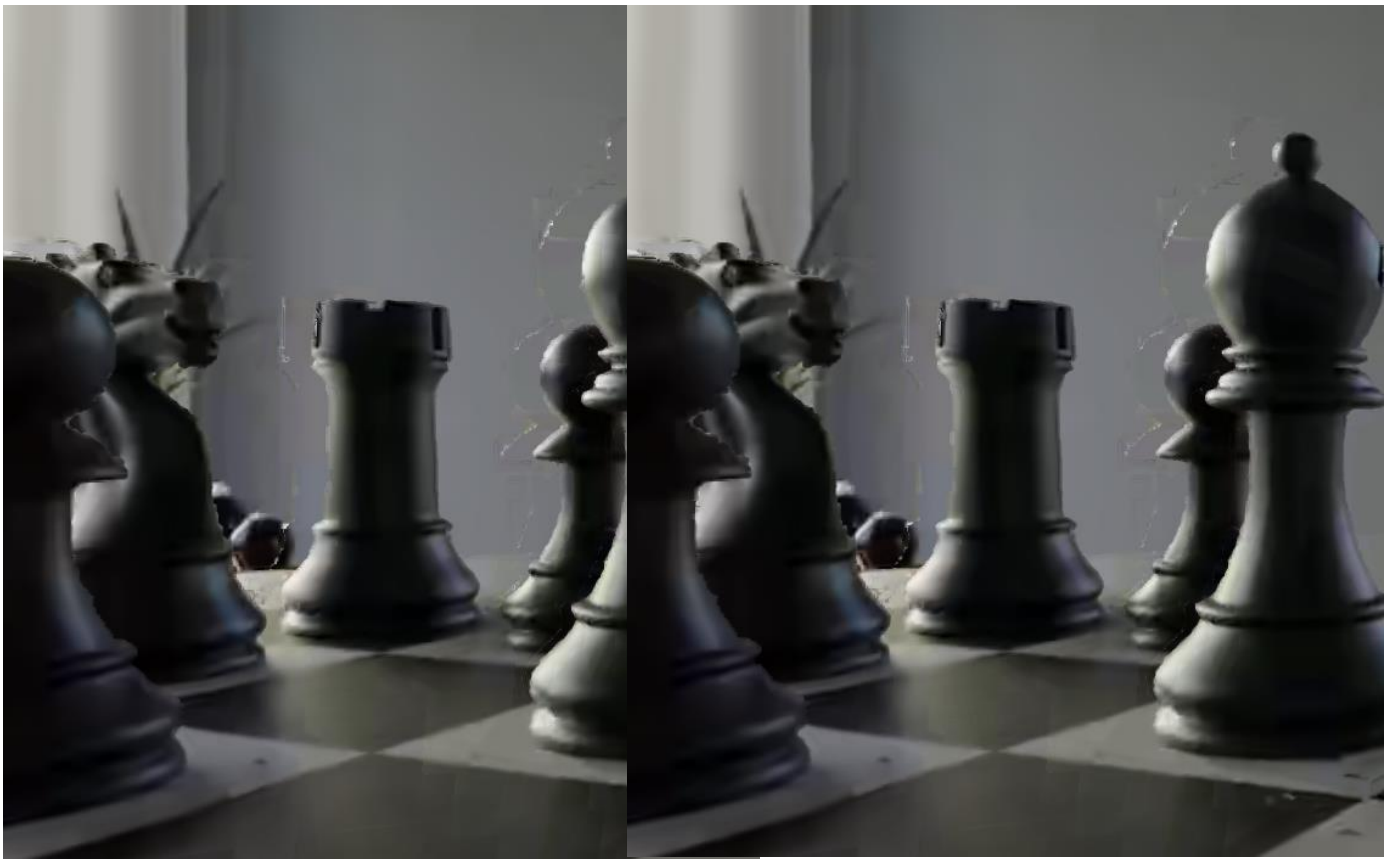

Fig. 2. The comparison of the current anchor (left) with the proposal (right).

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).

5 Recommendations

We recommend to continue the works on CE1.

References

[1] R. Achanta and S. Süsstrunk, “Superpixels and Polygons using simple non-iterative clustering,” in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, pp. 4895–4904.