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

ISO/IEC JTC1/SC29/WG11
MPEG2020/m53572
April 2020, Online

Source **Poznań University of Technology, Poznań, Poland**
Electronics and Telecommunications Research Institute, Daejeon, Republic of Korea

Status **Input**

Title **[MPEG-I Visual] Segmentation-based geometry upsampling in TMIV**

Author **Dawid Mieloch*, Adrian Dziembowski*, Jakub Stankowski*, Marek Domański*, Gwangsoon Lee****, **Jun Young Jeong****
* - Poznań University of Technology,
**- Electronics and Telecommunications Research Institute

Contact dawid.mieloch@put.poznan.pl

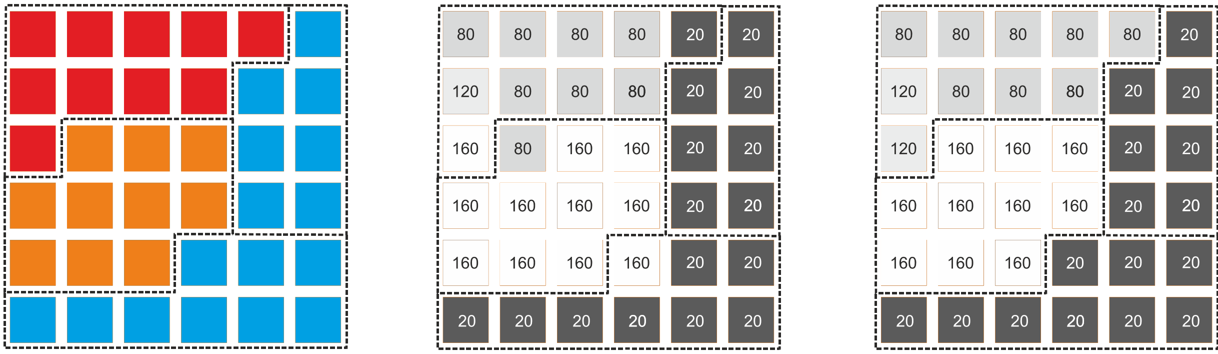
Abstract

This document describes the segmentation-based method of upsampling geometry in TMIV. The method refines geometry basing on segmented texture received by TMIV decoder.

1 Introduction

In the proposed method, the low-resolution geometry is enhanced using superpixel segmentation [1]. 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 potentially good tool for the removal of errors caused both by downscaling and encoding of geometry.

At first, geometry atlases are upscaled. Then, superpixels are computed on the decoded texture. For all superpixels, all pixels of geometry that belong to a patch are sorted. If any value of depth is in nearest or farthest 10% of depth range for given superpixel, then this depth is changed to the nearest value in the range.



The example of how the method works is presented above. The image on the left shows a 6×6 patch. The patch was segmented into 4 smaller superpixels, basing on the texture information. As it can be seen in the middle image, the "red" segment has such values of depth: 20, 80, 80, 80, 80, 80, 80, 80, 80, 120, 160. After performing the depth enhancement (top and bottom 10% of depth values in each segment are changed), value 160 will become value 120 and value 20 will become value 80, removing the obvious error in depth.

2 Experimental results

2.1 *Without TMIV upscaling*

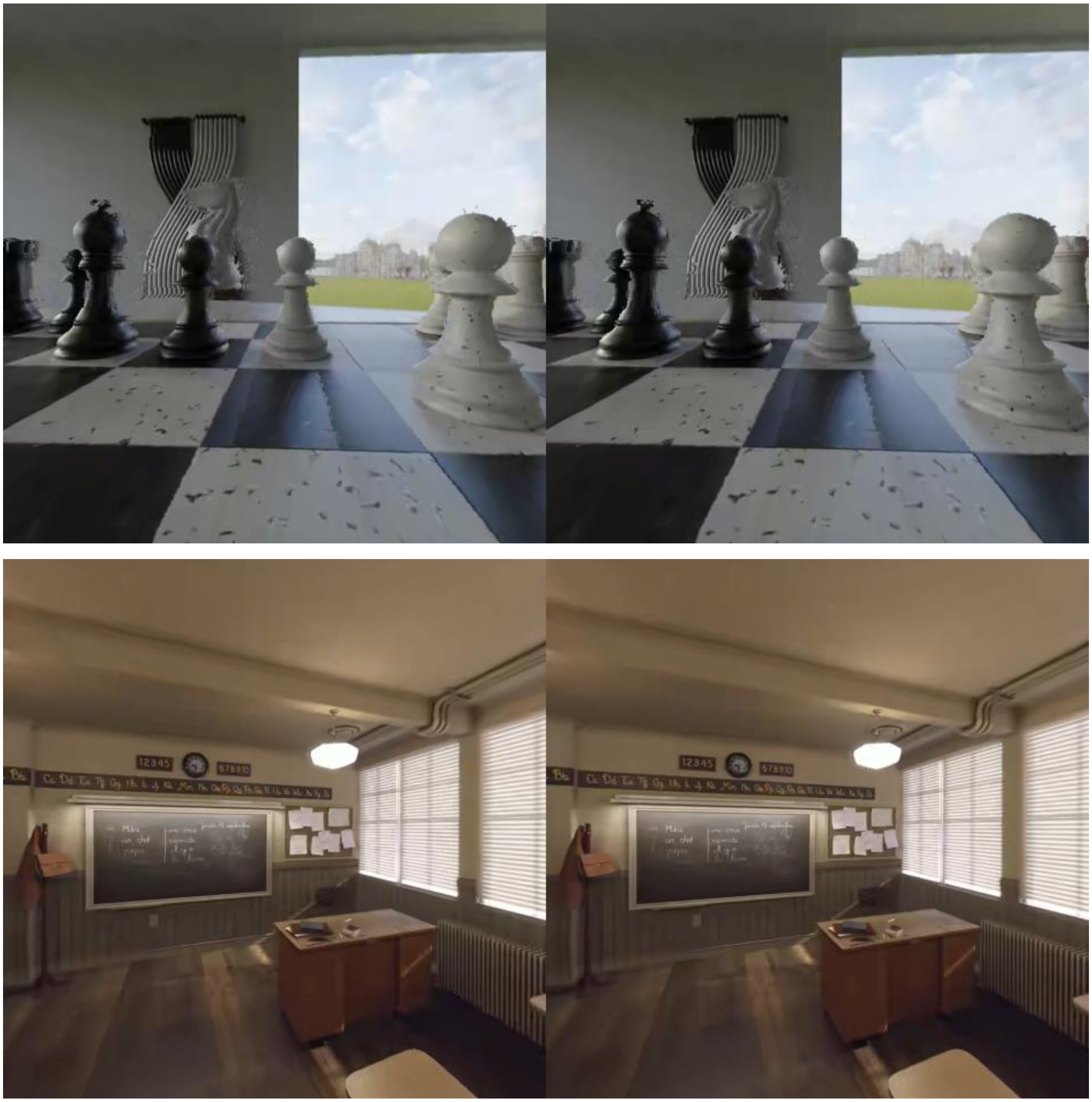
In the first experiment, the TMIV upscaler was removed. In this experiment, geometry atlases were upscaled using nearest neighbor interpolation. Then, proposed segmentation-based refinement was applied.

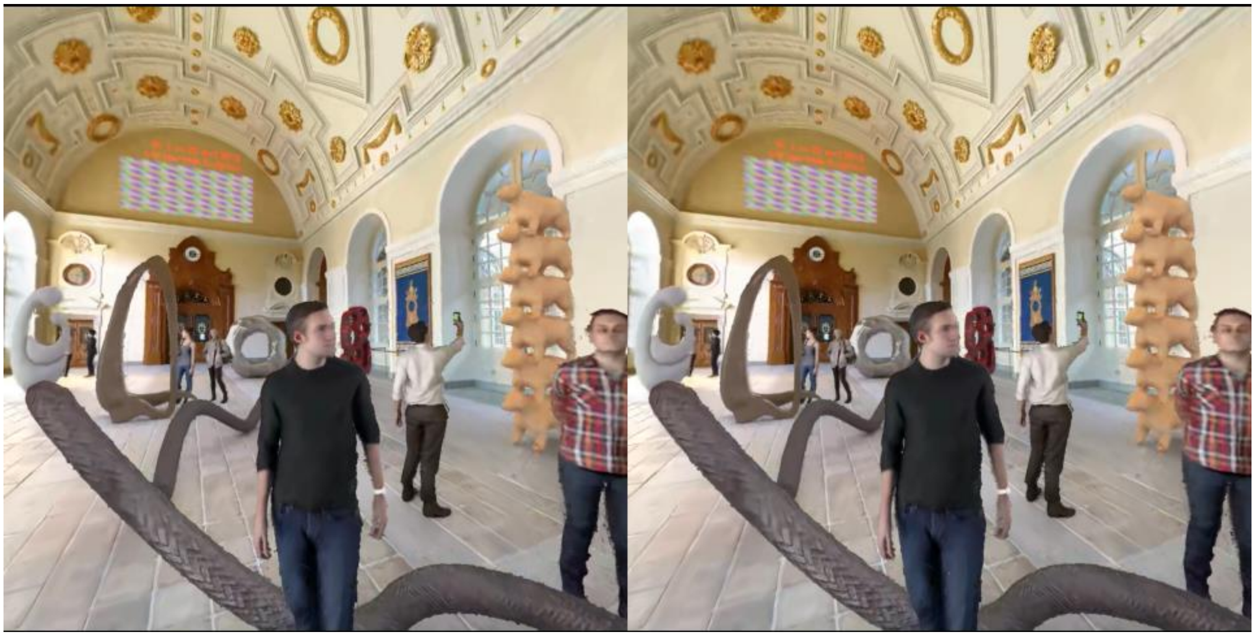
Results presented below show that the proposal works better with lower rates. Posetraces (uploaded to MPEG content server http://mpegfs.int-evry.fr/mpegcontent/ws-mpegcontent/MPEG-I/Poznan/m53572) confirm this dependency, moreover, for some sequences the proposal shows subjective quality improvements both for QP1 and QP5 (e.g. SN, SC, SA).

Test class	Sequence	Anchor (ff)	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 ratio
CG	ClassroomVideo	AA97 (MIV)	-9.1%	-3.4%	3.22	-1.0%	-0.5%	1.00
	TechnicolorMuseum	BA97 (MIV)	39.9%	20.1%	14.57	3.5%	2.0%	1.00
	InterdigitalHijack	CA97 (MIV)	5.1%	1.3%	12.58	0.7%	0.1%	1.00
	OrangeKitchen	JA97 (MIV)	21.4%	10.9%	14.47	2.6%	1.3%	1.00
	NokiaChess (*)	NA97 (MIV)	0.6%	-3.0%	16.87	0.3%	-1.1%	1.00
		All MIV Anchor	14.3%	7.3%	14.57	1.4%	0.7%	

NC	TechnicolorPainter	DA97 (MIV)	3.8%	2.1%	8.18	0.8%	0.5%	1.00
	IntelFrog	EA97 (MIV)	14.4%	6.5%	11.26	3.7%	1.5%	1.00
	PoznanFencing	LA97 (MIV)	3.4%	2.1%	13.95	-1.0%	-0.6%	1.00
		All MIV Anchor	7.2%	3.5%	13.95	1.2%	0.5%	

Examples of side-by-side comparisons with anchor (on the left) for QP5 are presented below.





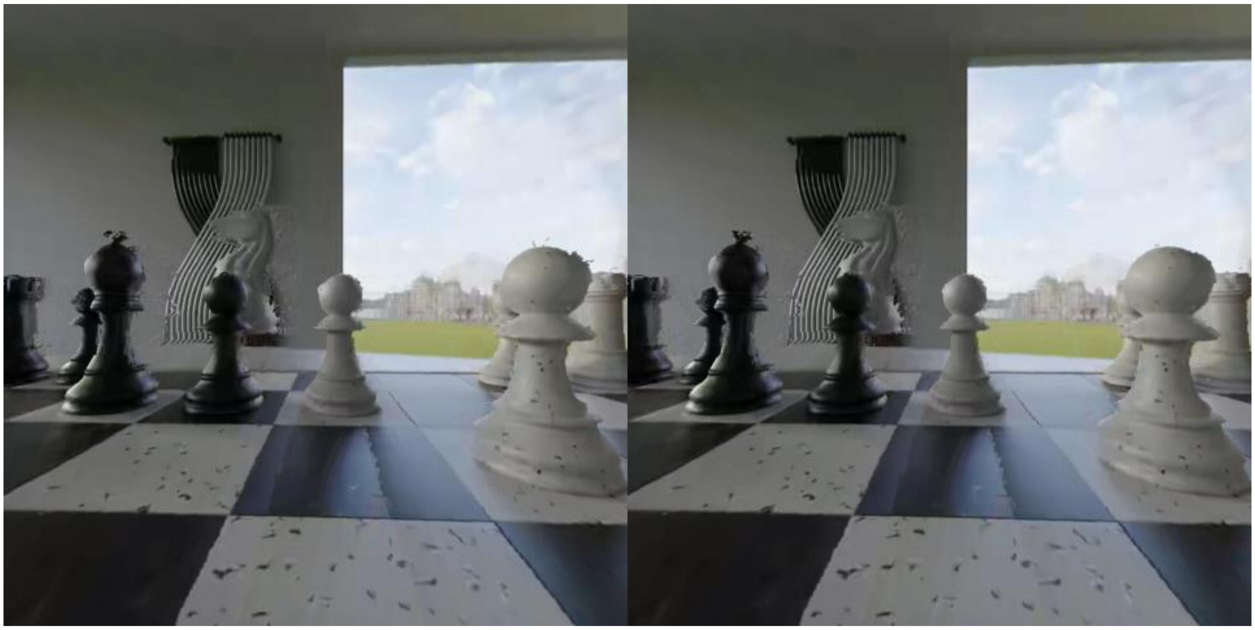
2.2 *With TMIV upscaling*

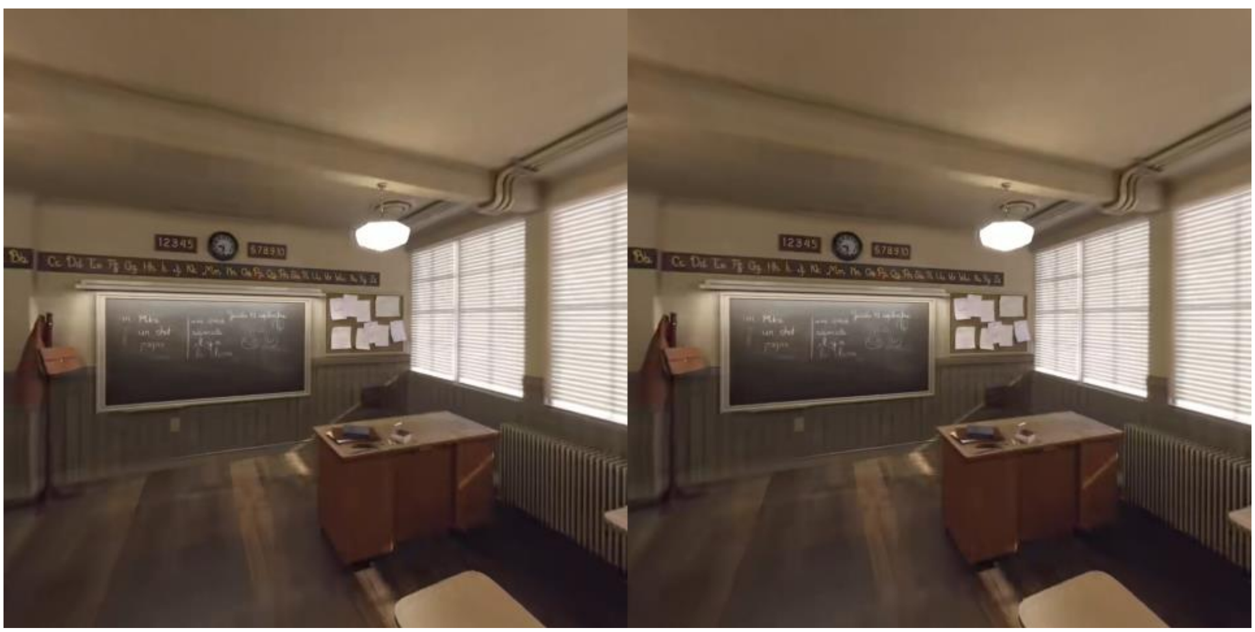
In the second experiment, proposed segmentation-based upscaling was performed on previously upscaled (using TMIV4 upscaling) geometry. Posetraces were also uploaded to MPEG content server. Examples of side-by-side comparisons with anchor (on the left) for QP5 are presented below.

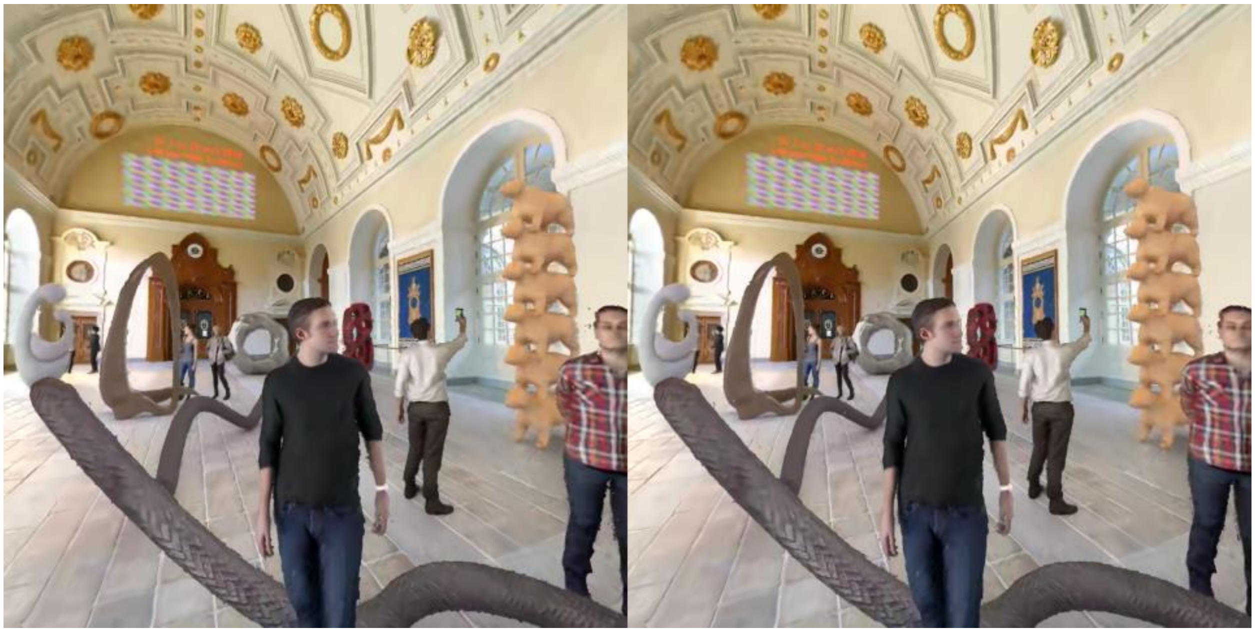
As posetraces show, merging on the proposal with the current TMIV4 upscaling has beneficial effects on the quality of synthesized videos not only for low bitrates. The objective quality was also increased in comparison with the first part of the experiment.

Test class	Sequence	Anchor (ff)	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 ratio
CG	ClassroomVideo	AA97 (MIV)	9.2%	3.4%	3.42	-0.1%	-0.1%	1.00
	TechnicolorMuseum	BA97 (MIV)	2.9%	0.9%	13.92	0.9%	0.3%	1.00
	InterdigitalHijack	CA97 (MIV)	-1.7%	-3.0%	12.26	-0.5%	-0.7%	1.00
	OrangeKitchen	JA97 (MIV)	10.5%	5.7%	14.78	0.8%	0.4%	1.00
	NokiaChess (*)	NA97 (MIV)	-7.2%	-7.4%	17.40	-2.3%	-2.9%	1.00
		All MIV Anchor	5.2%	1.7%	14.78	0.3%	0.0%	

NC	TechnicolorPainter	DA97 (MIV)	0.3%	0.3%	8.24	0.3%	0.2%	1.00
	IntelFrog	EA97 (MIV)	-1.4%	-0.7%	11.27	-0.6%	-0.2%	1.00
	PoznanFencing	LA97 (MIV)	-0.3%	0.0%	14.02	-0.4%	-0.1%	1.00
		All MIV Anchor	-0.4%	-0.1%	14.02	-0.2%	0.0%	







3 Conclusions

The document provides a brief description of the technique for upsampling the geometry in TMIV with the use of superpixel segmentation. The posetraces show improvement of the quality of final synthesis.

4 Recommendations

We recommend to:

- include proposed technique into TMIV5,
- continue works on the geometry upsampling and refinement.

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).

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.