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/M51603
January 2020, Brussels, Belgium

Source	**Poznań University of Technology (PUT), Poznań, Poland** **Electronics and Telecommunications Research Institute (ETRI), Daejeon, Republic of Korea**
Status	**Input**
Title	**Immersive Video CE3.2: Temporal patch redundancy removal**
Author	Adrian Dziembowski*, Dawid Mieloch*, Adam Grzelka*, Jakub Stankowski*, Marek Domański*, Gwangsoon Lee** * – Poznań University of Technology, ** – Electronics and Telecommunications Research Institute

1 Introduction

This document presents a technical description of one of the PUT/ETRI experiments on the atlas preparation (Immersive Video CE-3 [1]).

2 Overview of the proposed technique

Atlas preparation algorithm in TMIV still has one major flaw – there is a lot of redundant parts of input views in the atlases. The same information is copied within numerous patches, thus repeated many times. There are two reasons of that redundancy:

1. spatial – patches are filled by copying information in the entire bounding box of the cluster,
2. temporal – size and shape of each patch is aggregated for the whole GOP.

The proposed approach reduces influence of both issues, significantly reducing data redundancy, thus the size of the final bitstream. Here, the proposed temporal redundancy technique will be presented.

2.1 Temporal redundancy reduction

In TMIV 3.0, the shape of each patch does not change during the entire GOP. However, if there is any movement in the scene, some pixels are redundant and do not have to be copied for all the frames but only for the frames where they are needed.

In the proposed approach, only the blocks that contain pixels belonging to the cluster for the current frame are copied from the source views to the atlases. In Fig. 1 the comparison between TMIV 3.0 (left) and proposed method (right) is presented. In TMIV, the patch is fully copied for all frames. In proposed approach, major part of the patch's area is empty (thus easier to encode).

Fig. 1. TMIV 3.0 (left) vs. proposed method (right).

3 Experimental results

The results of the proposed enhancement are presented in the table below.

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 VMAF	Low-BR BD rate VMAF	High-BR BD rate MS-SSIM	Low-BR BD rate MS-SSIM	High-BR BD rate IV-PSNR	Low-BR BD rate IV-PSNR	Pixel rate ratio
CG	ClassroomVideo	AA97 (MIV)	9.0%	5.8%	4.51	4.0%	2.8%	-2.6%	-0.4%	-11.4%	-5.9%	0.00%
	TechnicolorMuseum	BA97 (MIV)	0.5%	-0.2%	13.03	-0.7%	-2.5%	-3.7%	-4.9%	-15.2%	-13.8%	0.00%
	TechnicolorHijack	CA97 (MIV)	5.4%	12.1%	11.90	-0.5%	2.4%	1.2%	9.5%	-2.2%	6.3%	0.00%
	OrangeKitchen	JA97 (MIV)	3.8%	2.9%	13.87	3.3%	0.3%	-1.2%	-0.8%	-7.1%	-3.2%	0.00%
		MIV	4.7%	5.2%	13.87	1.5%	0.8%	-1.6%	0.9%	-9.0%	-4.2%	0.00%
		All anchors	4.7%	5.2%	13.87	1.5%	0.8%	-1.6%	0.9%	-9.0%	-4.2%	0.00%
NC	TechnicolorPainter	DA97 (MIV)	6.8%	9.7%	6.72	-1.9%	0.9%	-7.6%	-0.5%	0.0%	4.5%	0.00%
	IntelFrog	EA97 (MIV)	40.6%	3.4%	10.97	2.9%	-13.3%	10.5%	-10.2%	39.9%	2.3%	0.00%
	PoznanFencing	LA97 (MIV)	93.1%	82.5%	13.97	134.6%	88.4%	301.3%	132.0%	35.8%	42.6%	0.00%
		MIV	46.8%	31.9%	13.97	45.2%	25.3%	101.4%	40.4%	25.2%	16.4%	0.00%
		All anchors	46.8%	31.9%	13.97	45.2%	25.3%	101.4%	40.4%	25.2%	16.4%	0.00%

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 VMAF	Low-BR BD rate VMAF	High-BR BD rate MS-SSIM	Low-BR BD rate MS-SSIM	High-BR BD rate IV-PSNR	Low-BR BD rate IV-PSNR	Pixel rate ratio
All		MIV	22.7%	16.6%	10.71	20.2%	11.3%	42.5%	17.8%	5.7%	4.7%	0.00%
		All anchors	22.7%	16.6%	10.71	20.2%	11.3%	42.5%	17.8%	5.7%	4.7%	0.00%

As the results show, use of the proposal decreases the IV-PSNR for CG sequences on average even by 9% (for high QPs).

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 Hyunho Kim from KAU for thorough crosscheck of the presented results.

5 Recommendations

Considering much better results than for TMIV 3.0 anchor for CG sequences, we suggest to include our technique into TMIV 4.0.

6 References

[1] Renaud Doré, "Description of Immersive Video Core Experiments 3 (Atlas preparation)", ISO/IEC JTC1/SC29/WG11 MPEG/N18934, October 2019, Geneva, CH.