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/M52347
January 2020, Brussels, Belgium

Source	Zhejiang University, Zhejiang, China Poznań University of Technology (PUT), Poznań, Poland
Status	Input for MPEG-I Visual
Title	MPEG-I Visual] Immersive Video CE3.6: Combination of Patch Merge and Patch Split
Authors	Bin Wang*, Yule Sun*, Lu Yu*, Adrian Dziembowski**, Dawid Mieloch** *—Zhejiang University, **—Poznań University of Technology,

1 Introduction

This contribution presents a description of ZJU and PUT's experiments on Immersive Video CE3.6 (Combination of Patch Merge and Patch Split) [1]-[2]. The goal of experiment is to improve coding efficiency, reduce pixel rate and reduce visual artifacts.

2 Description of ZJU and PUT's experiments

(1) ***Patch merging***

Before patch packing, we add a process "patch merging".

There are 2 patches (A and B), we judge whether the patches are need to be merged by the following function:

```
If(     patchA.width * patchA.height + patchB.width * patchB.height
        >= Mergedpatch.width * Mergedpatch.height  )
        is_PatchMerge = 1;
else
        is_PatchMerge = 0;
```

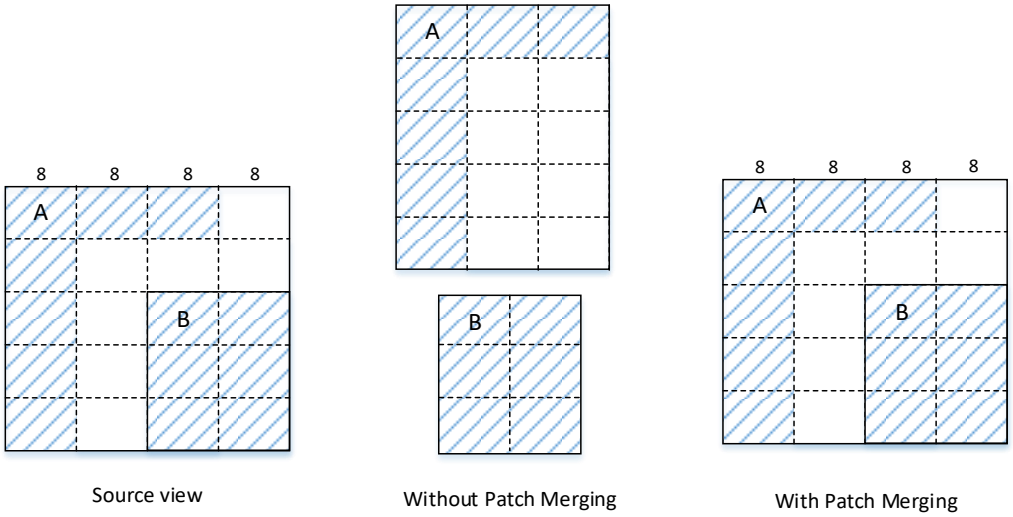

Fig. 1. Illustration of Patch Merging

(2) ***Patch split***

If the shape of the cluster is similar to the rectangle, copying information from the entire bounding box does not introduce much redundancy. However, if the cluster is L-shaped, the quantity of data copied from the source view could be significantly reduced (Fig. 1).

In order to decide how to split an L-shaped cluster, the total area of two subpatches is being minimized. The split line is always parallel to the shorter side of the patch. If splitting would not decrease the total area more than 10%, the split is not performed.

This approach allows to efficiently divide an L-shaped cluster. However, for other cluster shapes (e.g. C-shape), such approach does not result in the division of a cluster. Therefore, we proposed an additional cluster splitting (Fig. 2).

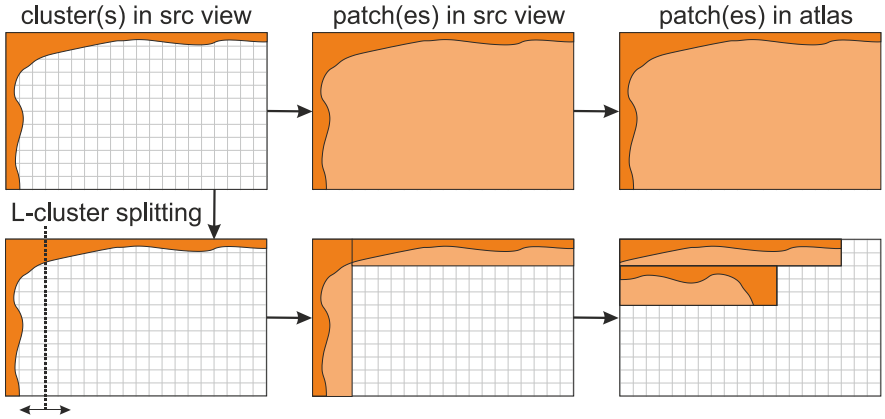

Fig. 2. L-shaped cluster splitting. The alignment grid is colored in grey.

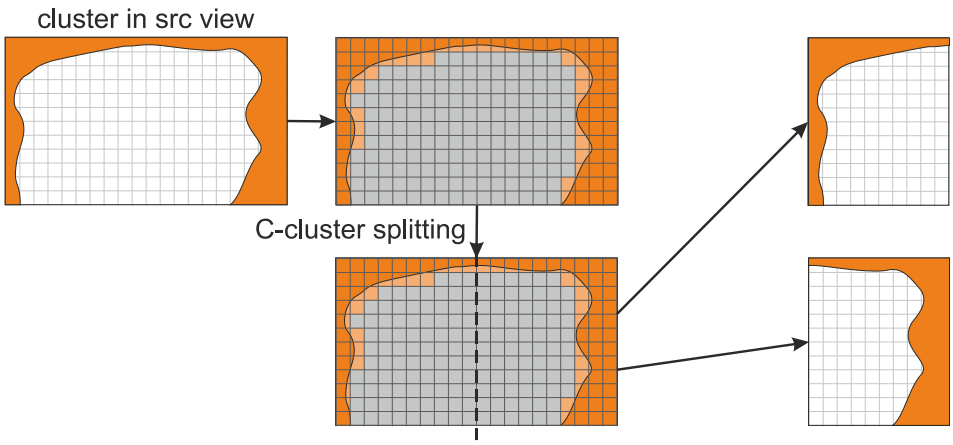

Fig. 2. C-shaped cluster splitting. The alignment grid is colored in grey.

Within the entire bounding box of the cluster, we calculate the number of 32×32 blocks that contain pixels belonging to the cluster (orange blocks in Fig. 2). Then, calculated number is divided by the total number of blocks within the analyzed bounding box. If that ratio is smaller than 30%, the cluster is split in half. Splitting of C-shaped cluster usually results in two L-shaped clusters.

Proposed cluster splitting is a recursive method. Example of the recursive splitting of an irregularly-shaped cluster is presented in Fig. 3.

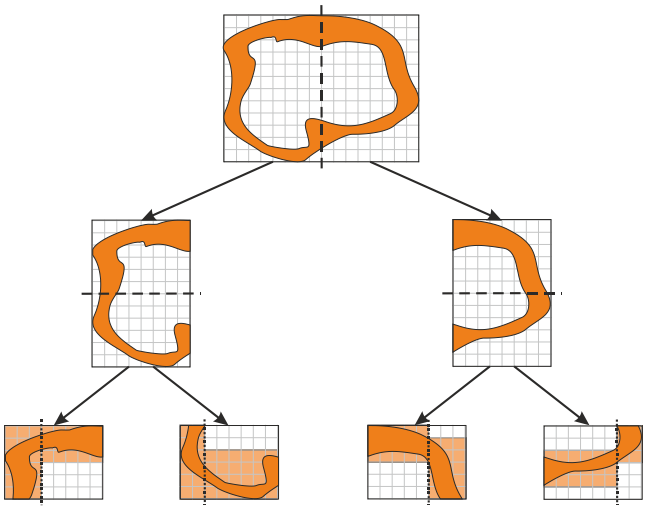

Fig. 3. Recursive splitting of the patch; dashed lines: C-splitting, dotted lines: L-splitting.

We decided that clusters smaller than 64×64 should not be split. It would result in a large number of very small clusters, smaller than a CU block, heavily increasing the required bitrate in HEVC encoding.

3 Experimental results

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)	-2.2%	4.5%	4.24	1.4%	7.1%	5.5%	7.7%	1.1%	4.9%	0.00%
	Technicolor Museum	BA97 (MIV)	-4.3%	-5.7%	12.08	-7.1%	-8.0%	-6.6%	-7.6%	#####	#####	0.00%
	Technicolor Hijack	CA97 (MIV)	4.6%	6.1%	11.54	-0.5%	-0.2%	5.2%	5.3%	0.8%	3.5%	0.00%
	OrangeKitchen	JA97 (MIV)	-5.4%	-5.3%	12.76	-8.5%	-5.7%	-1.5%	-1.3%	-4.5%	-4.7%	0.00%
	NokiaChess	NA97 (MIV)	-6.8%	-3.4%	14.43	3.4%	3.4%	6.1%	5.7%	-5.3%	-2.7%	0.00%
		MIV	-1.8%	-0.1%	12.76	-3.7%	-1.7%	0.7%	1.0%	-3.7%	-2.0%	0.00%
		All anchors	-1.8%	-0.1%	12.76	-3.7%	-1.7%	0.7%	1.0%	-3.7%	-2.0%	0.00%
NC	Technicolor Painter	DA97 (MIV)	6.6%	5.6%	6.10	3.4%	3.3%	3.6%	3.8%	3.2%	3.4%	0.00%
	IntelFrog	EA97 (MIV)	7.1%	4.8%	9.02	4.3%	3.4%	5.6%	3.8%	4.5%	3.5%	0.00%
	PoznanFencing	LA97 (MIV)	-8.5%	-5.9%	13.09	#####	-6.7%	-6.4%	-2.6%	-2.3%	-0.9%	0.00%
		MIV	1.7%	1.5%	13.09	-1.2%	0.0%	1.0%	1.6%	1.8%	2.0%	0.00%
		All anchors	1.7%	1.5%	13.09	-1.2%	0.0%	1.0%	1.6%	1.8%	2.0%	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	-0.3%	0.6%	9.83	-2.6%	-1.0%	0.8%	1.3%	-1.4%	-0.3%	0.00%
		All anchors	-0.3%	0.6%	9.83	-2.6%	-1.0%	0.8%	1.3%	-1.4%	-0.3%	0.00%

Fig. 2. Result of CE3.6

4 CE3.6 experiments for reduced pixel rate

4.1 *Introduction*

Patch splitting reduces spatial redundancy of patches, what results in more empty space in atlases. Therefore, it is possible to reduce pixel rate without significant change of the quality of synthesized views.

We performed two experiments with lowered pixel rates. In order to present fair comparison, we used the same pixel rates as in m51602.

4.2 *Merge + split*

In the first experiment we preserved operation order from CE3.6: patches are being merged at first and splitted then.

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)	-2.2%	4.5%	4.24	1.4%	7.1%	5.5%	7.7%	1.1%	4.9%	0.00%
	TechnicolorMuseum	BA97 (MIV)	-4.8%	-6.5%	12.32	-7.7%	-9.4%	-7.8%	-9.1%	-15.0%	-14.5%	-13.51%
	TechnicolorHijack	CA97 (MIV)	4.6%	6.1%	11.54	-0.5%	-0.2%	5.2%	5.3%	0.8%	3.5%	-20.00%
	OrangeKitchen	JA97 (MIV)	4.1%	0.0%	13.51	-4.7%	-6.2%	2.1%	-0.9%	4.9%	0.0%	-11.11%
	NokiaChess	NA97 (MIV)	-6.8%	-3.4%	14.43	3.4%	3.4%	6.1%	5.7%	-5.3%	-2.7%	-13.51%
		MIV	0.4%	1.0%	13.51	-2.9%	-2.2%	1.3%	0.7%	-2.0%	-1.5%	-11.44%
		All anchors	0.4%	1.0%	13.51	-2.9%	-2.2%	1.3%	0.7%	-2.0%	-1.5%	-11.44%
NC	TechnicolorPainter	DA97 (MIV)	6.6%	5.6%	6.10	3.4%	3.3%	3.6%	3.8%	3.2%	3.4%	0.00%
	IntelFrog	EA97 (MIV)	7.1%	4.8%	9.02	4.3%	3.4%	5.6%	3.8%	4.5%	3.5%	0.00%
	PoznanFencing	LA97 (MIV)	-8.5%	-5.9%	13.09	-11.3%	-6.7%	-6.4%	-2.6%	-2.3%	-0.9%	0.00%
		MIV	1.7%	1.5%	13.09	-1.2%	0.0%	1.0%	1.6%	1.8%	2.0%	0.00%
		All anchors	1.7%	1.5%	13.09	-1.2%	0.0%	1.0%	1.6%	1.8%	2.0%	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	1.0%	1.2%	9.97	-2.1%	-1.2%	1.1%	1.1%	-0.4%	0.0%	-6.71%
		All anchors	1.0%	1.2%	9.97	-2.1%	-1.2%	1.1%	1.1%	-0.4%	0.0%	-6.71%

Fig. 3. Result of CE3.6 with lowered pixel rate

4.3 *Split + merge*

In the second experiment we used inverted order of operations: split first, then merge.

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)	-5.4%	1.3%	4.21	0.9%	4.5%	4.4%	5.5%	1.3%	3.6%	0.00%
	TechnicolorMuseum	BA97 (MIV)	0.1%	-4.2%	12.92	-6.2%	-9.5%	-4.8%	-8.4%	-7.3%	-10.6%	-13.51%
	TechnicolorHijack	CA97 (MIV)	0.9%	0.2%	11.47	-6.5%	-7.3%	2.2%	0.1%	-3.3%	-2.9%	-20.00%
	OrangeKitchen	JA97 (MIV)	5.6%	-0.2%	14.05	-2.0%	-5.5%	3.8%	-1.4%	9.0%	1.6%	-11.11%
	NokiaChess	NA97 (MIV)	10.2%	3.6%	15.86	6.9%	1.6%	7.6%	2.6%	13.9%	4.2%	-13.51%
		MIV	0.3%	-0.7%	14.05	-3.4%	-4.5%	1.4%	-1.1%	-0.1%	-2.1%	-11.44%
		All anchors	0.3%	-0.7%	14.05	-3.4%	-4.5%	1.4%	-1.1%	-0.1%	-2.1%	-11.44%
NC	TechnicolorPainter	DA97 (MIV)	4.3%	4.0%	6.22	5.2%	4.2%	4.3%	3.9%	2.7%	3.0%	0.00%
	IntelFrog	EA97 (MIV)	4.2%	1.2%	8.93	1.4%	-0.1%	0.2%	-1.2%	2.9%	0.2%	0.00%
	PoznanFencing	LA97 (MIV)	26.7%	19.2%	13.69	14.9%	6.0%	33.3%	14.3%	18.2%	14.9%	0.00%
		MIV	11.7%	8.1%	13.69	7.2%	3.4%	12.6%	5.7%	7.9%	6.1%	0.00%
		All anchors	11.7%	8.1%	13.69	7.2%	3.4%	12.6%	5.7%	7.9%	6.1%	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	5.2%	3.1%	10.21	1.1%	-1.1%	6.2%	1.8%	3.4%	1.4%	-6.71%
		All anchors	5.2%	3.1%	10.21	1.1%	-1.1%	6.2%	1.8%	3.4%	1.4%	-6.71%

Fig. 4. Result of CE3.6 (with inverted order: split first) with lowered pixel rate

4.4 *Conclusions*

In general, it is better to perform merging at first. Merge + split does not have the problem of significant quality decrease (BD-rate higher than 10%).

On average, presented approach allows to preserve similar BD-rates as anchor while pixel rate is decreased.

5 Acknowledgement

We would like to thank Gwangsoon Lee from ETRI, Adrian Dziembowski and Dawid Mieloch for thorough crosscheck of the presented results.

6 Conclusions and recommendations

Considering the Pixel rate reduction and BD-rate gain in comparison with TMIV 3.0, we suggest to include our technique into TMIV 4.0.

7 References

[1] B. Salahieh, B. Kroon, J. Jung, M. Domański (Eds.), "Test Model 3 for Immersive Video," ISO/IEC JTC1/SC29/WG11, N18577, July. 2019.

[2] R. Doré, "Description of Immersive Video Core Experiments 3," ISO/IEC JTC1/SC29/WG11, N18707, July 2019.