INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG04
MPEG VIDEO CODING

ISO/IEC JTC1/SC29/WG04 MPEG/M55750
January 2021, Online

Source Poznań University of Technology, Poznań, Poland
Electronics and Telecommunications Research Institute, Daejeon, Rep. of Korea
Status Input document
Title MIV CE3.2-related: Temporal patch splitting
Author Adrian Dziembowski, Dawid Mieloch, Marek Domański (PUT)
Gwangson Lee, Jun Young Jeong (ETRI)

Abstract

This document presents a technical description of the PUT/ETRI experiment on temporal patch splitting and packing (MIV CE3.2). In the proposed solution, patches are allowed to be split temporally in order to provide more efficient packing.

1 Proposed technique

In TMIV7, the duration of all the patches is the same, equal to the intra period. When a patch contains an area with moving objects, a significant part of the area of this patch can be empty/unoccupied because of temporal redundancy removal. However, this area cannot be used for packing other patches if it was occupied even for one frame.

We propose a solution that allows using such an area, thus, to pack patches more efficiently. The idea of the proposed solution is presented in Fig. 1.

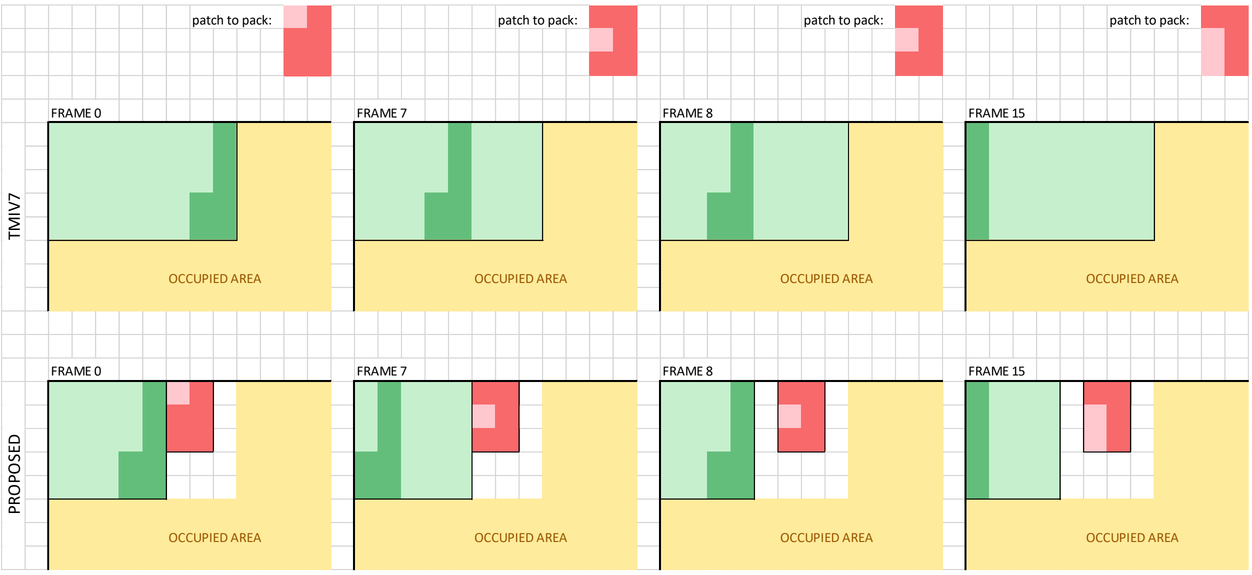

Fig. 1. Proposed approach vs. TMIV7.

In Fig. 1, the case of a patch containing a moving object is presented (green patch; dark green indicates area occupied in the current frame). Let us assume that other parts of the atlas are occupied (yellow area).

In TMIV 7, the light green area would be greyed out, but the red patch could not be packed there. Therefore, the red patch cannot be packed into the atlas thus this patch would be discarded.

In the proposed solution, the green patch would be split into two subpatches of equal duration (original duration: frames 0 – 15, result: one subpatch for frames 0 to 7 and second subpatch for frames 8 to 15). In this case, the area needed for two green subpatches would be smaller than without splitting, allowing to properly pack the red patch. Moreover, the red patch does not need to be temporally split.

The proposed solution required changes in the packing algorithm. Two main changes were made. At first, the aggregatedPruningMask is not used. Instead, the packer uses nonAggregatedPruningMask. When assessing the overlap between a new patch and patches already packed, the packer compares masks for each frame. If the block is occupied in both patches for any frame, the patch cannot be packed in the analyzed position.

Secondly, the pushInFreeSpace function is not used, thus the packing algorithm is no longer the maxRect one. All the patches are packed into the first available position.

The temporal splitting of the patch is performed recursively, together with L-shaped and C-shaped cluster spatial splitting, as written using pseudocode below:

```
for each cluster:
      performRecursiveSplit();

performRecursiveSplit(){
      if(splitLClusterVertically())    performRecursiveSplit();
      if(splitCClusterVertically())    performRecursiveSplit();
      if(splitLClusterHorizontally())  performRecursiveSplit();
      if(splitCClusterHorizontally())  performRecursiveSplit();
      if(splitClusterTemporally())     performRecursiveSplit();
}
```

The patch is split temporally if the total area of subpatches after splitting (summarized for all frames) is smaller than 80% of the area of the original patch.

The maximum number of temporal splits can be set (e.g. in the configuration file) to avoid splitting into short (e.g. 1-frame) subpatches, which will increase the bitrate of video atlases (big changes between frames) and metadata (more patches).

2 Experimental results

In the experiments, 3 cases were tested. In each, a different number of temporal splits of a patch was allowed. 1, 2 and 3 temporal patch splits were tested.

Table 1. Objective quality evaluation – 1 temporal split allowed.

Mandatory content - Proposal vs. Low/High-bitrate Anchors

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
ClassroomVideo	SA	2.0%	1.7%	1.72	1.8%	1.7%	1.0%	1.4%
Museum	SB	0.6%	0.4%	16.80	0.4%	0.3%	0.1%	0.1%
Fan	SO	6.5%	-0.0%	7.66	-5.8%	-7.8%	-1.7%	-5.9%
Kitchen	SJ	-0.0%	-0.0%	16.33	-0.0%	-0.0%	-0.0%	-0.0%
Painter	SD	-4.4%	-5.4%	8.11	-6.0%	-6.2%	-5.6%	-6.0%
Frog	SE	-2.0%	-3.7%	5.45	-3.3%	-4.4%	-5.0%	-5.1%
Carpark	SP	5.4%	0.0%	7.26	1.6%	-1.8%	0.8%	-1.8%
Chess	SN	3.1%	1.2%	15.72	2.3%	1.3%	1.6%	1.6%
Group	SR	1.1%	1.2%	11.81	0.9%	0.9%	0.6%	0.7%
MIV		**1.4%**	**-0.5%**	**10.10**	**-0.9%**	**-1.8%**	**-0.9%**	**-1.7%**

Optional content - Proposal vs. Low/High-bitrate Anchors

Fencing	SL	59.9%	16.8%	11.01	20.2%	5.6%	14.3%	3.2%
Street	SU	6.8%	-1.3%	8.52	5.1%	-2.3%	-2.2%	-4.3%
Hall	ST	33.5%	15.9%	11.73	11.8%	5.6%	13.0%	5.5%
ChessPieces	SQ	-6.4%	-1.7%	15.77	-3.6%	-0.8%	-0.4%	0.1%
Hijack	SC	1.2%	1.3%	9.42	0.3%	1.0%	1.2%	1.4%
MIV		**19.0%**	**6.2%**	**11.29**	**6.8%**	**1.8%**	**5.2%**	**1.2%**

Table 2. Objective quality evaluation – 2 temporal splits allowed.

Mandatory content - Proposal vs. Low/High-bitrate Anchors

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
ClassroomVideo	SA	1.6%	1.8%	1.71	2.2%	1.8%	1.9%	1.3%
Museum	SB	0.7%	0.5%	16.80	0.5%	0.4%	0.2%	0.2%
Fan	SO	6.2%	-0.3%	7.68	-6.1%	-8.2%	-2.0%	-6.3%
Kitchen	SJ	0.9%	0.6%	16.37	1.0%	0.8%	0.7%	0.5%
Painter	SD	-5.0%	-6.0%	8.13	-6.7%	-6.9%	-6.3%	-6.7%
Frog	SE	-2.2%	-3.8%	5.46	-3.3%	-4.4%	-5.2%	-5.2%
Carpark	SP	5.8%	0.3%	7.28	2.1%	-1.4%	1.0%	-1.7%
Chess	SN	3.0%	1.8%	15.74	3.0%	2.2%	1.8%	2.0%
Group	SR	1.1%	1.4%	11.76	1.0%	1.1%	0.5%	0.7%
MIV		**1.3%**	**-0.4%**	**10.10**	**-0.7%**	**-1.6%**	**-0.8%**	**-1.7%**

Optional content - Proposal vs. Low/High-bitrate Anchors

Fencing	SL	61.2%	17.0%	11.05	20.4%	5.7%	14.3%	3.1%
Street	SU	7.3%	-1.3%	8.54	5.5%	-2.3%	-2.3%	-4.6%
Hall	ST	35.4%	16.5%	11.73	12.7%	6.3%	14.3%	6.1%
ChessPieces	SQ	-5.1%	-0.5%	15.65	-3.4%	-0.3%	0.6%	1.0%
Hijack	SC	2.0%	2.2%	9.40	1.9%	2.1%	1.9%	2.3%
MIV		**20.1%**	**6.8%**	**11.28**	**7.4%**	**2.3%**	**5.8%**	**1.6%**

Table 3. Objective quality evaluation – 3 temporal splits allowed.

Mandatory content - Proposal vs. Low/High-bitrate Anchors

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**
ClassroomVideo	SA	2.8%	2.1%	1.73	2.2%	2.0%	1.2%	1.5%
Museum	SB	0.7%	0.5%	16.80	0.5%	0.4%	0.2%	0.2%
Fan	SO	6.1%	-0.4%	7.71	-6.4%	-8.4%	-2.2%	-6.5%
Kitchen	SJ	1.3%	0.5%	16.37	1.2%	0.7%	0.8%	0.4%
Painter	SD	-5.5%	-6.6%	8.13	-7.3%	-7.6%	-6.8%	-7.3%
Frog	SE	-2.4%	-4.2%	5.48	-3.6%	-4.8%	-5.5%	-5.7%
Carpark	SP	6.2%	0.4%	7.30	2.3%	-1.3%	1.1%	-1.6%
Chess	SN	4.6%	2.5%	15.72	3.6%	2.8%	2.7%	2.6%
Group	SR	1.4%	1.3%	11.81	1.0%	1.0%	0.9%	0.9%
MIV		**1.7%**	**-0.4%**	**10.12**	**-0.7%**	**-1.7%**	**-0.9%**	**-1.7%**

Optional content - Proposal vs. Low/High-bitrate Anchors

Fencing	SL	62.4%	17.1%	11.08	20.9%	5.6%	14.4%	3.1%
Street	SU	7.6%	-1.3%	8.55	5.6%	-2.5%	-2.4%	-4.8%
Hall	ST	37.5%	17.5%	11.75	14.2%	7.4%	15.2%	6.6%
ChessPieces	SQ	-7.1%	-0.9%	15.78	-3.6%	-0.1%	0.1%	1.3%
Hijack	SC	3.1%	3.2%	9.42	2.7%	3.0%	2.9%	3.2%
MIV		**20.7%**	**7.1%**	**11.32**	**8.0%**	**2.7%**	**6.1%**	**1.9%**

Table 5. Bitrate savings – 1 temporal split allowed.

SA	SB	SC	SJ	SN	SD	SE	SP	SL	ST	SU	SR	SO	SQ	CG	NC	All
0.1%	-0.1%	0.6%	0.0%	0.7%	-7.2%	-7.3%	-4.6%	-8.8%	-4.5%	-4.2%	0.1%	-12.5%	0.6%	-1.3%	-6.1%	-3.4%
0.4%	0.0%	0.8%	0.0%	0.8%	-6.9%	-6.2%	-4.5%	-8.5%	-4.4%	-5.0%	0.1%	-12.1%	0.6%	-1.2%	-5.9%	-3.2%
0.8%	0.0%	1.0%	0.0%	0.9%	-6.5%	-5.0%	-4.1%	-7.6%	-3.9%	-5.5%	0.2%	-11.5%	0.7%	-1.0%	-5.4%	-2.9%
1.2%	0.0%	1.4%	0.0%	1.0%	-6.1%	-4.6%	-3.8%	-6.7%	-3.4%	-5.7%	0.3%	-10.9%	0.8%	-0.8%	-5.0%	-2.6%
1.6%	0.2%	1.6%	0.0%	1.2%	-5.6%	-4.7%	-3.2%	-3.9%	-2.4%	-5.7%	0.4%	-10.0%	1.0%	-0.5%	-4.3%	-2.1%

Table 6. Bitrate savings – 2 temporal splits allowed.

SA	SB	SC	SJ	SN	SD	SE	SP	SL	ST	SU	SR	SO	SQ	CG	NC	All
0.0%	0.0%	0.8%	0.0%	1.1%	-7.8%	-7.3%	-4.8%	-9.0%	-4.2%	-4.4%	0.2%	-12.8%	0.8%	-1.2%	-6.3%	-3.4%
0.3%	0.0%	1.2%	0.1%	1.2%	-7.5%	-6.3%	-4.6%	-8.7%	-4.2%	-5.2%	0.2%	-12.4%	1.0%	-1.0%	-6.1%	-3.2%
0.7%	0.1%	1.6%	0.2%	1.3%	-7.3%	-5.0%	-4.2%	-7.8%	-3.5%	-5.9%	0.3%	-11.8%	1.1%	-0.8%	-5.6%	-2.9%
1.1%	0.2%	2.2%	0.3%	1.4%	-6.9%	-4.6%	-3.8%	-6.9%	-2.9%	-6.1%	0.4%	-11.3%	1.2%	-0.6%	-5.2%	-2.5%
1.5%	0.3%	2.5%	0.4%	1.8%	-6.3%	-4.7%	-3.2%	-4.0%	-2.1%	-6.0%	0.6%	-10.2%	1.4%	-0.2%	-4.4%	-2.0%

Table 7. Bitrate savings – 3 temporal splits allowed.

SA	SB	SC	SJ	SN	SD	SE	SP	SL	ST	SU	SR	SO	SQ	CG	NC	All
0.0%	-0.1%	0.9%	-0.1%	1.5%	-8.2%	-7.6%	-4.8%	-9.2%	-4.2%	-4.6%	0.2%	-13.1%	1.2%	-1.2%	-6.4%	-3.4%
0.3%	0.0%	1.6%	-0.1%	1.6%	-8.1%	-6.6%	-4.6%	-8.9%	-4.2%	-5.5%	0.3%	-12.6%	1.4%	-0.9%	-6.3%	-3.2%
0.8%	0.1%	2.1%	0.0%	1.8%	-7.9%	-5.4%	-4.2%	-8.1%	-3.5%	-6.1%	0.4%	-12.0%	1.5%	-0.7%	-5.9%	-2.9%
1.2%	0.2%	2.9%	0.1%	1.9%	-7.5%	-5.0%	-3.8%	-7.1%	-2.8%	-6.3%	0.6%	-11.5%	1.7%	-0.4%	-5.4%	-2.5%
1.7%	0.3%	3.4%	0.3%	2.3%	-6.9%	-5.1%	-2.9%	-4.2%	-1.7%	-6.3%	0.9%	-10.5%	2.0%	0.1%	-4.5%	-1.9%

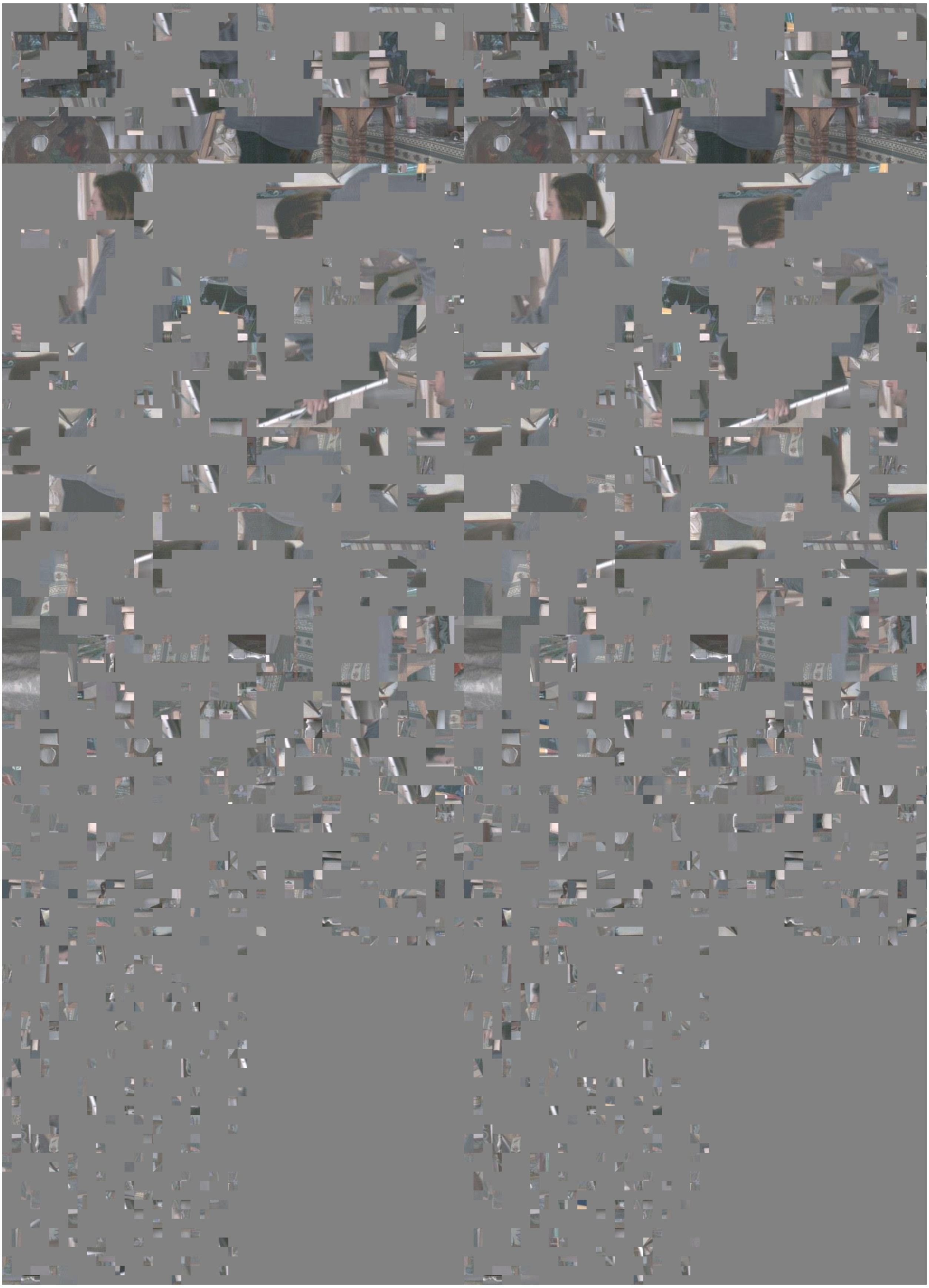

Fig. 2. Anchor: A97, SD, atlas 1, frame 15 (left) and 16 (right).

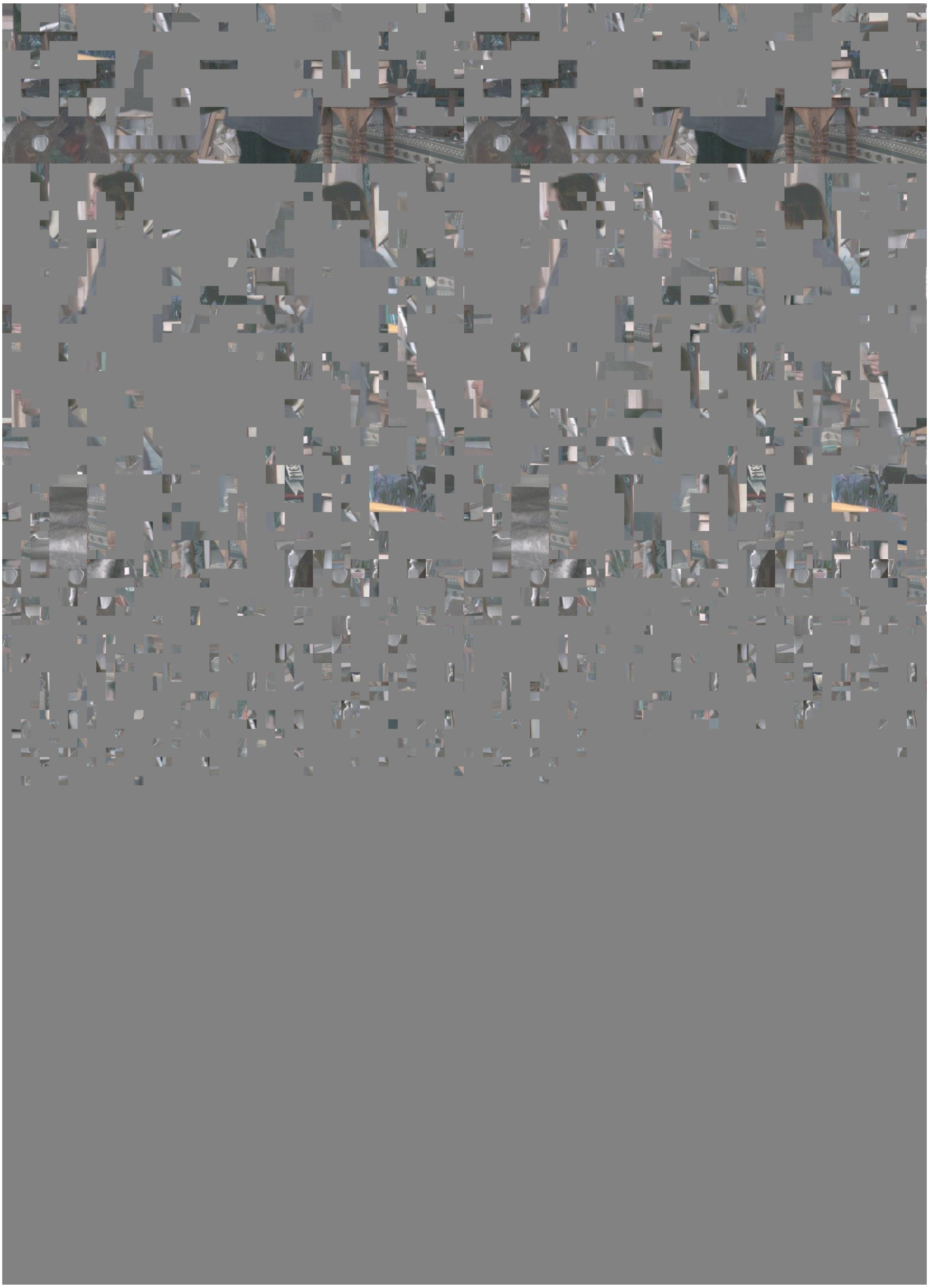

Fig. 3. Proposed: A97, SD, atlas 1, frame 15 (left) and 16 (right).

3 Syntax and semantics

8.1.1.2 Atlas sequence parameter set MIV extension syntax

asps_miv_extension( ) {	**Descriptor**
asme_group_idx	u(v)
asme_auxiliary_atlas_flag	u(1)
if(vme_embedded_occupancy_enabled_flag)	
asme_depth_occ_threshold_flag	u(1)
if(vme_geometry_scale_enabled_flag) {	
asme_geometry_scale_factor_x_minus1	ue(v)
asme_geometry_scale_factor_y_minus1	ue(v)
}	
if(!vme_embedded_occupancy_enabled_flag && vme_occupancy_scale_enabled_flag) {	
asme_occupancy_scale_factor_x_minus1	ue(v)
asme_occupancy_scale_factor_y_minus1	ue(v)
}	
asme_patch_constant_depth_flag	u(1)
asme_patch_attribute_offset_enabled_flag	u(1)
if(asme_patch_attribute_offset_enabled_flag)	
asme_patch_attribute_offset_bit_depth_minus1	ue(v)
asme_max_num_of_temporal_patch_splits	u(v)
}	

8.2.1.2 Atlas sequence parameter set MIV extension semantics

asme_max_num_of_temporal_patch_splits specifies a limit of number of possible temporal splits of a single patch. The number of bits used for the representation of asme_max_num_of_temporal_patch_splits is Ceil(Log2(*intraPeriod*)).

8.1.1.5 Patch data unit MIV extension syntax

pdu_miv_extension(tileID, p) {	**Descriptor**
if(vme_max_entity_id > 0)	
pdu_entity_id[tileID][p]	u(v)
if(asme_depth_occ_threshold_flag)	
pdu_depth_occ_threshold[tileID][p]	u(v)
if(asme_patch_attribute_offset_enabled_flag)	
for(c = 0; c < 3; c++)	
pdu_attribute_offset[tileID][p][c]	u(v)
if(asme_max_num_of_temporal_patch_splits > 0)	
pdu_temp_split_flag[tileID][p]	u(1)
if(pdu_temp_split_flag[tileID][p])	
pdu_temp_split_from[tileID][p]	u(v)
pdu_temp_split_to[tileID][p]	u(v)
}	

8.2.1.5 Patch data unit MIV extension semantics

pdu_temp_split_flag[tileID][p] equal to 1 indicates that the patch with index equal to p in a tile with ID equal to tileID does not last for the entire intra period.

pdu_temp_split_from[tileID][p] specifies the number of the first frame when patch with index equal to p in a tile with ID equal to tile ID is active. The number of bits used for the representation of pdu_temp_split_from[tileID][p] is the same as for asme_max_num_of_temporal_patch_splits.

pdu_temp_split_to[tileID][p] specifies the number of the last frame when patch with index equal to p in a tile with ID equal to tile ID is active. The number of bits used for the representation of pdu_temp_split_to[tileID][p] is the same as for asme_max_num_of_temporal_patch_splits.

Decoding process changes

AtlasBlockToPatchMap is being updated for each frame.

4 Recommendations

We recommend to:

- Include the proposed technique into TMIV8,
- Adopt proposed syntax and semantics,
- Continue the MIV CE3.

5 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).