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/M54176
June 2020, Online

Source	**Poznań University of Technology (PUT), Poznań, Poland** **Electronics and Telecommunications Research Institute (ETRI), Daejeon, Republic of Korea**
Status	**Input**
Title	**Immersive Video CE1.2: Geometry scaling**
Author	Adrian Dziembowski*, Dawid Mieloch*, 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 scaling (Immersive Video CE1.2). In the proposed solution, the dynamic range of geometry is being scaled to the limited range: [64, 511].

2 Proposed technique

In the proposed approach, the dynamic range of geometry atlases is unified for all content. All the normalized disparity values are linearly scaled to the limited range: [64, 511]. Lower range (< 64) is used for occupancy signaling, range [512, 1023] is unused.

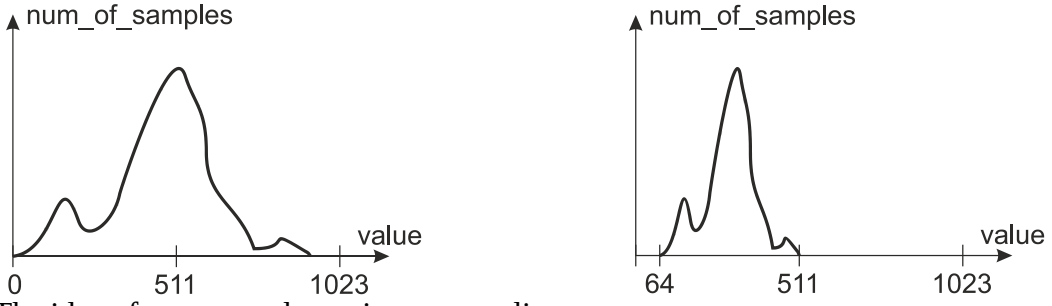

Fig. 1. The idea of geometry dynamic range scaling.

The motivation for such an approach was to decrease the total bitrate without reducing rendering quality, especially for natural content, where depth maps are not perfectly smooth.

Geometry atlases are being rescaled to the original range at the decoder side, thus information about geometry scaling has to be included in the bitstream. As presented in section 4, we propose to add three additional fields in the AtlasTileHeader:

- ath_depth_range_change_flag,
- ath_original_depth_range_start,
- ath_original_depth_range_end.

Parameters of geometry scaling are estimated once per GOP, separately for each atlas.

Encoder-side code changes:

```
if (ath_depth_range_change_flag) {
  for (int h = 0; h < depth.height(); h++) {
    for (int w = 0; w < depth.width(); w++) {
      if (depth(h, w) > 0) {
        depth(h, w) = (depth(h, w) + 0.5 – ath_original_depth_start) /
                      (ath_original_depth_range_end – ath_original_depth_start) * (511 – 64) + 64;
      }
    }
  }
}
```

Decoder-side code changes:

```
if (ath_depth_range_change_flag)) {
  for (int h = 0; h < depth.height(); h++) {
    for (int w = 0; w < depth.width(); w++) {
      if (depth(h, w) > expandedMin) {
        depth(h, w) = (depth(h, w) + 0.5 - 64) / (511 – 64) *
                      (ath_original_depth_range_end – ath_original_depth_start) + ath_original_depth_start;
      }
    }
  }
}
```

Example of geometry atlases with and without proposed geometry scaling are presented in Figs. 2 and 3.

Fig. 2. SP, geometry atlas: anchor (left) vs. proposal (right).

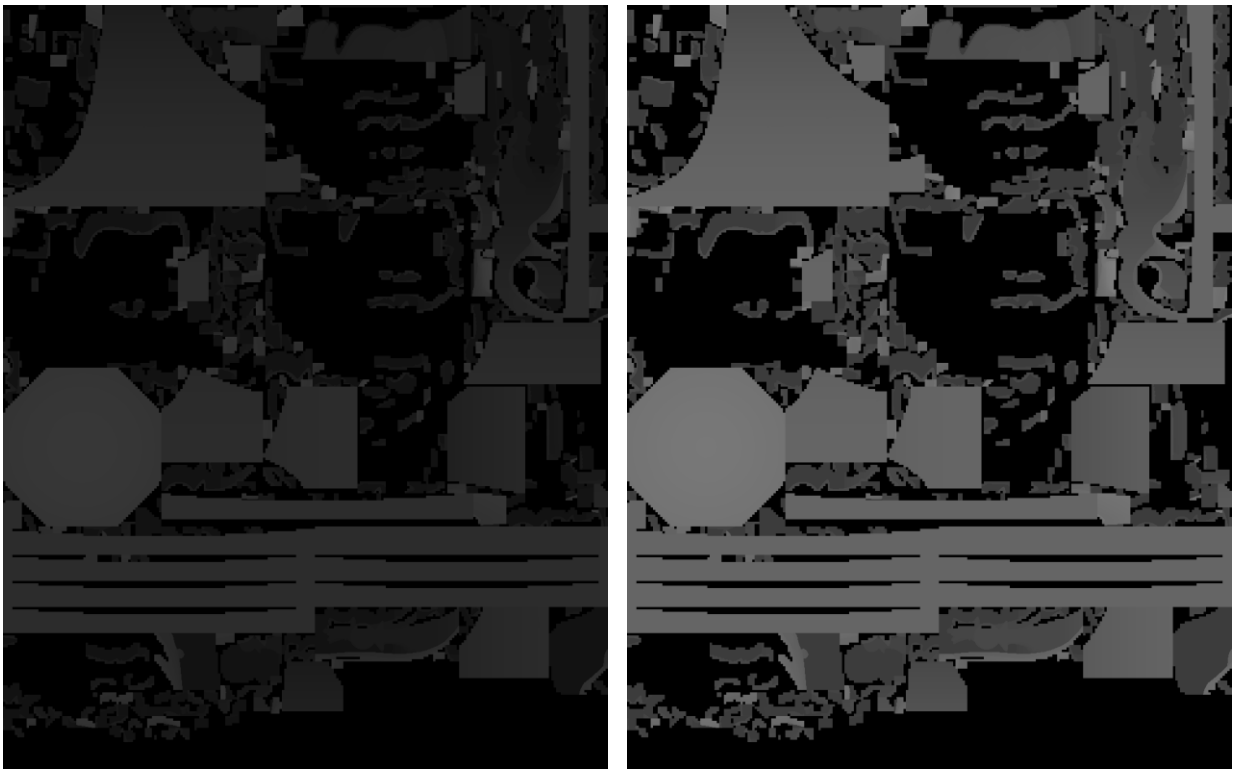

Fig. 3. SN, geometry atlas: anchor (left) vs. proposal (right).

3 Experimental results

Table 1. Depth range changed to [64, 511].

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	Pixel rate ratio
ClassroomVideo	SA	-3.4%	-8.9%	3.39	-3.0%	-9.9%	-6.8%	-10.6%	0.63
TechnicolorMuseum	SB	25.5%	16.0%	14.10	24.4%	12.1%	2.5%	1.4%	0.63
InterdigitalHijack	SC	-0.5%	0.9%	12.33	0.1%	1.2%	1.4%	1.9%	0.63
OrangeKitchen	SJ	20.4%	12.9%	15.26	30.4%	12.4%	-6.1%	-7.6%	0.62
TechnicolorPainter	SD	-31.8%	-33.6%	7.04	-31.7%	-33.6%	-32.5%	-34.2%	0.63
IntelFrog	SE	-0.4%	-2.6%	12.41	-0.2%	-3.4%	-4.9%	-5.2%	0.62
PoznanFencing	SL	-25.4%	-23.3%	15.06	-24.1%	-22.3%	-26.0%	-24.6%	0.52
MIV		**-2.2%**	**-5.5%**	**11.37**	**-0.6%**	**-6.2%**	**-10.3%**	**-11.3%**	

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

NokiaChess	SN	-14.7%	-7.7%	16.99	-16.0%	-3.8%	-5.3%	-2.6%	0.63
PoznanCarpark	SP	-19.3%	-20.8%	12.85	-20.9%	-21.8%	-21.9%	-22.8%	0.52
PoznanHall	ST	-11.2%	-10.9%	14.82	-11.4%	-11.1%	-11.9%	-11.4%	0.52
PoznanStreet	SU	-15.9%	-18.4%	12.98	-17.6%	-19.8%	-18.3%	-19.8%	0.52
MIV		**-15.3%**	**-14.4%**	**14.41**	**-16.5%**	**-14.1%**	**-14.3%**	**-14.1%**	

As presented in table above, proposed depth range changing allows to significantly improve BD-rates for all natural content. On the other hand, it decreases rendering quality for CG content because it destroys valid depth information (exception: SN, where range [64, 511] is wider than original depth dynamic range).

In order to preserve quality for CG sequences, depth dynamic range for such content may be expanded to wider range: [64, 1023]. For sequences where original depth range was wide enough (highest normalized disparity value higher than 900), depth expanding was skipped and ath_depth_range_change_flag set to 0.

Table 2. Depth range changed to [64, 1023].

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	Pixel rate ratio
ClassroomVideo	SA	0.7%	0.9%	3.38	0.5%	0.9%	1.4%	1.3%	0.63
TechnicolorMuseum	SB	-5.9%	-4.7%	14.40	-5.8%	-3.0%	0.5%	1.8%	0.63
InterdigitalHijack	SC	-8.1%	-1.1%	12.33	-6.3%	2.0%	8.4%	12.0%	0.63
OrangeKitchen	SJ	0.0%	0.0%	15.43	0.0%	0.0%	0.0%	0.0%	0.62
TechnicolorPainter	SD	0.0%	0.0%	7.01	0.0%	0.0%	0.0%	0.0%	0.63
IntelFrog	SE	18.7%	15.9%	23.53	21.2%	16.3%	22.7%	17.9%	0.62
PoznanFencing	SL	0.0%	0.0%	15.15	0.0%	0.0%	0.0%	0.0%	0.52
MIV		**0.8%**	**1.6%**	**13.03**	**1.4%**	**2.3%**	**4.7%**	**4.7%**	

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

NokiaChess	SN	-46.1%	-38.8%	18.36	-49.7%	-30.2%	-12.3%	-6.8%	0.63
PoznanCarpark	SP	0.4%	0.4%	12.88	0.5%	0.4%	0.4%	0.3%	0.52
PoznanHall	ST	34.9%	32.0%	14.80	34.9%	32.1%	35.7%	33.1%	0.52
PoznanStreet	SU	8.9%	9.1%	13.01	8.6%	9.0%	9.0%	9.0%	0.52
MIV		**-0.5%**	**0.7%**	**14.76**	**-1.4%**	**2.8%**	**8.2%**	**8.9%**	

The decision for changing depth range to [64, 511] or expanded range [64, 1023] may be taken based on vme_depth_low_quality_flag. Joint results are presented in Table 3.

Table 3. Depth range changed to [64, 511] or [64, 1023] based on low quality flag.

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	Pixel rate ratio
ClassroomVideo	SA	0.7%	0.9%	3.38	0.5%	0.9%	1.4%	1.3%	0.63
TechnicolorMuseum	SB	-5.9%	-4.7%	14.40	-5.8%	-3.0%	0.5%	1.8%	0.63
InterdigitalHijack	SC	-8.1%	-1.1%	12.33	-6.3%	2.0%	8.4%	12.0%	0.63
OrangeKitchen	SJ	0.0%	0.0%	15.43	0.0%	0.0%	0.0%	0.0%	0.62
TechnicolorPainter	SD	-31.8%	-33.6%	7.04	-31.7%	-33.6%	-32.5%	-34.2%	0.63
IntelFrog	SE	-0.4%	-2.6%	12.41	-0.2%	-3.4%	-4.9%	-5.2%	0.62
PoznanFencing	SL	-25.4%	-23.3%	15.06	-24.1%	-22.3%	-26.0%	-24.6%	0.52
MIV		**-10.1%**	**-9.2%**	**11.44**	**-9.6%**	**-8.5%**	**-7.6%**	**-7.0%**	

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

NokiaChess	SN	-46.1%	-38.8%	18.36	-49.7%	-30.2%	-12.3%	-6.8%	0.63
PoznanCarpark	SP	-19.3%	-20.8%	12.85	-20.9%	-21.8%	-21.9%	-22.8%	0.52
PoznanHall	ST	-11.2%	-10.9%	14.82	-11.4%	-11.1%	-11.9%	-11.4%	0.52
PoznanStreet	SU	-15.9%	-18.4%	12.98	-17.6%	-19.8%	-18.3%	-19.8%	0.52
MIV		**-23.1%**	**-22.2%**	**14.75**	**-24.9%**	**-20.7%**	**-16.1%**	**-15.2%**	

Fig. 4. Anchor vs. proposal at similar bitrate, synthesized input views.

4 Syntax and semantics

atlas_tile_header( ) {	**Descriptor**
ath_atlas_frame_parameter_set_id	ue(v)
ath_atlas_adaptation_parameter_set_id	ue(v)
ath_id	u(v)
ath_type	ue(v)
if(afps_output_flag_present_flag)	
ath_atlas_output_flag	u(1)
ath_atlas_frm_order_cnt_lsb	u(v)
if(asps_num_ref_atlas_frame_lists_in_asps > 0)	
ath_ref_atlas_frame_list_sps_flag	u(1)
if(ath_ref_atlas_frame_list_sps_flag == 0)	
ref_list_struct(asps_num_ref_atlas_frame_lists_in_asps)	
else if(asps_num_ref_atlas_frame_lists_in_asps > 1)	
ath_ref_atlas_frame_list_idx	u(v)
for(j = 0; j < NumLtrAtlasFrmEntries; j++) {	
ath_additional_afoc_lsb_present_flag[j]	u(1)
if(ath_additional_afoc_lsb_present_flag[j])	
ath_additional_afoc_lsb_val[j]	u(v)
}	
if(ath_type != SKIP_TILE) {	
if(asps_normal_axis_limits_quantization_enabled_flag) {	
ath_pos_min_z_quantizer	u(5)
if(asps_normal_axis_max_delta_value_enabled_flag)	
ath_pos_delta_max_z_quantizer	u(5)
}	
if(asps_patch_size_quantizer_present_flag) {	
ath_patch_size_x_info_quantizer	u(3)
ath_patch_size_y_info_quantizer	u(3)
}	
if(afps_raw_3d_pos_bit_count_explicit_mode_flag)	
ath_raw_3d_pos_axis_bit_count_minus1	u(v)
if(ath_type == P_TILE && num_ref_entries[RlsIdx] > 1) {	
ath_num_ref_idx_active_override_flag	u(1)
if(ath_num_ref_idx_active_override_flag)	
ath_num_ref_idx_active_minus1	ue(v)
}	
ath_depth_range_change_flag	u(1)
if(ath_depth_range_change_flag) {	
ath_original_depth_range_start	u(16)
ath_original_depth_range_end	u(16)
}	
}	
byte_alignment( )	
}	

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).

6 Recommendations

We recommend to:

- include proposed technique into TMIV6,
- adopt proposed syntax and semantics,
- continue the Core Experiment 1.

7 References

[N19215] "Description of Immersive Video Core Experiments 1: Bitstream Adaptation" ISO/IEC JTC1/SC29/WG11 MPEG/N19215, May 2020, Alpbach, Austria.