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

ISO/IEC JTC1/SC29/WG04 MPEG/M56950
April 2021, Online

Source	Orange, PUT, ETRI
Status	Input document
Title	[MIV] Combination of m56626 and m56335 for Geometry Assistance SEI message
Author	Gordon Clare, Patrick Garus, Félix Henry (Orange) Błażej Szydełko, Dawid Mieloch, Adrian Dziembowski, Marek Domański (PUT), Gwangsoon Lee, Jun Young Jeong (ETRI)

Abstract

This contribution presents the SEI syntax from m56626 modified by adding the handling of rectangular block splitting from m56335, and a 1Mb/s limit to the amount of SEI message. The speed up of the decoder runtime is similar to that of m56626 and it is asserted that the added support for rectangular blocks improves the synthesis BDRate.

1 Introduction

During the surrent MPEG 134 meeting, the following contributions were examined:

- m56626, which proposes to send depth related features as SEI messages to help the depth estimator reduce its search space in the Geometry Absent profile, and uses square blocks to convey this information locally
- m56335, which adds the support for rectangular blocks while using a different syntax for convey the features.

During the presentation of these contributions, the group expressed the desire to examine the performance of a combined contribution with the follow aspects:

- Use the syntax of m56626
- Add support syntax for handling rectangular blocks of m56335
- Limit the amount of SEI message being sent to 1Mb/s

This contribution presents this combination and its performance, and proposes to adopt the corresponding SEI message.

2 Experimental results

2.1 G17 anchor vs. G17 with proposed SEI

2.1.1 Compression performance

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	A	-58.4%	-46.5%	5.75	-64.4%	-54.6%	-25.9%	-16.8%
Museum	B	---	-62.6%	14.27	-55.7%	-33.5%	-21.3%	-10.8%
Fan	O	17.5%	16.3%	10.71	45.5%	34.2%	15.7%	16.8%
Kitchen	J	6.3%	9.8%	13.10	4.0%	5.4%	11.4%	13.4%
Painter	D	-45.4%	-22.4%	7.06	-20.2%	-1.0%	-41.4%	-21.4%
Frog	E	2.2%	10.3%	6.87	4.8%	12.2%	9.9%	15.5%
Carpark	P	-54.9%	-29.8%	9.77	-41.4%	-22.3%	-59.7%	-35.4%
Chess	N	---	---	24.06	---	253.5%	---	---
Group	R	---	---	13.56	-71.4%	-49.0%	---	---
MIV		**---**	**---**	**11.68**	**---**	**16.1%**	**---**	**---**

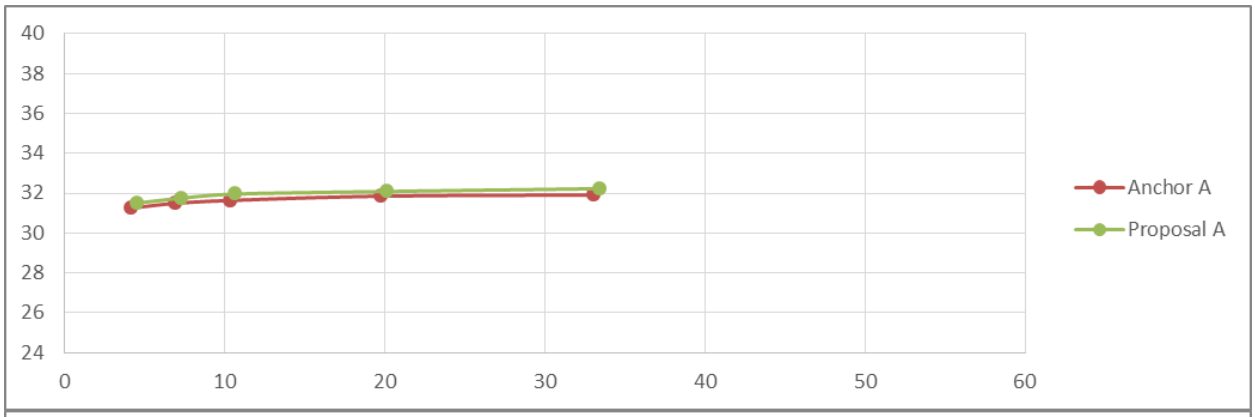
Optional content - Proposal vs. Low/High-bitrate Anchors

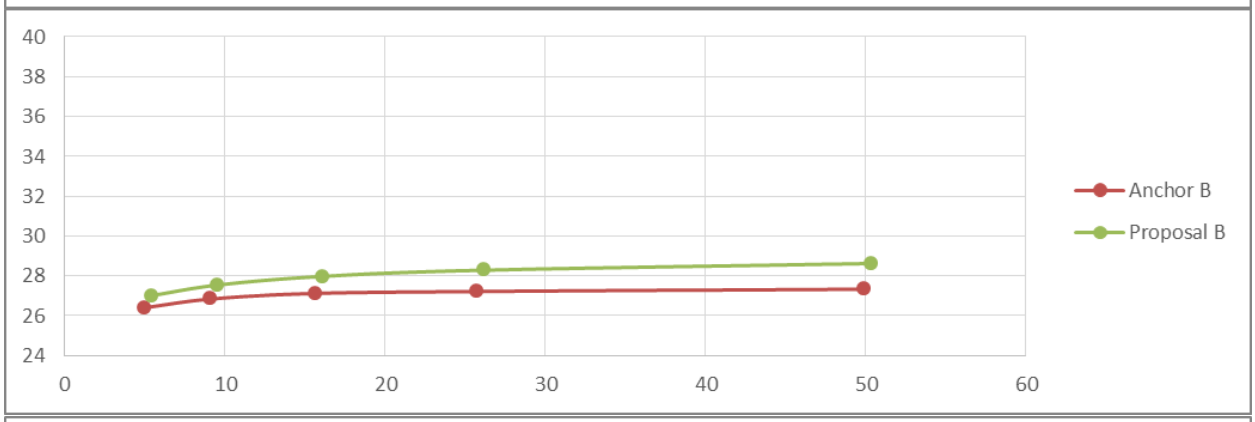
Fencing	L	-34.6%	-7.2%	13.13	-32.3%	-0.3%	-0.0%	8.9%
Hall	T	---	---	17.92	-71.8%	-38.8%	---	---
Street	U	27.6%	24.0%	8.20	1.1%	12.2%	-5.0%	6.0%
ChessPieces	Q	-73.6%	---	31.67	---	---	-18.6%	-12.1%
Hijack	C	---	---	21.14	-56.2%	-35.7%	---	---
Mirror	I	---	228.0%	14.86	---	187.6%	---	346.4%
MIV		**---**	**---**	**17.82**	**---**	**---**	**---**	**---**

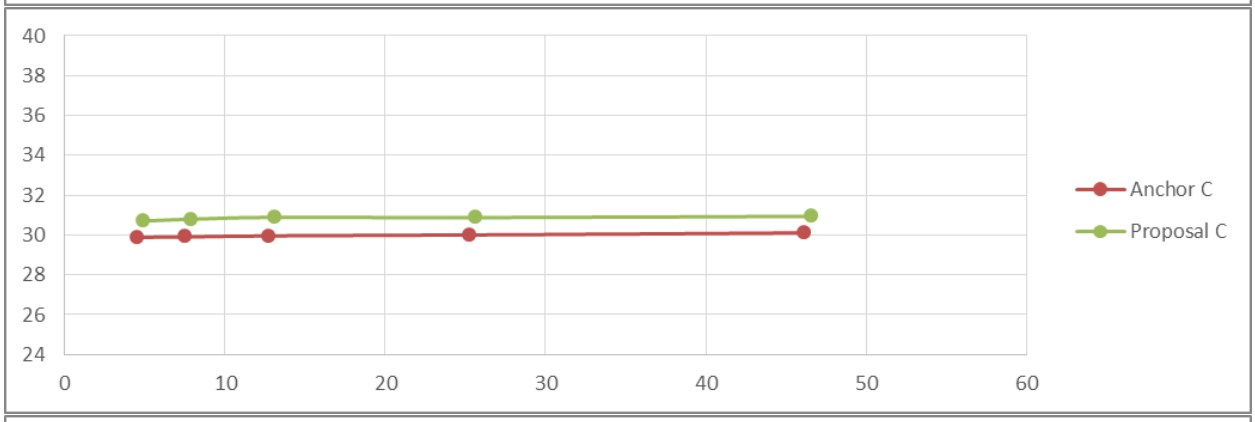
The following figure shows the PSNR curves for all sequences, mandatory and optional. Among the 15 sequences:

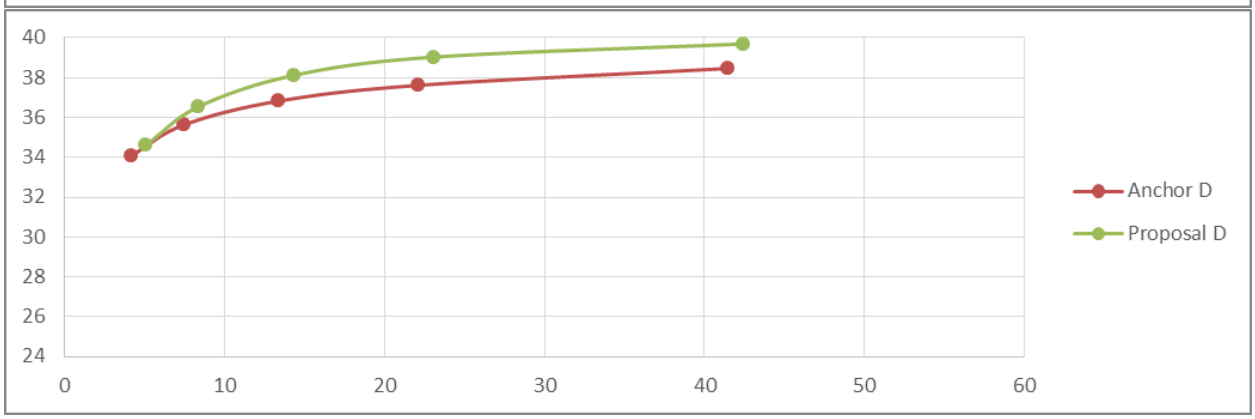
- 6 show a PSNR improvement of 1dB or more
- 8 show a similar PSNR curve
- 1 shows a loss of 1 dB or more (sequence I, due to the wrong depth in the mirror)

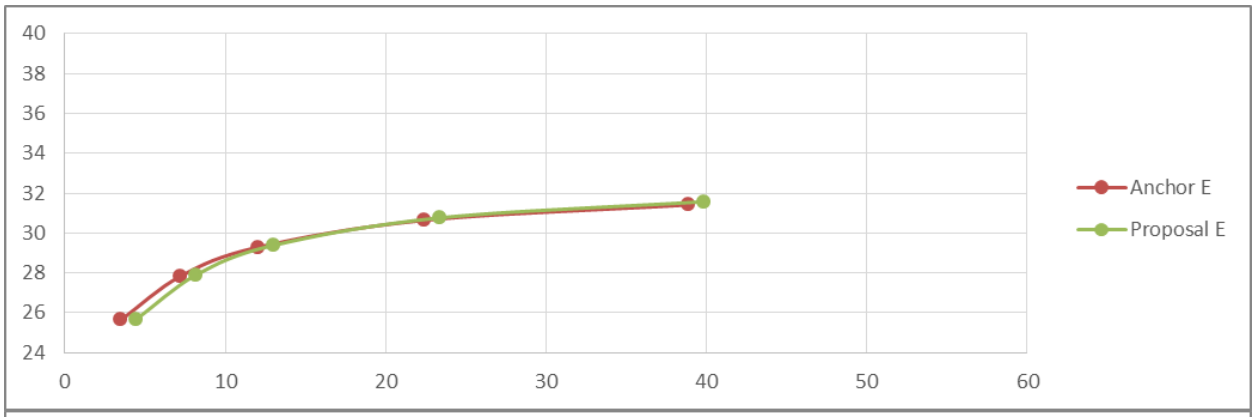
It is asserted that this is an improvement over m56626.

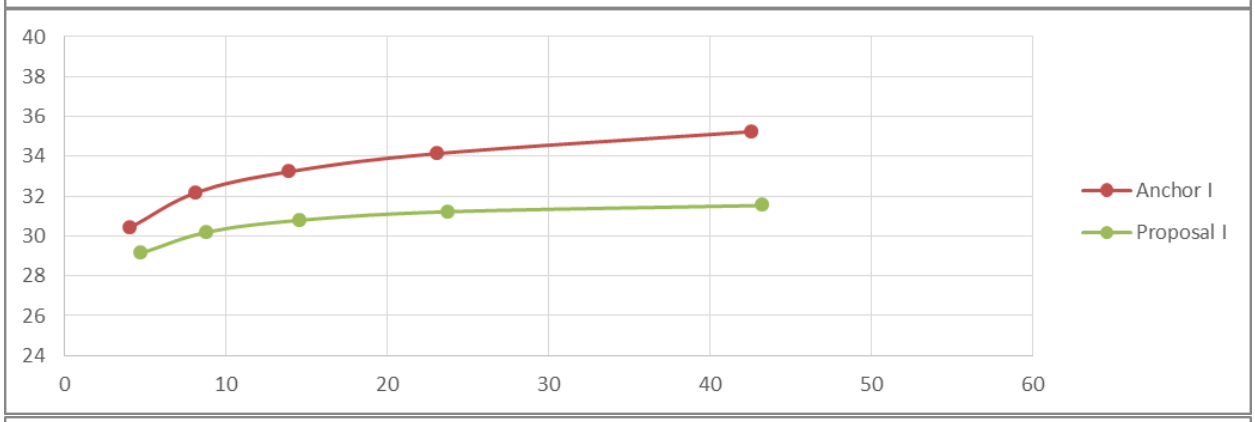

40
38
36
34
32
30
28
26
24
0
10
20
30
40
50
60
Anchor A
Proposal A

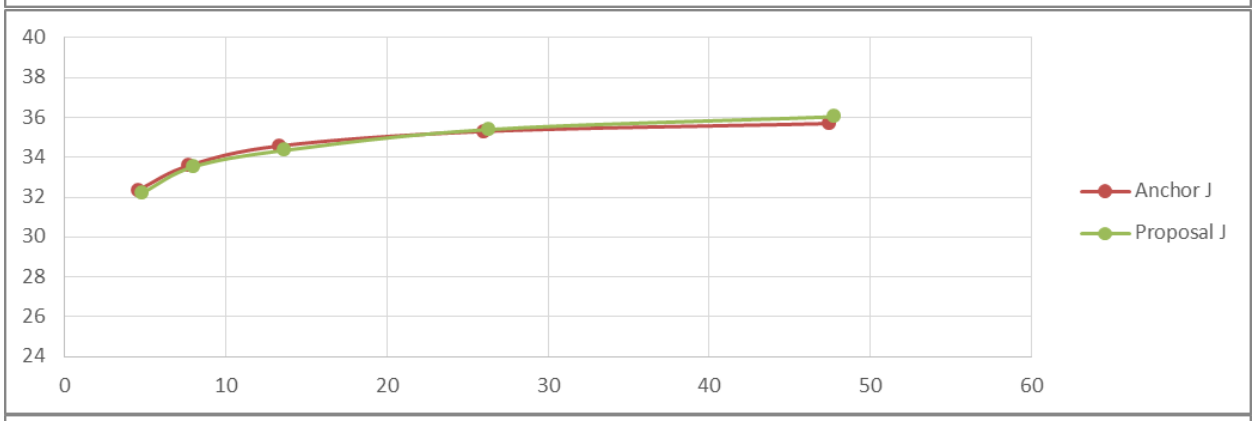

40
38
36
34
32
30
28
26
24
0
10
20
30
40
50
60
Anchor B
Proposal B

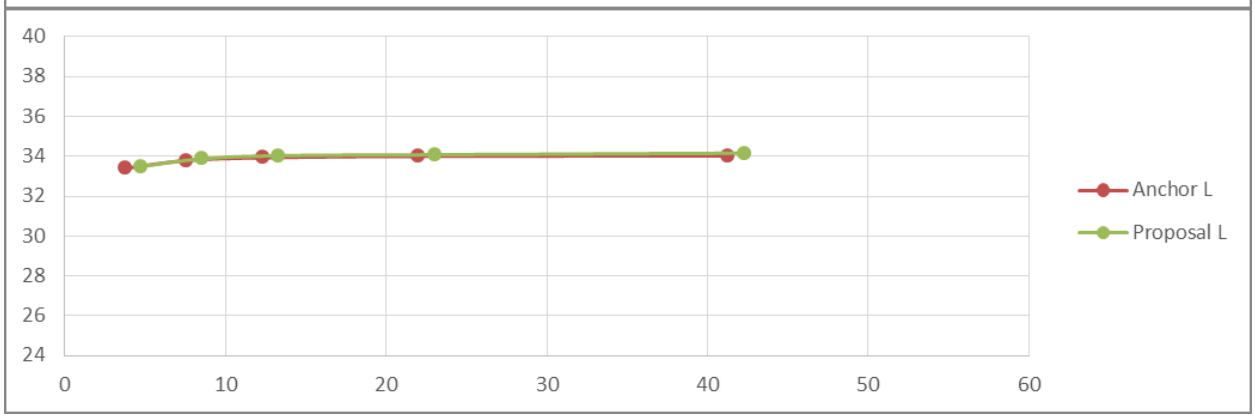

40
38
36
34
32
30
28
26
24
0
10
20
30
40
50
60
Anchor C
Proposal C

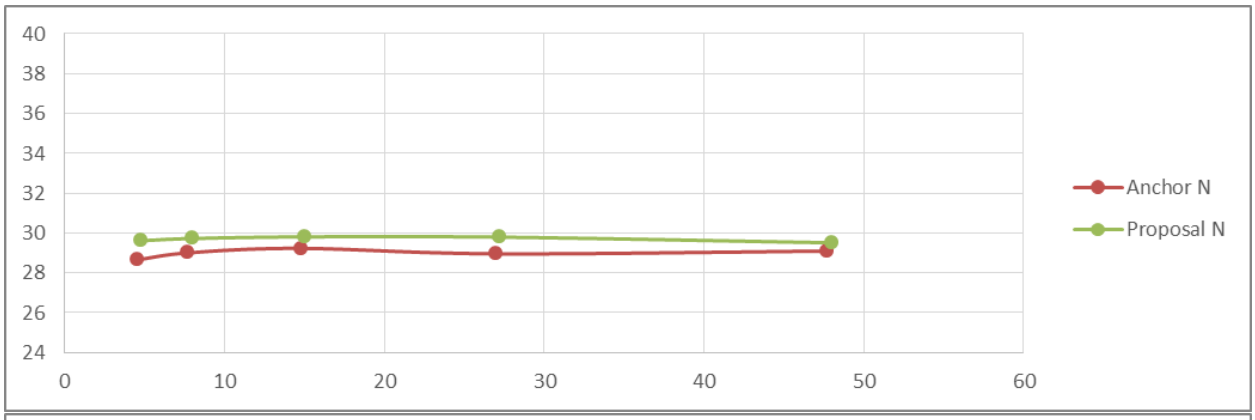

40
38
36
34
32
30
28
26
24
0
10
20
30
40
50
60
Anchor D
Proposal D

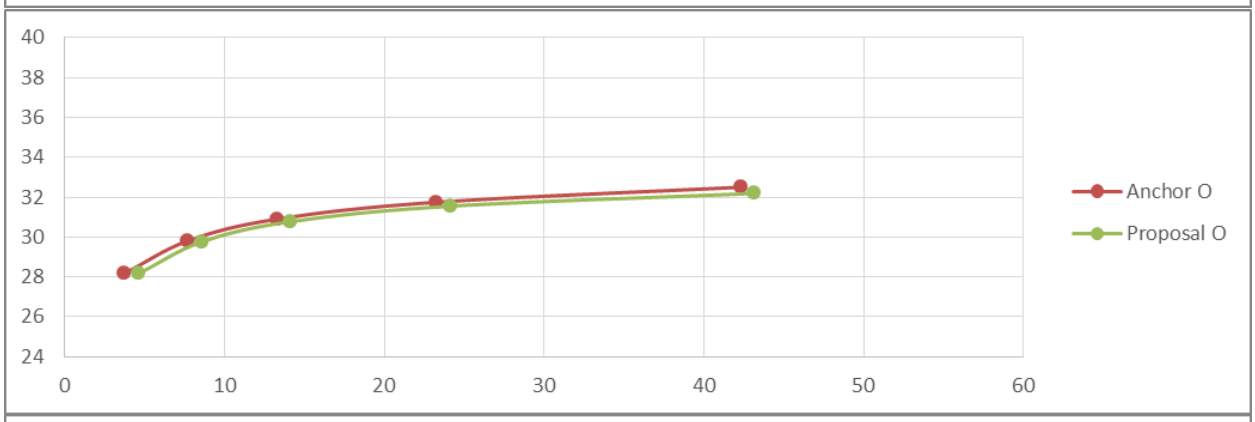

40
38
36
34
32
30
28
26
24
0
10
20
30
40
50
60
Anchor E
Proposal E

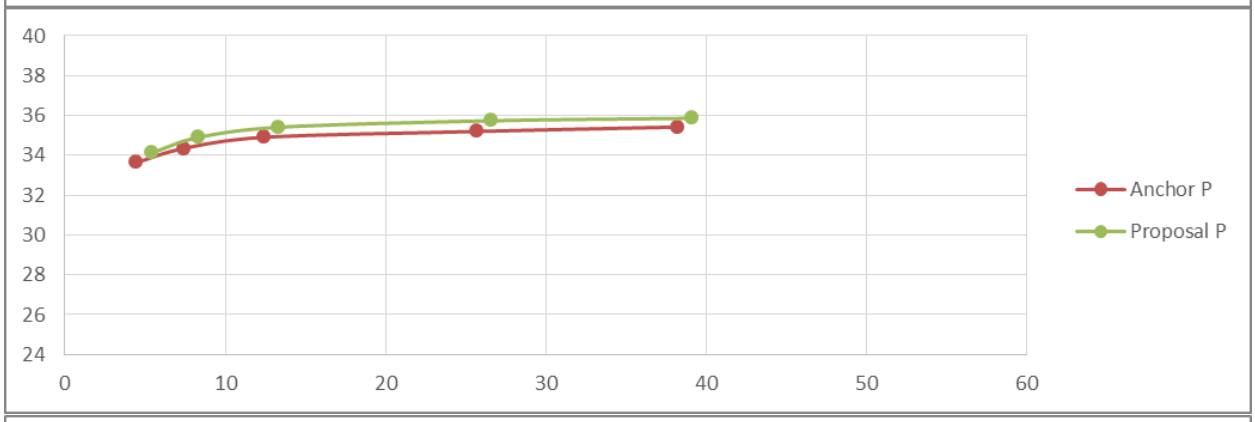

40
38
36
34
32
30
28
26
24
0
10
20
30
40
50
60
Anchor I
Proposal I

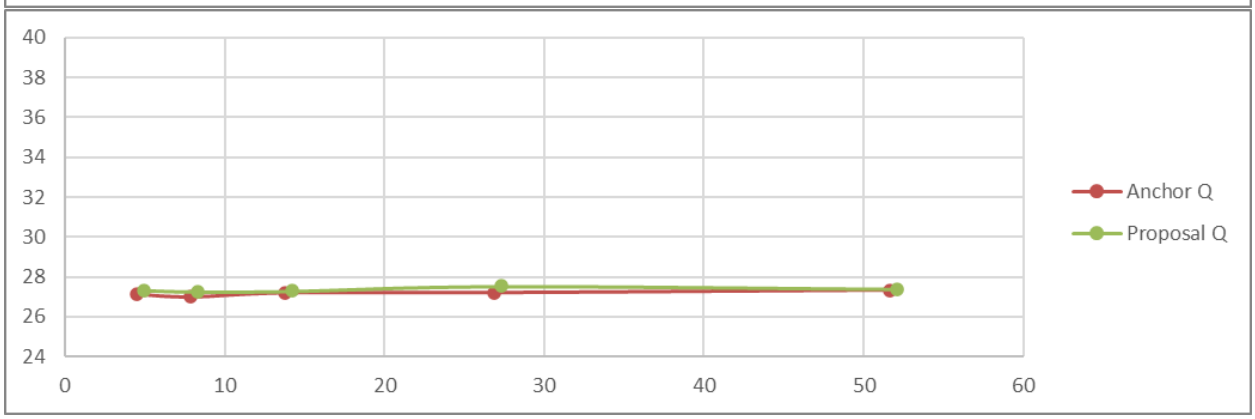

40
38
36
34
32
30
28
26
24
0
10
20
30
40
50
60
Anchor J
Proposal J

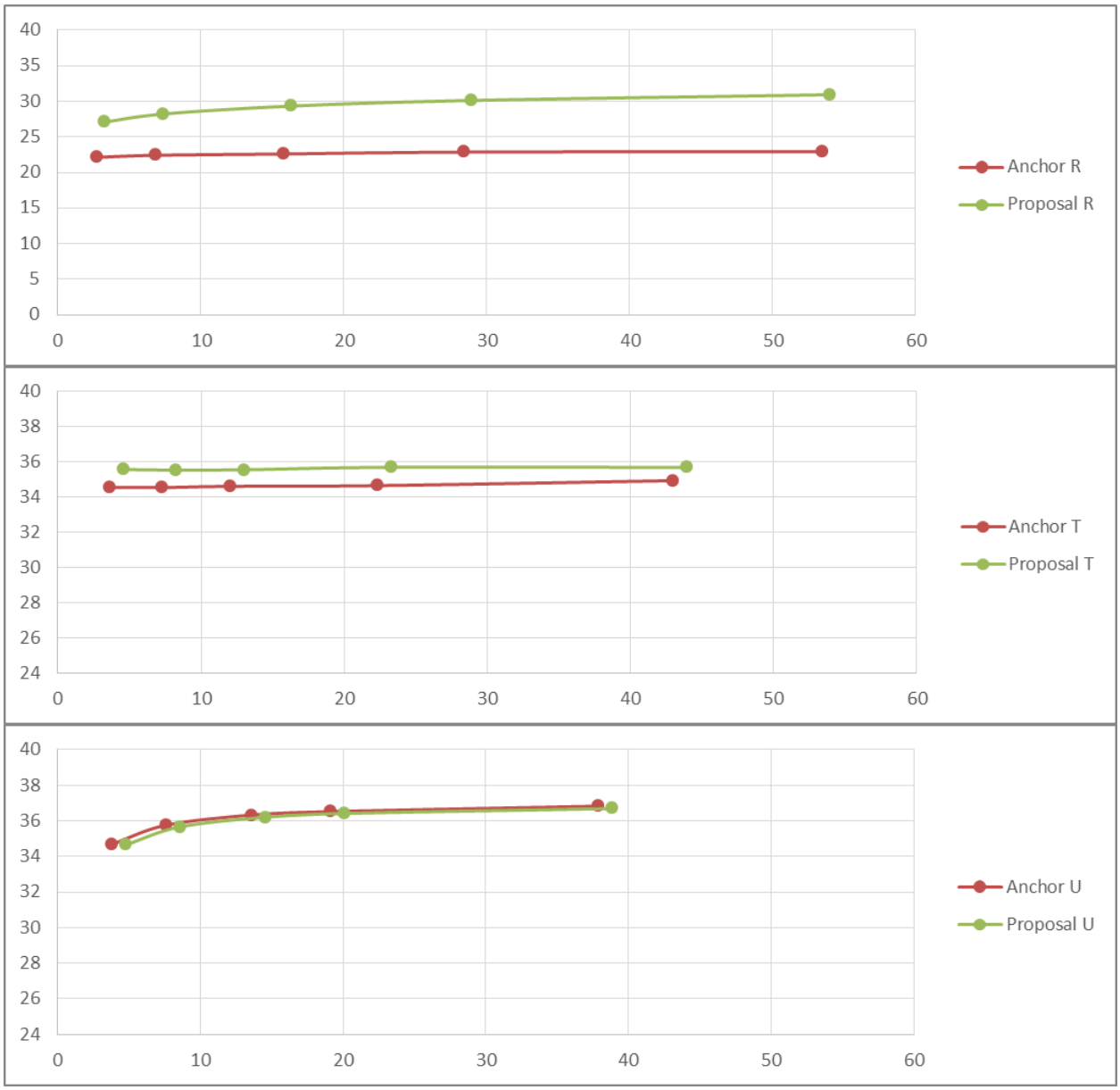

40
38
36
34
32
30
28
26
24
0
10
20
30
40
50
60
Anchor L
Proposal L


40
38
36
34
32
30
28
26
24
0
10
20
30
40
50
60
Anchor N
Proposal N


40
38
36
34
32
30
28
26
24
0
10
20
30
40
50
60
Anchor O
Proposal O


40
38
36
34
32
30
28
26
24
0
10
20
30
40
50
60
Anchor P
Proposal P


40
38
36
34
32
30
28
26
24
0
10
20
30
40
50
60
Anchor Q
Proposal Q



2.1.2 Runtime performance

The following table shows the runtime change in the decoder (time spent at the decoder when compared to G17 anchor):

Mandatory content

Sequence		TMIV decoding (incl. IVDE)	Depth map estimation
ClassroomVideo	A	30.8%	24.2%
Museum	B	63.1%	57.9%
Fan	O	46.2%	43.5%
Kitchen	J	67.4%	64.2%
Painter	D	49.0%	46.5%
Frog	E	23.3%	22.3%
Carpark	P	71.3%	69.1%
Chess	N	78.1%	73.7%
Group	R	56.5%	53.6%
MIV		54.0%	50.6%

Optional content

Fencing	L	28.8%	26.2%
Hall	T	44.5%	42.9%
Street	U	49.3%	47.8%
ChessPieces	Q	83.8%	83.2%
Hijack	C	78.7%	63.8%
Mirror	I	59.3%	55.8%
MIV		57.4%	53.3%

2.1.3 Bitrate limitation of SEI message

It was the intention of the group to evaluate the performance of the proposed SEI message under the bitrate condition of 1 Mb/s.
This is the size of encoded features for all sequences (with the quantization step used for the depth interval boundaries):

```
/A/FEATURESENC.txt qp=1:    kbps: 358.969 at 30fps
/B/FEATURESENC.txt qp=1:    kbps: 439.115 at 30fps
/C/FEATURESENC.txt qp=1:    kbps: 392.329 at 30fps
/D/FEATURESENC.txt qp=1024: kbps: 938.456 at 30fps
/E/FEATURESENC.txt qp=256:  kbps: 973.228 at 30fps
/I/FEATURESENC.txt qp=1:    kbps: 657.431 at 30fps
/J/FEATURESENC.txt qp=1:    kbps: 287.562 at 30fps
/L/FEATURESENC.txt qp=4:    kbps: 996.788 at 25fps
/N/FEATURESENC.txt qp=1:    kbps: 240.635 at 30fps
/O/FEATURESENC.txt qp=1:    kbps: 857.04  at 30fps
/P/FEATURESENC.txt qp=8:    kbps: 896.553 at 25fps
/Q/FEATURESENC.txt qp=1:    kbps: 446.513 at 30fps
/R/FEATURESENC.txt qp=1:    kbps: 542.922 at 30fps
/T/FEATURESENC.txt qp=4:    kbps: 965.8   at 25fps
/U/FEATURESENC.txt qp=2:    kbps: 970.318 at 25fps
```

2.2 *Split type statistics*

split type / sequence								
SA	73%	7%	14%	2%	2%	2%	1%	1%
SB	38%	19%	13%	3%	3%	14%	5%	5%
SC	50%	11%	6%	1%	1%	22%	4%	4%
SD	50%	15%	5%	2%	2%	12%	7%	7%
SE	60%	11%	7%	4%	4%	8%	4%	4%
SI	52%	13%	9%	2%	2%	11%	6%	6%
SJ	30%	13%	24%	4%	4%	17%	4%	4%
SL	58%	9%	7%	2%	2%	12%	5%	5%
SN	78%	2%	8%	2%	2%	3%	2%	2%
SO	34%	31%	7%	2%	2%	14%	5%	5%
SP	36%	15%	11%	12%	12%	7%	3%	3%
SR	31%	22%	5%	2%	2%	22%	8%	8%
ST	49%	10%	11%	3%	3%	15%	5%	5%
SU	44%	11%	18%	9%	9%	5%	2%	2%
average	49%	13%	10%	4%	4%	12%	4%	4%

According to non-uniform probability of different split types, number of bits used for signaling of different split types should be different:

split type / flag								
gas_split_flag	0	1	1	1	1	1	1	1
gas_quad_split_flag		1	0	0	0	0	0	0
gas_split_orientation_flag			0	0	0	1	1	1
gas_split_symmetry_flag			1	0	0	1	0	0
gas_split_first_block_bigger				0	1		0	1

3 Syntax

geometry_assistance(payloadSize) {	**Descriptor**
gas_qs	ue(v)
gas_bw	ue(v)
for(v = 0; v <= mvp_num_views_minus1; v++) {	
viewID = ViewIndexToID[v]	
for(l = 0; l < (ci_projection_plane_height_minus1[viewID]+gas_bw)/gas_bw; l++) {	
for(c = 0; c < (ci_projection_plane_width_minus1[viewID] +gas_bw)/gas_bw; c++) {	
gas_split_flag	u(1)
subBlocksHorizontally = 1	
subBlocksVertically = 1	
if(gas_split_flag) {	

gas_quad_split_flag	u(1)
if(gas_quad_split_flag) {	
subBlocksHorizontally = 2	
subBlocksVertically = 2	
} else {	
gas_split_orientation_flag	u(1)
if(gas_split_orientation_flag) {	
subBlocksHorizontally = 2	
} else {	
subBlocksVertically = 2	
}	
gas_split_symmetry_flag	u(1)
if(!gas_split_symmetry_flag) {	
gas_split_first_block_bigger	u(1)
}	
}	
}	
for(sbl = 0; sbl < subBlocksVertically; sbl++) {	
for(sbc = 0; sbc < subBlocksHorizontally; sbc++) {	
gas_skip_flag	u(1)
if(gas_skip_flag) {	
if(l==0 && c==0&& sbl ==0&& sbc ==0) { /*None*/	
LTMinFlag = 2	
LTMaxFlag = 2	
}	
else if(l==0 && sbl==0) /*Left*/	
LTMinFlag = 0	
LTMaxFlag = 0	
}	
else if(c==0 && sbc==0)/*Top*/	
LTMinFlag = 1	
LTMaxFlag = 1	
}	
else {	
gas_ltmin_flag	u(1)
gas_ltmax_flag	u(1)
LTMinFlag = gas_ltmin_flag	
LTMaxFlag = gas_ltmax_flag	
}	
zmin_delta	ue(v)
zmax delta	ue(v)
}	
}	
}	
}	

}	
}	
}	

4 Semantics

The geometry assistance SEI message indicates suggested information associated with each view that can be used by a depth estimation process occurring at decoder side. This SEI message can be useful when the Geometry Absent profile is used. Its main purpose is to reduce the complexity of the decoder side depth estimation process.

A view is uniformly divided into square blocks, and each block can be further divided once into smaller subblocks of square or rectangular shapes. The top left corner of the first block coincides with the top left corner of a view. One or more syntax elements are associated with each block. The syntax may indicate that the block is skipped, thereby suggesting that the depth of the current block does not need to be updated. If the syntax indicates that the current block is not skipped, the remaining syntax indicates the suggested minimum and maximum depths present in the block. A depth estimation process can take benefit of this information to reduce the search space of depth candidates, and to avoid producing depths that are outside of the suggested range. Table X indicates the available split types (including no split) of a block, with the associated values of the block division syntax. Figure Y shows an example of the block subdivision of a view.

Table X: The different block split types and associated values of the syntax elements

split type / flag								
gas_split_flag	0	1	1	1	1	1	1	1
gas_quad_split_flag		1	0	0	0	0	0	0
gas_split_orientation_flag			0	0	0	1	1	1
gas_split_symmetry_flag			1	0	0	1	0	0
gas_split_first_block_bigger				0	1		0	1

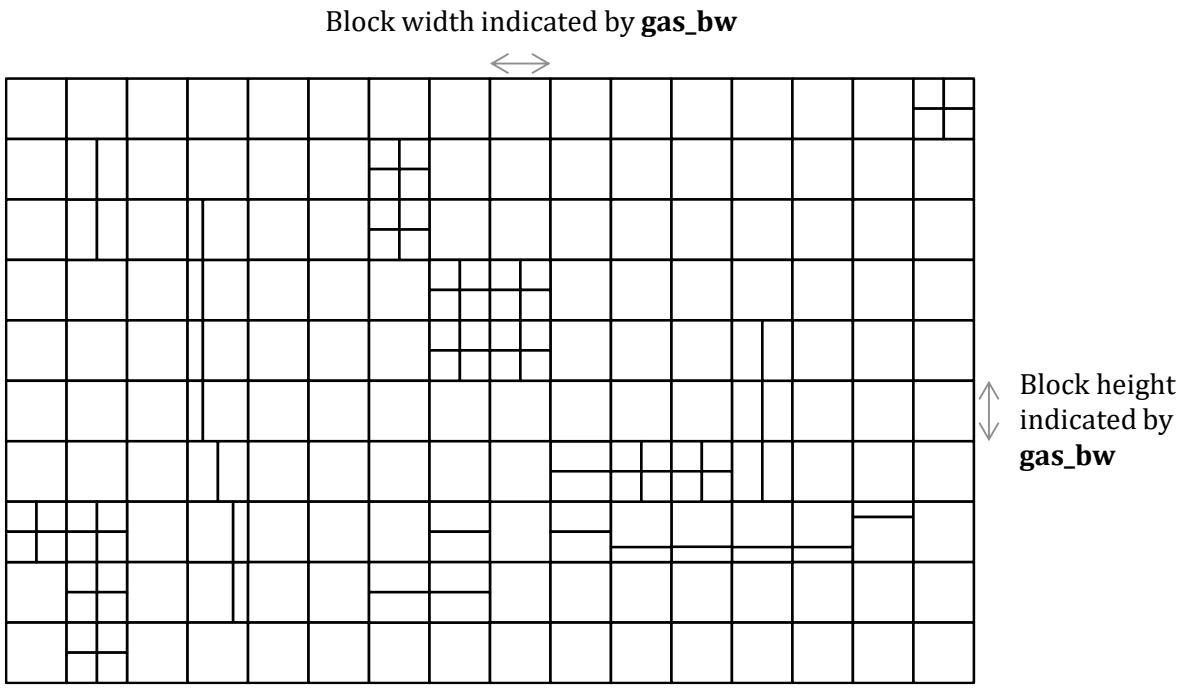


Figure Y: example of a partition of a view into all possible block divisions

gas_sq specifies the quantization step to be used for the suggested depth range boundaries. The variable QStep is set to the value specified by gas_sq.

gas_bw specifies the block width to be used on the view identified by the variable ViewID. The value specified by gas_bw can only be an integer power of 2 and must be larger than 4.

gas_split_flag specifies if the current block is split into smaller subblocks.

gas_quad_split_flag equal to 0 signals that the current block is split into two rectangular subblocks. When it is equal to 1, the current block is split into four square subblocks of identical sizes.

gas_split_orientation_flag equal to 0 signals that the current block is split horizontally. When it is equal to 1, the current block is split vertically.

gas_split_symmetry_flag equal to 0 signals that the area of the two sublocks is different. When it is equal to 1, the area of the two sublocks is equal.

gas_split_first_block_bigger equal to 0 signals that the first subblock (top subblock if gas_split_orientation_flag is equal to 0, and left subblock if gas_split_orientation_flag is equal to 1) is bigger than the second subblock. When it is equal to 1, the first subblock is smaller than the second subblock.

gas_skip_flag equal to 0 signals that a zmin_delta and a zmxa_delta syntax elements are present in the bitstream, and that a gas_ltmin_flag and a gas_ltmax_flag may be present. When it is equal to 1, it signals that no other syntax elements are present in the bitstream for the current block, and it suggests that the geometry information in this block has not changed since the previous frame in display order.

gas_ltmin_flag equal to 0 signals that the prediction of the current minimum depth is to be taken from the left block, otherwise from the top block.

gas_ltmax_flag equal to 0 signals that the prediction of the current maximum depth is to be taken from the left block, otherwise from the top block.

zmin_delta specifies the remainder to be added to the prediction to obtain the minimum depth value suggested for the current block.

zmax_delta specifies the remainder to be added to the prediction to obtain the maximum depth value suggested for the current block

Variables ZMinLeft and ZMaxLeft are set to the minimum and maximum depth range of the left block, respectively, and if available

Variables ZMinTop and ZMaxTop are set to the minimum and maximum depth range of the top block, respectively, and if available.

The suggested minimum depth range ZMin and maximum depth range ZMax of the current block are derived by the following equantions

ZMin = (LTMinFlag == 2 ? 0 : LTMinFlag == 1 ? ZMinTop : ZMinLeft) + QStep*zmindelta

ZMax = (LTMaxFlag == 2 ? 0 : LTMaxFlag == 1 ? ZMinTop : ZMinLeft) + QStep*zmaxdelta

5 Recommendations

We propose a new SEI message which helps a decoder-side depth estimation process to reduce the search space and thus reduces the runtime of the decoder in the Geometry Absent profile. We suggest that the SEI message is adopted in the MIV standard.

6 References

[m56335] Rectangular blocks in encoder-derived features for decoder-side depth estimation, B. Szydełko, D. Mieloch, A. Dziembowski (PUT), G. Lee, J.Y. Jeong (ETRI).

[m56626] [MIV] Geometry Assistance SEI message
G. Clare, P. Garus, F. Henry (Orange)