INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11 MPEG129/m51560
Jan 2020, Brussel, Belgium

Source	**Intel, Philips, ETRI, PUT, InterDigital**
Status	**Input document**
Title	**MIV anchors and crosscheck**
Author	Basel Salahieh, Bart Sonneveldt, Gwangsoon Lee, Adrian Dziembowski, Renaud Doré

1 Introduction

The Anchor iteration includes a new synthetic content NokiaChess CG1-N, bringing the number of contents to 8.

The *Common Test Conditions for Immersive Video* [1] specify four anchors:

a. MIV anchors categorized as A97 when computed on 97 frames,
b. MIV anchors categorized as A17 when computed on 17 frames,
c. MIV view anchors categorized as V17 because computed on 17 frames,
d. MV-HEVC anchors.

Only a) b) and c) are reported here. The Excel template file with “ff” radical (full frame) relates to A97 category, while the Excel template file with “rf” radical (reduced frame) relates to A17 and V17 categories. The purpose of A17 is to compare with V17.

The MIV anchors (for objective evaluation) are based on TMIV 3.0 announced on the MPEG-I Visual reflector on November 8, 2019.

We have divided the work to generate the MIV anchors among 5 participants according to Table 1.

Table 1: Division of anchor generation work

Sequence	Short name	Anchor	Anchor by	Crosscheck by
CG1 – A	SA	a), b) & c)	Philips	PUT
CG1 – B	SB	a), b) & c)	PUT	ETRI
CG1 – C	SC	a), b) & c)	InterDigital	Intel
CG2 – J	SJ	a), b) & c)	Intel / InterDigital	Philips
CG1 – N	SN	a), b) & c)	ETRI	InterDigital
NC1 – D	SD	a), b) & c)	Intel	ETRI
NC1 – E	SE	a), b) & c)	Intel	Philips
NC2 – L	SL	a), b) & c)	Philips	PUT

We have reported in Table 2 the information provided by the participants on their used computation resources.

Table 2: System information

	Operating system	Compiler	C++ standard library
ETRI	Windows10	VC15	
PUT	Windows10	VC15	
Intel	Windows 10	VC15	
InterDigital	Centos Linux 7	GCC 9.1.0	GCC 9.1.0
Philips	Scientific Linux release 7.6 (Nitrogen)	GCC 9.1.0	GCC 9.1.0

2 Notes and observations

2.1 Adding Analysis sheet

Two new analysis sheets “Analysis A97” and “Analysis A17” have been added in ff Excel file and in rf Excel file respectively, with extra information related to all content classes:

- Percentage of total bitrate related to texture, depth and metadata
- A (PSNR, bitrate) curves (in non-logarithmic horizontal axis).

2.2 Number of metrics

The VIF metrics have not been filled by all participants, making it more questionable for the next CTC iteration. Reducing the number of metrics below 5 could be debated.

2.3 Anchor generation time

The table below shows the cumulated computation time for all 8 contents classes, A97 and A17, V17, as reported by the participants (not the crosscheckers). As a matter of fact, there is a significant discrepancy of the computation times reported by participant and crosschecker:

- The computation difference can be as high as 30%, even 50% for BA97 Rendering time
- the faster for one content may be slower for another (cf for example Intel & Philips for JA97 and EA97)
- the faster for 1 QP may be slower for other QP (cf for example EA97 encoding time)
- the faster for one part like the encoder may be slower for other part like the renderer (cf for example JA97 between encoder and renderer)

The difference in memory size allocated to each computation core and number of threads running on same computing node / workstation may explain some of these discrepancies, but not all.

Also, it happens that the Generation Time has not been reported sometimes, while it may be a useful information. When the Gen time figure of the participant were missing, it has been replaced by the crosschecker’s one.

This computation time report remains anyway a useful information. The table 3 of TMIV 3.0 Anchor computation time can be compared with Table 4 of previous TMIV2.0, and one can see that

- V17 computation time has been drastically reduced, due to 5.7× less frames to compute, leading to faster global Anchor computation, which was the intention.
- The generation time Gen for A97 has been increased since TMIV2.0, probably due to Atlas pruning sophistication.

Table 3: TMIV3.0 Anchor generation time in compute hours

	Enc T [h]	Dec T [h]	Gen T [h]	Ren T [h]	Total [h]
MIV A97	505	1	118	355	979
MIV A17	81	0	21	65	167
MIV view V17	178	0	0	216	394
All anchors	764	2	139	636	1540

Table 4:TMIV2.0 Anchor generation time in compute hours (source: Philips, from m49961)

	Enc T [h]	Dec T [h]	Gen T [h]	Ren T [h]	Total [h]
MIV	462	1	17	217	698
MIV view	820	2	1	1137	1960
All anchors	1282	4	18	1354	2658

3 Anchor Crosscheck

We have a full check between any Anchor computation participant and its crosschecker for the 3 class of computation A97, A17 and V17, in the sense that the BD rates are all 0.00%. Note that md5 on the generated bitstream has not been computed this time.
Exception to this was a very slight difference on JA17 between Intel and Philips, but Interdigital has successfully crosschecked Philips results, and its results are present in the JA97 template.
The synthesis results differences are lower than 0.005 dB, except 1 value for NA17 equal to 0.01dB.

4 Extra Excel template for comparing A17 versus V17

In addition to the 2 previously mentioned templates, a third one has been added for highlighting the comparison of A17 (left) versus V17 (right). The rate-distortion curves (i.e. Y-PSNR [dB] vs. Bitrate [Mbps]) across various sequences are shown in Fig. 1. The saving in pixel rate for the A17 compared to A17 is shown in Table 5.

Table 5: The Pixel rate ratio revealing the saving achieved in A17 compared to V17.

SA	SB	SC	SJ	SN	SD	SE	SL
-83.33%	-56.64%	-50.00%	-33.33%	-42.19%	-29.41%	-14.29%	6.67%

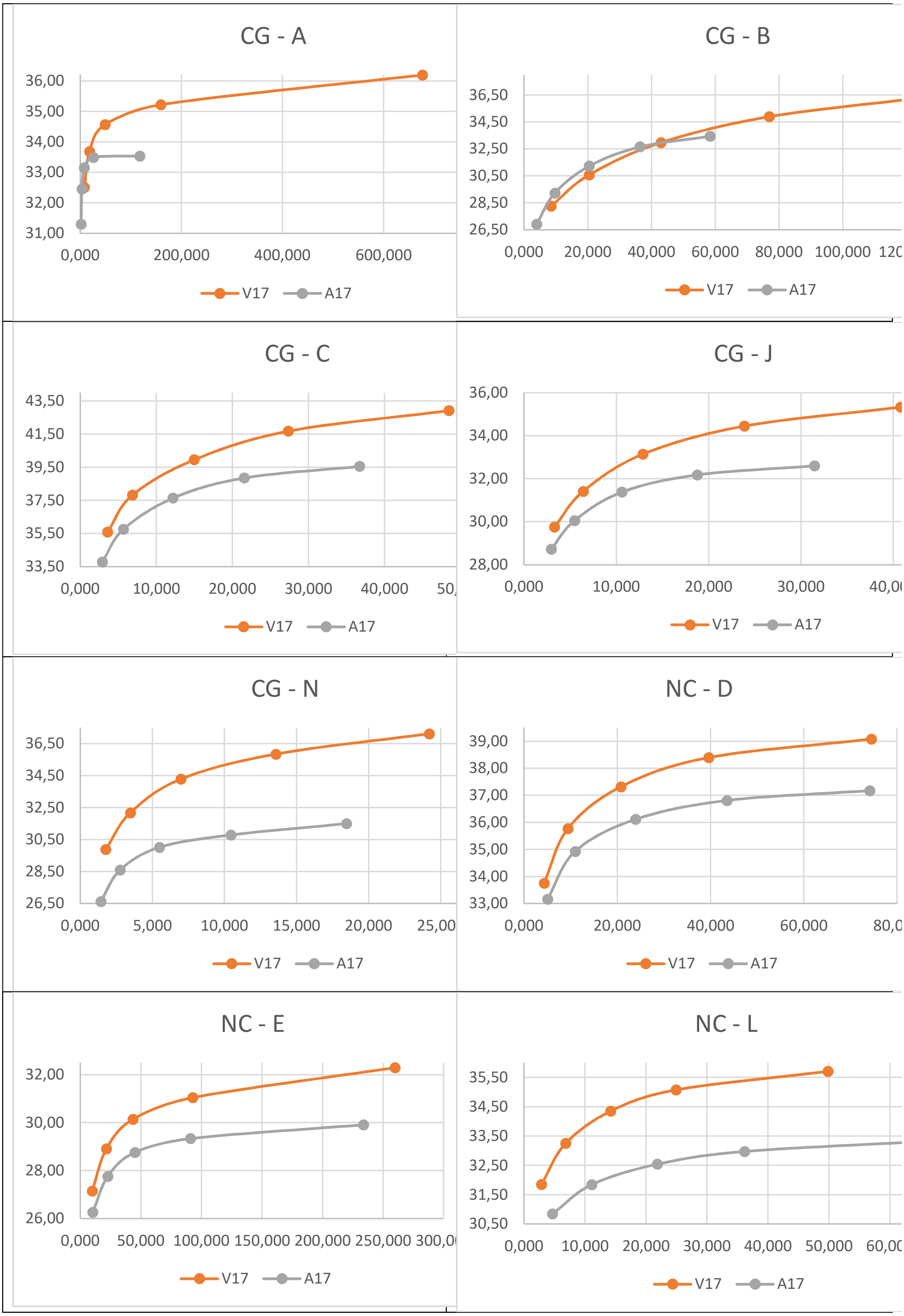

Figure. 1. Rate-distortion curves for various content comparing A17 & V17 anchors.

5 Discussion on R97

In addition to previously mentioned 3 Anchors for objective metrics, Intel has generated 300 frames pose traces from all source views (R97 like "raw") using Multi-Pass Renderer (with 3 passes of 2, 4, all views per pass). Location of the pose traces video from Intel are here: http://mpegfs.int-evry.fr/mpegcontent/ws-mpegcontent/MPEG-I/Intel/TMIV3_R97PoseTraces

This location includes a subdirectory with side-by-side mp4 videos with R97 on the left and A97 on the right. These videos have been produced using the following command (where "anchor_yuv" below relates to R97 and "proposed_yuv" relates to A97):

ffmpeg.exe -f rawvideo -pix_fmt yuv420p10le -s:v {width}x{height} -r {fps} -i {anchor yuv} -f rawvideo -pix_fmt yuv420p10le -s:v {width}x{height} -r {fps} -i {proposed yuv} -c:v libx264 -crf 10 -pix_fmt yuv420p -filter_complex "[0:v]scale=1920:-1[v0];[1:v]scale=1920:-1[v1];[v0][v1]hstack=inputs=2" {output mp4}

The purpose is to define a reference which Anchors pose traces should get close to, but will not be able to overperform. This also helps evaluating the rendering side in isolation of the other TMIV components.

One possibility would be to create a "near ground truth" pose traces that can be used as a reference to compute the objective metrics on the anchors' and proponents' pose traces to better evaluate the delivered quality in non-source positions and made such info part of the CTC templates. For that, it will be helpful to have the ray traced ground-truth pose traces for the synthetic content while keep improving the ones for natural content (i.e. better depth maps, better tuning for rendering parameters).

6 Recommendations

- We recommend that the attached template is used by all proponents.
- We propose that the attached template forms the basis for the next CTC reporting template.
- We recommend addressing the question of R97 usage in the next Ad Hoc meeting or confcall.
- We invite volunteering providers of the synthetic content used in MIV to study ray-traced pose traces to include as part of R97.

7 Reference

[1] J. Jung, B. Kroon, J. Boyce, M. Domański (Eds) Common Test Conditions for Immersive Video, ISO/IEC JTC1/SC29/WG11 MPEG/N18789, November 2019, Geneva Switzerland