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

ISO/IEC JTC1/SC29/WG04 MPEG/M59126
January 2022, Online

Source	MPEG Video coding
Status	Input document
Title	BoG on MPEG Immersive video
Authors	Bart Kroon (Philips), Vinod Kumar Malamal Vadakital (Ofinno), Joel Jung (Tencent), Dawid Mieloch (PUT), Basel Salahieh (Vimmerse)

1 Introduction

The related AHG report is MPEG/M58463 *AHG on MPEG Immersive video*.

Offline discussion is at: http://mpegx.int-evry.fr/software/MPEG/MIV/InputDocuments

2 BoG notes

2.1 Late documents

2.1.1 m58997 Deep Learning Based Specular Pruning

- previous proposal used pix2pix network fo specularity highlight detection
- this proposal has trained pix2pix on three different datasets
  - [2] G. Fu, Q. Zhang, Q. Lin, L. Zhu, C. Xiao, “Learning to Detect Specular Highlights from Real-world Images,” ACM MM, Oct. 2020.
  - [4] M. D. Elkhouly, T. Tsesmelis, A. D. Bue, S. James, "Lights: Light Specularity Dataset For Specular Detection In Multi-View," ICIP, Sept. 2021.
  - [5] G. Fu, Q. Zhang, L. Zhu, P. Li, C. Xiao, “A Multi-Task Network for Joint Specular Highlight Detection and Removal,” CVPR, June. 2021.
- 7 h training time
- Dataset 1 and 2: BD-rate loss (7%) for mandatory sequences, gain (14%) for optional sequences
- Dataset 3: BD-rate loss (4%), and gain 1.1% respectively

Discussion:

Q: Are white colours due to saturation detected as specular due to learning-based approach? Is it possible to differentiate?
A: Really difficult to distinguish. The labels in the datasets are already like this and it was difficult to correct this due to the large number of images.

Q: Learning on images, but apply on video?
A: Only one frame was tested but with coding, so "A1" CTC anchor

Informative contribution.

2.2 Specifications

2.2.1 WD1 of ISO/IEC 23090-12(2E) MIV

There are only two editors because Lukasz does not want to continue due to other obligations. Should there be a third editor? -> yes.

One open issue was not discussed: clause 9.2.5.2 *Atlas data decoding process* should not depend on the VPS (#441). It is not necessary to fix, and the issue was closed.

Based on input from PUT and ETRI, and having talked with Vinod (@vinod_mv), we propose as editors for MIV edition-2:

- Jill (@jillboyce)
- Bertrand (@ChupeauB)
- Vinod (@vinod_mv)
- Bart (@bartkroon)

2.2.2 ISO/IEC 23090-5(2E) V3C

- Adopted m58677 V3C CodecGroup Entry for HEVC Main profile in joint meeting with WG 04 and WG 07.
- Most of the ballot comments are disposed, only three are open

2.3 Use cases and requirements for MIV edition-2

2.3.1 m59085 Draft use cases and requirements for MIV – edition-2

WG 02 BoG:

- v1 of the input document was accepted without modifications, apart from the title and document number.

Offline:

- v2 of the input document is cleaned up, and ready apart from the document number.
- To be presented in a WG 02 plenary.

2.3.2 m58861 heterogenous sources in MIV edition 2

Joint meeting with WG 07:

- The proposal is summarized, as already discussed in the MIV AHG
- Some technical work would be needed to implement the requirement
- Question about the level of integration: per patch, or higher level
- This level of integration is possible, it mainly requires a V3C profile, but probably also a small amount of syntax work
- This will be V3C(3E) because V3C(2E) will go to FDIS

- One expert claims that the number of decoders is not important as long as the total number of pixels is similar, because decoder sessions can be time-multiplexed
- Combination of MIV and V-PCC at MPEG scene description level is possible, and this is the reference. An EE could try to build a solution based on this proposal, to study the benefits. It is suggested to study the influence of the number of decoder instantiations.
- The WG 07 convenor has no issue with having a V3C-level requirement in the MIV edition-2 use cases and requirements document
- The joint EE's will have a WG 04 output document

2.4 Promotion of MIV

2.4.1 Whitepaper

- v2 of m58942 is a cleaned up draft of the final white paper.
- There is a wish to have results in the white paper. Bart has looked up, and for the SBTVD proposal, results from [m57096] were used. To be discussed in the plenary (done):

[m57096] J. Jung, A. Giraud, X. Li, S. Liu (Tencent), B. Kroon (Philips), V.K. Malamal Vadakital (Nokia), MOS analysis of new CTC QPs for MIV, MPEG 135 input document m57096, July 2021

-> Only provide bitrate ranges.

AG 3 meeting:

- The lack of performance numbers was immediately noted.
- I have explained that the verification test is not ready and we do not want to show objective results because with 6DoF viewports we have a correlation problem with objective metrics and expert viewing. Mathias was in the call, and I mentioned that Mathias would agree with me, and he did.
- It was discussed if the white paper should be postponed until after the verification test
- Instead, it was decided to have the white paper now, and a new white paper to report on the verification test when ready.
- The final version (by stripping of the cover page and removing all tracked changes) was sent to and received by the AG 3 convenor.

2.4.2 Press release on the CD of 23090-23

1. A draft press release was prepared in time for the next plenary session: #479.
2. The press release was accepted in the plenary session.
3. AG 3 meeting: The press release was accepted w/o further review.
4. The AG 3 and WG 04 convenor will review the press release once more before publication.

2.5 Experiments

2.5.1 Verification test preparations (#480)

- Run experiments on CTC sequences and then use the new configuration with the VT sequences.
    - Use a subset of CTC sequences: Museum, Chess, Fencing, Frog
- Focus only on the quality of posetraces.

- Try using RVS with TMIV to check if using the same synthesizer is possible.
- Can we use different profiles/configurations with different ratepoints/content?
  - Possible split on the CG/natural, perspective/omnidirectional.
  - Everyone agrees it can be done.
- The conditions of experiments were agreed and will be implemented in the description of this issue.
- The communication with AG5 will be done only in another specified issue to avoid sending too many unnecessary messages.

Sufficient volunteers have registered to participate in the VT dry run and EE 6 expert viewing. Those interested in joining, please contact Joel Jung.

2.5.2 Exploration experiments

- Offline discussions are progressing well on average
- MIV EE-1 has four participants
- MIV EE-4 has four participants
- MIV EE-5 has two participants and as a continuation that is acceptable
- MIV EE-6 has four participants
- MIV EE-7 has three participants

All considered EE's have enough participation!

2.5.3 Joint exploration experiments with WG 07

- V3C JEE-1 was discussed in a JBoG on Tue 15:30 UTC
  - Sufficient Video and 3DG experts have attended: **six participants**
  - Due to the mesh CfP, 3DG experts cannot commit to additional experiments
  - This EE will be spread out over two meeting cycles
  - Draft text of the description: m59135
- V3C JEE-2 was discussed in a JBoG on Wed 23:00 UTC
  - Currently there are no solutions for combining V-PCC and MIV patches in one video sub-bitstream
  - The goal is to create a platform that allows for demonstration that there is a benefit in integrating heterogeneous sources in a bitstream
  - There is no need for objective evaluation or expert viewing, thus there is also no need for an anchor and it is not important if CTC's are followed
  - Because V-PCC has OVD, it is better to also use MIV with OVD
  - How is the work divided? What is the mode of collaboration?
  - There are **three participants**
  - The participants have had an offline call to work out the JEE description: notes
  - The JEE document (JEE 1 and JEE 2) will be reviewed on Friday
  - Based on the results of this EE, in the next iteration, it may be possible to define meaningful evaluation criteria including an anchor
  - Draft text of the description is shared at #491

2.6 Call for Test Material

~~A first draft text is available.~~ A third draft text is available: WG04Nxxxx_v3.zip

Jun Young Jeong will replace Renaud Doré as contact person.

MIV BoG 2: extend source format (Wed 15:30 UTC):

Edits to the call for test material document have been reviewed. These include:

1. Emphasize on collecting indoor and outdoor scenes for inward and outward looking cameras.
2. Allowing cameras with temporally varying intrinsics and extrinsic parameters.
3. Allowing colourized depth representation.
4. The capture depth video need not be aligned with cameras that capture other attributes (e.g. texture).
5. Requesting good quality object segmentation maps corresponding to the captured videos.

Discussion on the json format, there is agreement on introducing csv file per view to reflect the changes in extrinsics and intrisics per frame:

UpdateAtFrame	FocalX	FocalY	X	Y	Z	yaw	pitch	roll	depth min	depth max
0	0.1	0.2	0	2	0.2	0	2	2	0	1
10	0.1	0.2	1	5	0.2	1	5	5	0.2	0.8
20	0.1	0.2	2	3	0.2	2	3	3	0.2	0.8
150	0.5	0.3	2	3	0.3	2	3	3	0.3	0.7

Recommendations of the BoG meeting:

- Approve the proposed edits to the CftM document.
- Agree on removing the json section from the document but attach as json file along with the document.
- For fisheye/wide-angle content, recommendation is to have raw and undistorted content if available.
- Request an input contribution to define the exact csv format for support varying intrinsics and extrinsics + how the TMIV integration should be.
- Request an input contribution to support "cameras added or dropped during the recording" feature.