

**ISO/IEC JTC 1/SC 29/WG 11**

**Coding of moving pictures and audio**

**Convenorship: UNI (Italy)**

**ISO/IEC JTC 1/SC 29/WG 11 N18996**

**Document type: Approved WG 11 document**

**Title: Call for MPEG-I Visual Test Materials**

**Status: Approved**

**Date of document: 2020-01-31**

**Source: Video**

**Expected action:**

**No. of pages: 5**

**Email of convenor: leonardo@chiariglione.org**

**Committee URL: mpeg.chiariglione.org**

**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC1/SC29/WG11**

**CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC1/SC29/WG11 MPEG2019/N18996**

**January 2020, Brussels,** **Belgium**

|  |  |
| --- | --- |
| **Source** | **Video** |
| **Status** | **Approved** |
| **Title** | **Call for MPEG-I Visual Test Materials** |
| **Editors** | Gauthier Lafruit, Mehrdad Teratani, Renaud Doré, Bart Kroon, Joel Jung, Vinod Kumar Malamal Vadakital |

# Introduction

During exploration experiments, MPEG has realized that there is a lack of test material for video experiments into MPEG-I Immersive Video [N18787], especially for 6DoF and Dense Light Field. This document is an update of the former **[**N18562**]** in order to take multiple camera architectures into account, e.g. ERP, linear perspective camera arrays and lenslet cameras. Moreover, in order to prepare for more advanced features of next generation Immersive Video, capture/creation of **non-Lambertian content** with specularities, transparent objects and reflections is also encouraged.

We hereby solicit new test material for 6DoF and Dense Light Field, **in particular using non-parallel camera setups**. Since all technologies in MPEG-I Visual are based on multiple camera views and their depth maps, we recommend that not only the camera views, but also depth maps (measured or estimated at high quality) are provided by the proponents.

Content may be provided as computer-generated/synthetic **3D models of static and dynamic scenes**, as this material can be used for rendering various viewpoints with computer graphics techniques, creating video footage required in all experiments and comparative studies (e.g. PSNR quality evaluation with reference views).

MPEG is also calling for **natural content**, both **indoor and outdoor**, directly captured with camera rigs. Content with **objects close to the camera** are also requested, since this will challenge the proposed technologies for parallax rendering, e.g. heavy motion parallax for nearby foreground objects.

Content should be provided in any **image-based representation format**, e.g. **lenslet** format, or **multiview + depth**. If content is generated in computer-generated/synthetic format, it should be rendered to create various image-based projections (and their depth maps) to be eligible as test material.

Please consult [N18787] for an overview of available test material.

# Immersive video content

Immersive video here refers to having a number of cameras capturing the scene, out of which any other viewpoint to the scene can be synthesized so that the viewer has the feeling of being immersed into the scene (also thanks to a stereoscopic head mounted display).

The purpose is to prepare content for MIV core experiments and 6DoF exploration experiments. Content will be used to define pose traces and common test conditions for evaluating competing coding technologies, as well as pre-processing (e.g. depth estimation) and post-processing (e.g. rendering by view synthesis) algorithms.

Typical content that would be required with high priority is:

* Multiview +Depth, with increased viewing volume dimension, more cameras, larger distance between cameras than current content used for MIV.
* High quality depth maps (provide confidence map, if available).
* Inward looking and outward looking camera setups, with arbitrary configuration possible (includes any non-trivial configuration).
* Natural content and outdoor scenes

Content with following properties is highly encouraged for preparing the second generation MIV and 6DoF experiments, cf. [m51479]:

* Any projection type corresponding to a physical camera rig (perspective, fisheye, …).
* Non-static content, with biological entities (people, cats, dogs, grass, hairs, …).
* Deep image (such as OpenEXR), with multiple properties per pixel, where properties includes in priority depth information. Looking for:
  + Particle based volumes like fog, water, snow, …
  + reflections
  + transparency
* Moving cameras with step-in/step-out motion and/or change in view orientation

Camera intrinsics and extrinsics should be provided in json format, cf. [N18068].

Textures should be provided in YUV420 8-bits, depth maps in YUV420 at least 8 bits, preferably 10 bits or more.

# Dense Light field content

Dense Light Fields can be either captured by dense multi-camera arrays or a lenslet camera, e.g. Lytro or Raytrix. Currently, two formats (multiview and lenslet) are considered. These data formats can be converted from one to another. Such a conversion can be invertible or not, depending whether we have plenoptic 1.0 or 2.0 camera modes.

This activity aims at comparing the coding performance of different potential representations of dense Light Field data (lenslet, multiview, etc). Currently, several test sequences are provided with dense multiview video, and only a few with lenslet cameras. New lenslet content in plenoptic 1.0 and 2.0 format is solicited, with priority for the latter. Moreover, we do not have any test material captured by an array of plenoptic cameras.

We hence encourage participants to provide MPEG test material for this activity with following contents, by MPEG130: (1) multiview lenslet video, dense multiview video, (2) camera parameters, and (2) the conversion tool for conversion from lenslet to multiview video, e.g. RLC [N18567].

# 3D synthetic models

In the case of photorealistic synthetic content, it is recommended that complementary to the rendered image data, also the 3D model itself plus any scripting to generate the required dataset is supplied. This enables other MPEG experts to render variants of the same scene. It is for instance possible to simulate a virtual camera array by placing multiple cameras in the scene. Preferably, scenes should include a timeline that allows for rendering of short movies with some dynamics. Due to the open source and easy accessibility of the Blender software, a .blend file would be the most preferred format for this type of content

The format of the 3D model should be in a suitable interchange format [N17252], or open source (e.g. Blender[[1]](#footnote-1)). Typical representations of 3D models are:

* Texture, mesh and lighting information,
* Procedural scenes (e.g. algorithms).

Content providers are nevertheless asked to render their 3D scene in order to create multiview image-based representations. Also, the depth maps should be created in the rendering process.

# Experimentation test set

In addition to photorealistic synthetic content that is solicited, simple Cornell-Box type scenes as illustrated in Figure 1 are also solicited. These types of focused scenes are useful to test individual aspects of a volumetric video and the experimentation and evaluation of algorithms would become easier.

|  |  |
| --- | --- |
| A picture containing wall, tree, indoor, green  Description automatically generated | A picture containing indoor, wall, green, ceiling  Description automatically generated |

Figure 1: (left) a Cornell-Box with a smoke simulation; (right) a Cornell-Box with a pure reflective cuboid and a glossy sphere.

# Copyright notice

Content owners should provide a copyright notice along with the dataset to inform MPEG about copyright and usage restrictions.

# Contact:

* Bart Kroon, [bart.kroon@philips.com](mailto:bart.kroon@philips.com)
* Gauthier Lafruit, [gauthier.lafruit@ulb.ac.be](mailto:gauthier.lafruit@ulb.ac.be)
* Mehrdad Teratani, [panahpour@nuee.nagoya-u.ac.jp](mailto:panahpour@nuee.nagoya-u.ac.jp)
* Vinod Kumar Malamal Vadakital, vinod.malamalvadakital@nokia.com

# References

[m51479] J. Jung, V.K. Malamal Vadakital, BoG report: future aspects on immersive video, ISO/IEC JTC1/SC29/WG11 MPEG2019/ m51479, Geneva, October 2019.

[N18068] Reference View Synthesizer (RVS) manual, ISO/IEC JTC1/SC29/WG11 MPEG2018/N18068, Macao, October 2018.

[N17252] Interchangeable Formats in VFX Industry (v1), ISO/IEC JTC1/SC29/WG11 MPEG2018/N17252, Macau, October 2018.

[N18562] Call for MPEG-I Visual Test Materials, ISO/IEC JTC1/SC29/WG11 MPEG2019/N18562, Gothenburg, July 2019.

[N18786] Summary on MPEG-I Visual Activities, ISO/IEC JTC1/SC29/WG11 MPEG2019/N18786, Geneva, October 2019.

[N18787] Overview of MPEG-I Visual Test Materials, ISO/IEC JTC1/SC29/WG11 MPEG2019/N18787, Geneva, October 2019.

[N18788] Call for MPEG-I Visual Test Materials, ISO/IEC JTC1/SC29/WG11 MPEG2019/N18788, Geneva, October 2019.

[N18793] Manual of Reference Lenslet Content Convertor (RLC 1.0), ISO/IEC JTC1/SC29/WG11 MPEG2019/N18793, Geneva, October 2019.

1. https://www.blender.org/ [↑](#footnote-ref-1)