 ISO/IEC JTC 1/SC 29/WG 04 N0751

**ISO/IEC JTC 1/SC 29/WG 04  
MPEG Video Coding   
Convenorship: CN**

**Document type:** Output Document

**Title:** Common test conditions for Gaussian splat coding

**Status:** Approved

**Date of document:** 2025-10-11

**Source:** ISO/IEC JTC 1/SC 29/WG 04

# Expected action: None

# Action due date: None

**No. of pages:** 9 (without cover page)

**Email of Convenor:** yul@zju.edu.cn

**Committee URL:** <https://sd.iso.org/documents/ui/#!/browse/iso/iso-iec-jtc-1/iso-iec-jtc-1-sc-29/iso-iec-jtc-1-sc-29-wg-4>

**INTERNATIONAL ORGANIZATION FOR STANDARDIZATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC 1/SC 29/WG 04 MPEG VIDEO CODING**

**ISO/IEC JTC 1/SC 29/WG 04 N** **0751**

**October 2025, Geneva (CH)**

|  |  |
| --- | --- |
| **Title** | **Common test conditions for Gaussian splat coding** |
| **Source** | **WG 04, MPEG Video Coding** |
| **Status** | **Approved** |
| **Serial Number** | **25782** |

# Introduction

Common test conditions are desirable for conducting coding experiments in a well-defined environment and easing the comparison of the outcome of experiments. This document provides the common test conditions (CTC) for the Gaussian splat coding (GSC) exploration into Gaussian splat coding, conducted jointly by WG 4 *MPEG Video Coding* and WG 7 *3D Graphics and Haptics* experts. This document replaces [N 1292]. The common test conditions are defined to evaluate the coding efficiency, objective and subjective quality, complexity, pixel rate, and user experience of Gaussian splat coding solutions for one or more frames.

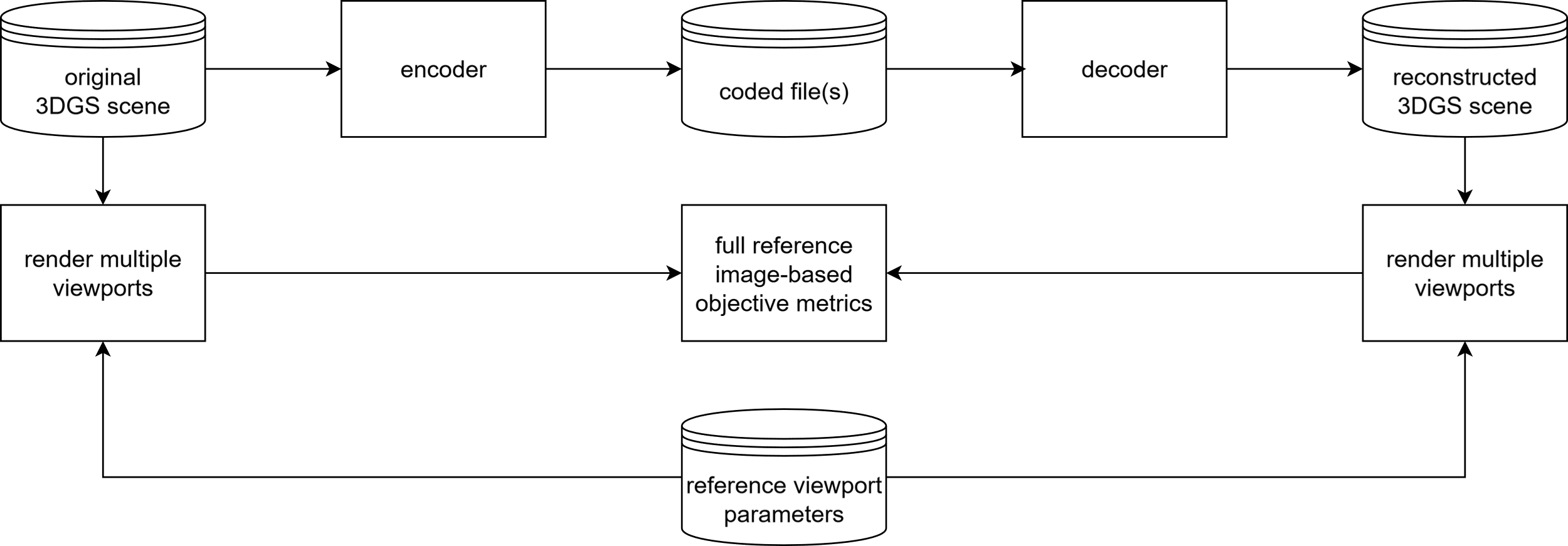


Figure 1: Objective evaluation

This document intentionally excludes test conditions for rendering or training. In general, joint exploration experiment descriptions may override this document by providing specific test conditions, under the condition that there is an explanation why that is needed. Ideally, such deviations should be minimized, and any divergence from this document should remain as limited as possible.

# Test sequences

There are two variants of the test conditions denoted 1F and NF:

* 1F: For static coding experiments (single frame) the frame count is equal to 1.
* NF: For dynamic coding experiments (all intra, random access, etc.) the frame count is equal to 32 for all sequences.

The set of mandatory sequences is different for 1F and NF, as listed in Table 1. A description of the test sequences is provided in Annex A. Both variants have the same start frame for all sequences. Test material is available on request. MPEG experts are expected to consult the newest description of JEE 6.1 *data preparation*.

Table 1: Test sequences for experiments in accordance with this document

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Class** | **ID** | **1F** | **NF** | **Start**  **frame** | **Frame size** |
| Forward  facing | bartender\_tracked | Optional | Optional | 0 | 1920 × 1080 |
| bartender\_semitracked | Mandatory | Mandatory | 0 | 1920 × 1080 |
| cinema\_tracked | Optional | Optional | 0 | 1920 × 1080 |
| cinema\_semitracked | Mandatory | Mandatory | 0 | 1920 × 1080 |
| breakfast\_tracked | Optional | Optional | 0 | 1920 × 1080 |
| breakfast\_semitracked | Mandatory | Mandatory | 0 | 1920 × 1080 |
| breakfast\_untracked | Mandatory | Mandatory | 0 | 1920 × 1080 |
| breakdance\_untracked | Mandatory | Mandatory | 0 | 1920 × 1080 |
| Object  centric | manwithfruit\_tracked | Mandatory | Mandatory | 81 | 3840 × 2160 |
| lego\_ferrari | Mandatory | N/A | 0 | 4594 × 5514 |
| lego\_bugatti | Mandatory | N/A | 0 | 3852 × 2868 |
| cricket\_player | Mandatory | N/A | 0 | 4520 × 2540 |
| plant | Mandatory | N/A | 0 | 2954 × 3968 |
| solo\_tango\_female | Mandatory | N/A | 0 | 4534 × 2542 |
| solo\_tango\_male | Mandatory | N/A | 0 | 4528 × 2544 |
| tango\_duo | Mandatory | N/A | 0 | 4522 × 2538 |
| tennis\_player | Mandatory | N/A | 0 | 4518 × 2540 |
| library | Mandatory | N/A | 0 | 3793 × 2131 |
| flowerdance | Mandatory | N/A | 0 | 2456 × 2054 |
| gymnast | Mandatory | N/A | 200 | 2456 × 2054 |

# Software and tools

This document provides the intended software tags of software to be used in Table 2. If during anchor generation a problem is found that prohibits the anchor to be generated in a reproducible way, then software may be fixed and the anchor report will provide new tags. It is not permitted to use non-stable references like branch names in MPEG contributions that follow the CTC.

Table 2: Reference software for this version of this document

|  |  |  |
| --- | --- | --- |
| **Name** | **URL** | **Tag** (SHA) |
| MPEG GSC metrics | [mpeg-gsc-metrics](https://git.mpeg.expert/MPEG/Explorations/GSC/gsc-software/mpeg-gsc-metrics) | release\_2.1  (215afff8) |
| MPEG 3D renderer | [mpeg-3d-renderer](https://git.mpeg.expert/MPEG/Explorations/GSC/gsc-software/mpeg-3d-renderer) | release\_mpeg152  (to be released) |
| Gaussian Splat Tools | [mpeg-gsc-tools](https://git.mpeg.expert/MPEG/Explorations/GSC/gsc-software/mpeg-gsc-tools/-/tree/main/gsTools?ref_type=heads)  /gsTools | release\_mpeg151  (42505ee9) |
| Pre- and post-processor | [mpeg-gsc-tools](https://git.mpeg.expert/MPEG/Explorations/GSC/gsc-software/mpeg-gsc-tools/-/tree/main/gsTools?ref_type=heads)  /pre\_post\_processing | release\_mpeg151  (42505ee9) |
| GSCodec Studio | [gscodec\_studio](https://git.mpeg.expert/MPEG/Explorations/GSC/gsc-software/gscodec_studio) | v1.2  (c3663580) |
| HEVC model (HM) | [HM](https://vcgit.hhi.fraunhofer.de/jvet/HM) | HM-18.0  (fb4486d5) |
| G-PCC for GSC | [g-pcc-for-gsc](https://git.mpeg.expert/MPEG/Explorations/GSC/gsc-software/g-pcc-for-gsc) | release\_mpeg151\_rc1  (ee95dbc1) |

# Anchor generation

## General

The purpose of an anchor is to provide a realistic reference that supports the evaluation of a proposal. Anchors apply to 1F and NF unless specified otherwise. A proposal needs to compare against the most relevant anchor. Compared to that anchor, proponents should aim to change only one thing in an experiment (ceteris paribus) or otherwise provide ablation studies to breakdown more complex proposals.

## QP tuning of anchors

It is within the mandate of anchor generators to tune QP's, and thus no QP values are provided in this document. Preferably QP tables are added to the software repositories of anchor and proposed encoder models to avoid unnecessary crosschecking mismatches. Target file sizes and bit rates for anchors and proposals are provided in Table 3. Rate points can be adjusted per sequence to ensure measuring a sufficient quality range while avoiding saturation.

Table 3: Rate points with target sizes [MB] for 1F and bit rates [Mbps] for NF

|  |  |  |
| --- | --- | --- |
| Rate point | 1F target [MB] | NF target [Mbps] |
| r01 | 5 | 10 |
| r02 | 10 | 20 |
| r03 | 20 | 50 |
| r04 | 50 | 100 |
| r05 | 100 | 200 |

[Ed.(m74626): Values in Table 3 are example values, to be updated. These values should ideally be derived from market input, use cases, and requirements.]

## Geometry-based anchor (1F-geo)

This geometry-based anchor maps splat coefficients to the attributes of a sparse point cloud, and codes them using ISO/IEC 23090-9:2023 Geometry-based point cloud compression (G-PCC).

This anchor is based on a fork of the test model that was created for the Gaussian splat coding activity, called *G-PCC for GSC* in combination with the *Pre- and postprocessor*. The correct URLs and versions of both are listed in Table 2, and pre- and postprocessing is performed with the bit-depths provided in Table 4.

Table 4: Pre- and postprocessing parameters for the geometry-based anchor

|  |  |
| --- | --- |
| **Parameter** | **Value** |
| bits\_pos | 18 |
| bits\_sh | 12 |
| bits\_opacity | 12 |
| bits\_scale | 12 |
| bit\_rot | 12 |

## Video-based anchor (1F-vid)

The video-based anchor based on *GSCodec Studio* uses packing of Gaussian splat positions and attributes to video planes that are encoded as separate videos using HEVC (HM). The correct software URLs and versions are listed in Table 2. Metadata is produced by the encoder to allow for the generation of the reconstructed 3DGS scene. Packing itself (reorganization) is considered as pre-processing and is fixed in the anchor. The general scheme is illustrated in Figure 2.

A screenshot of a computer screen

AI-generated content may be incorrect.

Figure 2 Video-based anchor based on GSCodec Studio

To compute the anchor, follow the instructions in [/scripts/README.md](https://git.mpeg.expert/MPEG/Explorations/GSC/gsc-software/gscodec_studio/-/blob/v1.2/scripts/README.md?ref_type=tags) to install all packages, and use the command-lines that are provided in [/examples/benchmarks/mpeg152/1f\_vid\_hm/](https://git.mpeg.expert/MPEG/Explorations/GSC/gsc-software/gscodec_studio/-/tree/v1.2/examples/benchmarks/mpeg152/1f_vid_hm?ref_type=tags) to run the experiments.

When asked to install PLAS from the JasonLSC fork, use this commit:

pip install git+https://github.com/JasonLSC/PLAS@002f46584edc8a0aa67b2deed1d47b10c64f9564

**WARNING**: The output of the anchor depends on the PyTorch version.

# Objective evaluation

Objective evaluation for one test point is conducted according to Figure 1. It is out of scope how the original 3DGS scene or reference viewport parameters were created.

## Running the encoder

The encoder receives the original 3DGS scene and a configuration, resulting in one or more coded files. The configuration is required to be sequence-independent. It is not permitted that the encoder uses reference camera parameters other than for adding supplemental information to a bitstream: reference camera parameters may not be used to improve the coding of splats. However, it is permitted that the encoder receives and uses viewing space information that indicates coarsly what positions and angles are suitable for viewing of the scene.

The total file size, including all metadata, is measured. All information required for decoding must be measured. For 1F this is reported as a size in bytes. For NF this is reported as a bit rate in kbps (1000 bps). It is permitted that the encoder calculates and reports the size or bit rate for convenience, but exactly the same number can be obtained by manually summing over the size of each coded file[[1]](#footnote-1). The encoder also measures and reports runtime complexity.

## Running the decoder

The decoder receives the one or more coded files, and outputs a reconstructed 3DGS scene. It is not permitted that the decoder uses the reference viewport parameters. The decoder also measures and reports runtime complexity. It is not permitted that the decoder uses any out-of-band information that is not included in the total size of the coded files. It is also not permitted that the decoder is configurable in any way that may impact the decoded result.

## Rendering and objective quality metric calculation

For the purpose of objective evaluation, the rendering and metric calculation steps have been combined in two software tools: MPEG GSC metrics and GsTools (Table 2). It thus requires two steps to calculate the objective quality metrics for a test point:

1. Run GsTools once to add the reference viewport parameters to source 3DGS scenes.
2. Run MPEG GSC metrics to render multiple viewports, run full-reference objective quality metrics, and average them over all evaluated frames and viewports.

Exactly the following parameters have to be provided to MPEG GSC metrics:

-a, -b, -o

--frameCount=*FRAME\_COUNT*

--startFrame=*START\_FRAME*

--width=*FRAME\_WIDTH*

--height=*FRAME\_HEIGHT*

--verbose=1

--cpu=1

--useCameraPosition=1

Providing other parameters may cause crosschecks to fail. The *START\_FRAME*, *FRAME\_WIDTH* and *FRAME\_HEIGHT* have to be set to the values in Table 1. The *FRAME\_COUNT* depends on the variant (1F or NF).

The reporting template has room for filling in the metrics that are listed in Table 5. For a crosscheck to be successful all mandatory objective metrics have to be filled in, and they need to match closely (<0.005% BD-rate), which coincides with having 0.00% on the summary tab of the reporting template. Optional objective metrics are for information only.

Table 5: Mandatory and optional objective metrics

|  |  |
| --- | --- |
| **Objective metric** | **Status** |
| RGB-PSNR | Mandatory |
| YUV-PSNR | Mandatory |
| YUV-SSIM | Mandatory |
| Custom A | First optional metric of choice |
| Custom B | Second optional metric of choice |

## Reporting of runtime complexity

Encoders, decoders and renderers need to report runtime complexity by measuring delta CPU time. If GPU processing is used, proponents must explicitly report it.

## Reporting template

The file sizes or bit rates, objective quality metrics, and runtimes have to be entered into an Excel reporting template. An empty template is attached to this document.

* Anchor generators will provide a template with the left side filled in.
* Anchor crosscheckers will provide a template with the anchor filled in on the left side, and the crosscheck on the right side.
* Proponents and their crosscheckers will provide templates with the anchor on the left side, and the proposal on the right side.

# Subjective evaluation

For subjective viewing, each sequence is synthesized according to a set of pose traces. A pose trace specifies for each frame the position and orientation of the viewport to synthesize. Pose traces are specified according to the MPEG 3D renderer data format, see for example [cfg/plane.txt](https://git.mpeg.expert/MPEG/Explorations/GSC/gsc-software/mpeg-3d-renderer/-/blob/release_mpeg151/cfg/plane.txt?ref_type=tags). The format of each synthesized view is an image with perspective projection with 1920 × 1080 pixels resolution and 10-bit YUV 4:2:0 color format. The purpose is to mimic natural viewing on a head-mounted display while using offline tools and a 2D monitor.

Use the following command-line for MPEG 3D renderer:

scripts/renderer.sh \

--input=input/ \

--output=video.mp4 \

--width=1920 \

--height=1080 \

--videotype=3 \

--gaussian=1 \

--camera=C*AMERA\_PATH*.txt

Because of the large difference in visual comfort between a viewer that voluntarily initiates head motion versus a viewer watching the same viewport on a 2D monitor, pose traces have a small amount of motion. For each sequence there are three pose traces, which are meant to represent a diversity of natural head movement compliant with the overall dimension of the capture rig. The pose traces will be provided with the sequences as part of JEE 6.1 *data preparation*. It is meaningful to define the pose traces according to the conditions of capture, and typically to define the related path within the volume of the camera rig. It is convenient to formulate this range as a volume in 3D space.

For adoption of a proposed method, the proponent must provide the pose trace videos of the proposed method, during a viewing session, and make it clear what the bitrate and pixel-rate differences with the anchor are.

# Rules for proponents

## General

A proposal needs to compare against the most relevant anchor. Compared to that anchor, proponents should aim to change only one thing in an experiment (ceteris paribus) or otherwise provide ablation studies to breakdown more complex proposals.

## QP tuning of proposals

It is the responsibility of each proponent to determine QP values that are compatible with their own algorithmic solution, in order to maintain sufficient bitrate overlap with the anchor. They are not required to use the same QPs as those used for the anchor.

## Reporting template

Proponents and their crosscheckers will provide templates with the anchor on the left side, and the proposal on the right side.

# Evaluation of alternative representations

Within the scope of the GSC exploration, there was an interest in evaluating codecs for proposed alternative representations because the field is developing quickly.

Unfortunately, objective comparison of two coding proposals is only possible when they use the same source and reconstructed format. Hence, when possible, it is advised that proposed codecs are outfitted to convert between 3DGS and the proposed source format, such that experiments can be meaningfully compared against the anchors in this document and other proposals. No access to ground-truth images is allowed in this format conversion.

If format conversion without loss of quality is not possible, then additional effort is required:

1. Prepare test material in the alternate source format, without restrictions of access to ground truth views, resulting in a new set of test sequences.
2. Test and select a suitable renderer, with acceptable software license, that accepts the proposed alternate source format and reference viewport parameters, and outputs a rendered viewport in the same way as MPEG GSC metrics does.
3. Perform crosschecks including subjective evaluation to confirm that test sequences can be rendered with good quality, and reproducibly on relevant platforms.

The above steps are outside of the scope of this document, but having done that, it is possible to perform coding experiments based on Figure 3 and attached reporting templates. The joint exploration experiment needs to provide an alternative for Table 1.

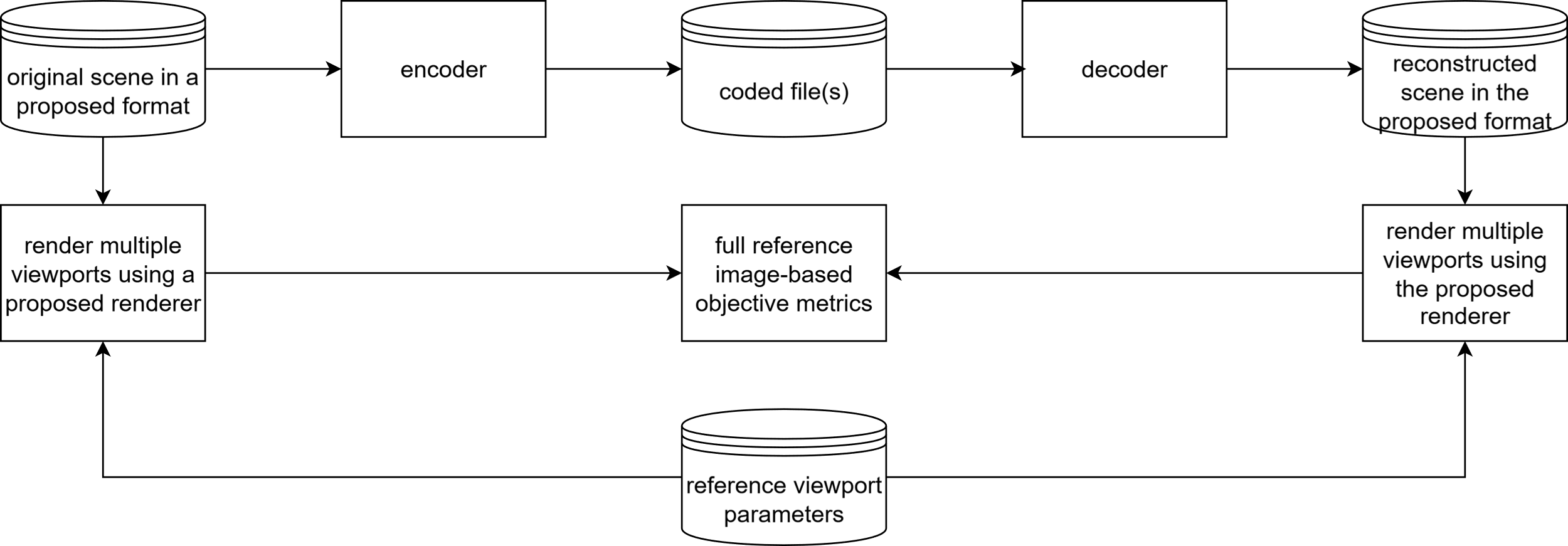


Figure 3: Objective evaluation of an alternative source format partially within the scope of this document

# References

[N 1292] *Draft CTC for Gaussian Splat Coding*, ISO/IEC JTC 1/SC 29/WG 7 N 1292, July 2025, Daejeon (KR)

# Attachments

* Empty reporting template

# Annex A: Description of test sequences

[Ed.(m74626): Annex A is work in progress.]

This annex contains an informative description of test sequences such as how the content was made, number of splat primitives, etc.

Relevant information:

|  |  |
| --- | --- |
| **Category** | Forward facing or object centric |
| **Input contributions** | mXXXXX |
| **Length & frame rate** | XXX frames (XXX fps) |
| **Number of captured views** | XXX |
| **Texture format** | YUV 4:2:0 10 bit |
| **Captured view resolution** | 4096 × 2048 |
| **Download link** | http://XXX |

It can also be of interest to have a visual representation of the location of the captured viewpoints in the 3D space.

1. Hence it's advisable to keep SEIDecodedPictureHash and similar parameters turned off [↑](#footnote-ref-1)