 ISO/IEC JTC 1/SC 29/WG 2 N452

**ISO/IEC JTC 1/SC 29/WG 2**

**MPEG Technical requirements   
Convenorship: SFS (Finland)**

**Document type:** Output Document

**Title:** Draft Gaussian splat coding requirements

**Status:** Approved

**Date of document:** 2025-04-03

**Source:** ISO/IEC JTC 1/SC 29/WG 2

# Expected action: none

**Action due date:** none

**No. of pages:** 4 (without cover page)

**Email of Convenor:** igor.curcio@nokia.com

**Committee URL:** <https://isotc.iso.org/livelink/livelink/open/jtc1sc29wg2>

**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC 1/SC 29/WG 2**

**MPEG TECHNICAL REQUIREMENTS**

**ISO/IEC JTC 1/SC 29/WG 2 N452**

**Online – April 2025**

|  |  |
| --- | --- |
| **Status** | Approved |
| **Title** | Draft Gaussian splat coding requirements |
| **Source** | WG2, MPEG Technical Requirements |
| **Serial Number** | 25075 |

# Abstract

This document provides Draft requirements for Gaussian splat coding. Two different paradigms are being investigated: the Gaussian splat input coding paradigm with a fixed representation called I-3DGS and the Alternative Gaussian splat paradigm where other potential representations are explored and called A-3DGS. Draft requirements are provided in the context of these tracks.

# Introduction

Gaussian splatting [1] is a recently developed immersive representation that provides photorealistic visual quality for novel view synthesis (with 6 degrees of freedom or 6DoF). Compared to other immersive representations such as meshes, multi-view or point clouds, this representation can handle complex visual content including reflective (non-Lambertian) surfaces or volumetric effects (such as fire, smoke etc.) with high visual quality. This representation is typically learned from single or multiple cameras inputs and yields compelling results with relatively fast training times. Furthermore, Gaussian splatting also allows for real-time rendering with photorealistic visual quality. However, the representation data size is large, and there is a need for efficient compression technologies for Gaussian splatting. In addition, mobile terminals are well deployed, and Gaussian splats are a good candidate for representing immersive content, therefore lightweight compression is also needed. For all these reasons, MPEG started an exploration activity on Gaussian splat coding (GSC) and this document provides draft use cases for it.

MPEG currently explores two different paradigms:

* **I-3DGS paradigm/track:** The I-3DGS process involves taking trained Gaussian splat (GS) data in the “INRIA” format [1] as input. The output is a symmetric result in the same format as the input.
* **A-3DGS paradigm/track:** this exploration track welcomes content in any GS format that is for example learned from multiple camera views. Notably, A-3DGS operates under the assumption that ground truth images are accessible, and training can be explored to generate GS inputs.

Figure 1 illustrates both I-3DGS and A-3DGS tracks.

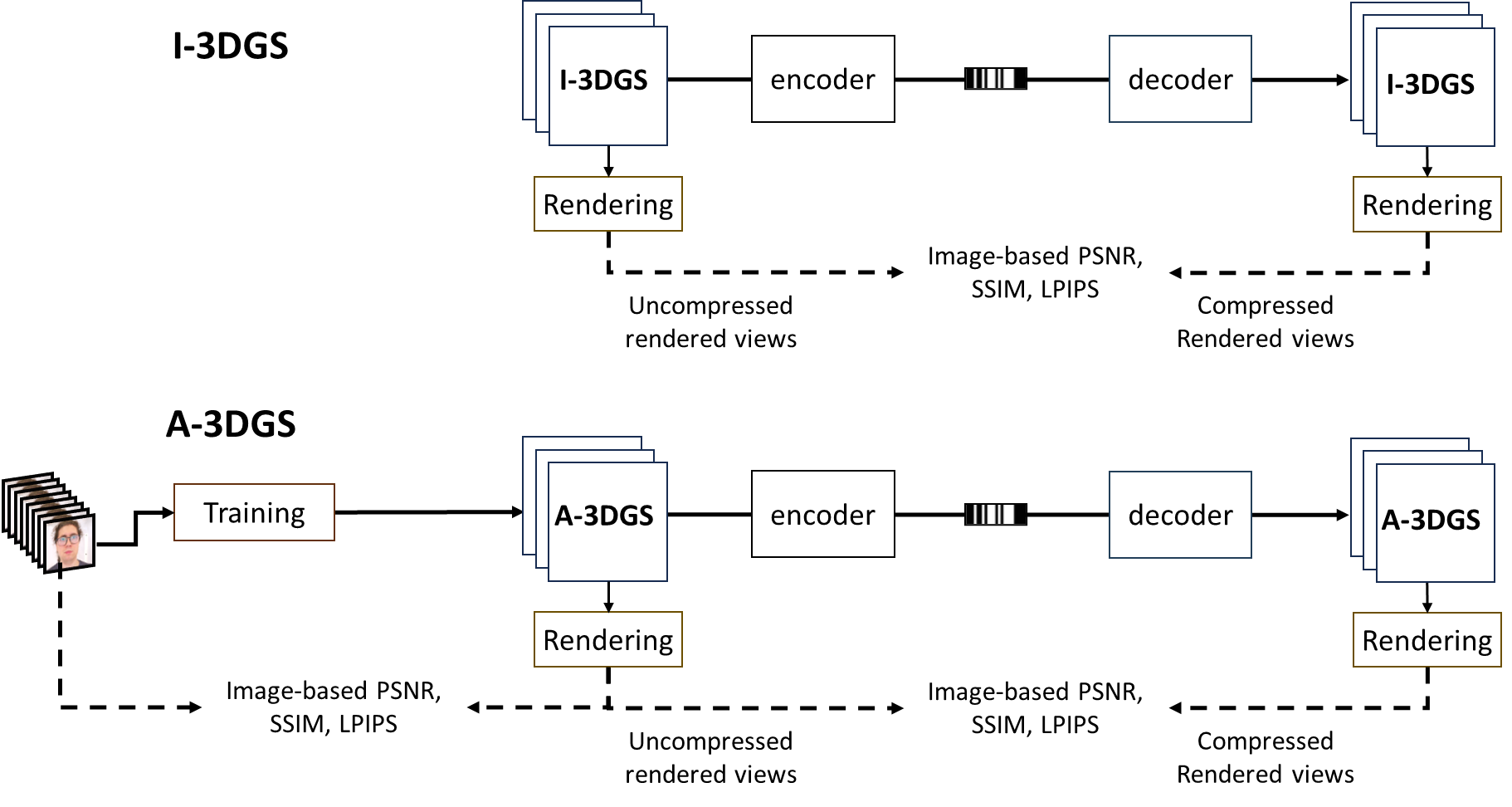


Figure I-3DGS and A-3DGS tracks for Gaussian splat coding

Along with these tracks, MPEG has identified a need for lightweight coding technology that would enable the decoding and consumption of such representation on mobile devices, such as smartphones, glasses, Head-mounted Displays etc.

A list of definitions of terms used in this document is provided in Table 1.

Table Definitions of terms used in this document

|  |  |
| --- | --- |
| **GS (Gaussian Splat)** | The primitive representation used by 3D Gaussian Splatting and its variants, typically represented in the form of geometry and attributes or in the form or a point cloud and additional attributes. |
| **GS frame** | A static GS representation at a given time instant. |
| **Geometry** | The locations in 3D space of the points in the GS, i.e., the (x,y,z) coordinates of the points. |
| **Attribute** | A value or set of values, other than geometry, associated with a point, e.g., spherical harmonics, scale, rotation, opacity. |
| **Lossless GS compression** | A GS is losslessly encoded if and only if its geometry and attributes are compressed in a lossless manner. |
| **Lossy GS Compression** | Lossy GS compression permits different tradeoffs between bitrate and quality. Such trade-offs could be achieved by encoding in lossy manner any components of the GS clouds (geometry and attributes). |
| **Tracked GS** | A GS sequence where the number of Gaussian Splats is the same for each frame and in which the correspondence between each primitive of successive frames is stable over at least a group of frames. |
| **Non-tracked GS** | A GS sequence where the number of Gaussian splats may vary for each frame. |
| **6-DoF** | Six degrees of freedom, refers to the six mechanical degrees of freedom of movement of a rigid body in three-dimensional space and includes three translation directions and three rotation angles. |
| **Non-Lambertian surfaces** | Surfaces that cause specific types of reflectance such that their color change with the observer’s viewpoint (sometimes referred to as view-dependent appearance). |

# Requirements

This section details the currently considered requirements for I-3DGS and A-3DGS tracks. Requirements are classified in two sections; representation requirements and coding requirements, respectively.

* 1. **Representation requirements**

Representation requirements are separated into common requirements and specific requirements. Common representation requirements apply to all tracks, i.e., I-3DGS and A-3DGS. Besides the common representation requirements, a list of specific requirements is provided for I-3DGS.

* + 1. **Common representation requirements**

|  |  |
| --- | --- |
| Requirement | Description |
| (r1) | The representation shall enable 6DoF photorealistic rendering. |
| (r2) | The representation shall be use-case-agnostic (the same representation shall be able to represent, e.g., scenes, actors or objects). |
| (r3) | The representation shall support static and dynamic scenes. |
| (r4) | The representation shall support non-Lambertian, semi-transparent content properties. |

* + 1. **Specific representation requirements for I-3DGS**

|  |  |
| --- | --- |
| Requirement | Description |
| (r5) | The I-3DGS frame representation shall support:   1. 3D positions: (X, Y, Z) coordinates with a specification of its precision and dynamic range. 2. Pre-defined attributes: the attributes associated with each 3D position including colour, rotation, scale, opacity, diffuse (sh0) and specular (sh1, sh2, sh3) spherical harmonics coefficients as illustrated in Equation 1. 3. User-defined attributes per 3D position. |
| (r6) | The I-3DGS sequence shall support a sequence of multiple I-3DGS frames. |
| (r7) | The I-3DGS sequence shall support tracked and non-tracked sequences |
| (r8) | The I-3DGS frame representation shall offer the option to support compact representations of its attributes and positions. |

* 1. **Coding requirements**

Coding requirements are separated into common coding requirements, which are common to all use cases that have been identified for Gaussian splat coding, and specific coding requirements, which apply to a subset of use cases. All coding requirements, common or specific, apply to both I-3DGS and A-3DGS tracks.

* + 1. **Common coding requirements**

|  |  |
| --- | --- |
| Requirement | Description |
| (c1) | The specification shall support lossy compression with variable bit-rate. |
| (c2) | The specification shall support decoding with low complexity, i.e., allow for real-time decoding on low power devices such as mobile devices, glasses and Head Mounted Displays |
| (c3) | The specification shall support decoding in real-time, i.e., the complexity shall allow for feasible implementation of decoding within the constraints of the available technology at the expected time of usage, on high-end devices such as workstation PCs, gaming laptops, rendering farms and cloud-based rendering services. |
| (c4) | The specification shall support means of efficient compression to save storage and/or to transmit compressed representation with various (fixed and mobile) networks. |
| (c5) | The specification shall support quality scalability (e.g., progressive coding). |
| (c6) | The specification shall support spatial random access. |
| (c7) | The specification shall provide support for metadata related to camera parameters |

* + 1. **Specific coding requirements**

|  |  |
| --- | --- |
| Requirement | Description |
| (c8) | The specification shall offer the option to support lossless compression. |
| (c9) | The specification shall support temporal scalability. |
| (c10) | The specification shall support spatial scalability, for example in terms of levels of detail (or levels of density) of primitives. |
| (c11) | The specification shall support temporal random access. |
| (c12) | The specification shall support mechanisms providing error resilience to transmission errors. |
| (c13) | The specification shall provide support for a preview (or thumbnail) mode that can be rendered fast. |

# References

[1] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 2023. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. ACM Transactions on Graphics 42, 4 (July 2023). https://repo-sam.inria.fr/fungraph/3d-gaussian-splatting/