 ISO/IEC JTC 1/SC 29/WG 2 N358

**ISO/IEC JTC 1/SC 29/WG 2  
MPEG Technical requirements   
Convenorship: SFS (Finland)**

**Document type:** Output Document

**Title:** Requirements for AI-based Point Cloud Coding

**Status:** Approved

**Date of document:** 2024-04-26

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

**Expected action:** None

**Action due date:** None

**No. of pages:** 7 (with cover page)

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

**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-2>

**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 N358**

**Rennes, France – April 2024**

|  |  |
| --- | --- |
| **Title:** | **Requirements for AI-based Point Cloud Coding** |
| **Source:** | **WG 2, MPEG Technical requirements** |
| **Status** | **Approved** |
| **Serial Number** | **23868** |
|  |  |

# Introduction

While so far, the most common way of representing the visual component of the world has been to take the output of a camera, compress it for transmission and storage using one of the MPEG video coding standards and eventually decode it and present it on 2D displays, there are now more and more devices that capture and present 3D representations of the world.

A point cloud is defined as a multi-set of points where a point is described by its 3D position with (x,y,z) coordinates(e.g., where x, y, and z have finite (or fixed) precision and dynamic range) and (optionally) a set of scalar/vector attributes. Typically, each point in a cloud has the same number of attributes attached to it. Point clouds can be used to reconstruct an object or a scene as a composition of such points. Point clouds can be captured using multiple cameras and depth sensors in various setups and may be made up of thousands up to billions of points in order to represent realistically reconstructed scenes.

As compression technologies are needed to reduce the amount of data required to represent a point cloud, MPEG is developing an **AI-based** Point Cloud Compression (PCC) standard targeting for use in real-time communications, applications such as geographic information system (GIS), Computer-aided design (CAD), and cultural heritage, autonomous driving & robotics, XR &Gaming, Maps, and Building Information Modeling (BiM). Given the above applications, MPEG has identified the following categories of PCC: AI-based static point cloud compression (Category 1), AI-based dynamic point cloud compression (Category 2), and AI-based dynamically acquired/fused point cloud compression (Category 3).

# Note

*In reading this document, the reader is advised to take an attention to the usage of the words: ‘shall’ and ‘should’. Whenever used, the term ‘shall’ indicates that it is a mandatory statement. Whenever used, the term ‘should’ indicates that it is not mandatory, but desirable.*

# Requirements for AI-based Point Cloud Coding

This document presents the requirements for MPEG AI-based Point Cloud Coding. The requirements capture the use cases defined in [1] and [2] but are not limited to these use cases only.

Note: Section 3.1 and 3.2 includes requirements previously defined for 3D Point Cloud Representation and for 3D Point Cloud Compression with traditional methods; section 3.3 provides new requirements for the AI-based Point Cloud Compression framework.

# 3D Point Cloud Representation

Requirement

MPEG AI-based PCC provides means for encoding and decoding 3D point clouds as defined in Table 1 for each category.

Specification

The 3D point cloud representation supports:

1. 3D positions: the (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, reflectance, normal vectors and transparency.
3. Generic (i.e., user-defined) attributes per 3D position
4. View-dependent attributes per 3D position
5. Time-varying point clouds: point clouds captured or represented with timed information.

# 3D Point Cloud Coding

Requirement

MPEG AI-based PCC shall provide means for efficient coding and compression for storage, streaming or downloading of 3D point clouds as defined in Table 2 for each category.

Specification

MPEG AI-based PCC support:

1. ***Lossy compression:*** means to control the bitrate shall be supported.
2. ***Lossless geometry compression:*** the reconstructed position shall be mathematically identical to the original. The number of points reconstructed from the compressed point cloud is the same as the original. The reordering of points during compression is permissible.
3. ***Lossless attribute compression:*** the reconstructed attributes shall be mathematically identical to the original. The number of points reconstructed from the compressed point cloud is the same as the original. The reordering of points during compression is permissible.
4. ***Near-lossless geometry compression***: The number of points after compression remains the same as the original, but the point locations after compression may not be mathematically identical, but the error between the original and compressed points is always less than the given error margin.
5. ***Near-lossless attribute compression:*** The number of points after compression remains the same as the original, but the point attributes after compression may not be mathematically identical, but the error between the original and compressed attributes is always less than the given error margin.
6. ***Temporal variations*** (e.g., dependency among temporal frames) of point clouds shall be supported.
7. ***Low latency:*** Encode plus decode as low as one point cloud frame duration shall be supported. For some applications, an even lower latency should be supported.
8. ***Low complexity:*** The complexity shall allow for feasible implementation of encoding and decoding within the constraints of the available technology at the expected time of usage.
9. ***Temporal scalability:*** The dependency of frames should be structured such that some frames can be dropped from the bitstream.
10. ***Spatial scalability:*** The compressed bitstream should be structured with more than one layer to decode the points of the current layer predicted from the points from the lower layer(s) which provides a coarse approximation (i.e., a lower number of points) of the entire point cloud.
11. ***Region-based spatial scalability:*** The compressed bitstream should be structured with more than one layer such that certain regions of interest may have a higher density with additional layers; where the layers may be predicted from the lower layer(s).
12. ***Quality scalability:*** A point cloud should be coded at a single spatial resolution but at different qualities (or bit depths). The data and decoded samples of lower qualities may be used to predict data or samples of higher qualities to reduce the bit rate to code the higher qualities.
13. ***Spatial random access:*** it shall be possible to decode the point-cloud corresponding to a region without having to decode the entire bitstream.
14. ***Temporal random access*** shall be possible.
15. ***Error resilience:*** it shall be possible to cope with packet loss without having to retransmit the entire point cloud.
16. ***Parallel encoding and decoding:*** The design should support parallel processing implementation with low cost in terms of bitrate overhead.
17. ***Separable attribute and geometry coding*** (e.g. appearance change: while geometry remain the same, the attribute can be replaced).
    * Geometry only coding is allowed.
    * Multiple attributes coding support is allowed.
18. ***Geometry Precision:*** at least up to 18 bits for category 3 (acquired) up to 22 bits for category 3 (fused), between 10 and 20 bits for category 1 and between 10 bits and 12 bits for category 2.

# AI-Based Point Cloud Coding requirements.

Requirement

MPEG AI-based PCC requirement provides an AI framework as defined inTable 2 for each category.

Specification

MPEG AI-based PCC supports:

1. ***Model architecture****:* shall be stable. It is not allowed to update the network architecture.
2. ***Model download and Update:***

* Model update on the fly (e.g. sequence specific weight update) shall not be required. If proposed, its size shall be counted in the bd rate.
* Model update/download on demand (e.g. model weight update) is allowed. The rate required to download a model (possibly per rate point) prior to operation shall not be counted in the bd rate.

1. ***Inference Reproducibility****:* The results shall be reproduceable and within a tolerance level (0.2% for overall distortion) on the same or on a different platform.

# Annex A – AI-based Point Cloud Coding Requirements per Category

All the requirements except for ‘not-applicable’ cases are expected to be supported during the AI-based Point-cloud coding standardization phase.

the Category 3 is split into Category 3A (acquired content) and Category 3B (fused content), and only Category 3A is considered.

# AI-based Point Cloud Representation requirements per category.

Table 1 defines per category, the AI-based point cloud representation requirements for AI-GC as per their definition in Section 3.1.

Table 1 –Requirements of Point Cloud Representation per Category

|  |  |  |  |
| --- | --- | --- | --- |
| Requirements | Category 1 | Category 2 | Category 3.A |
| a) 3D positions | P | P | P |
| b) Pre-defined attributes | P (colors) | P (colors) | P  (colors, reflectance) |
| c) Generic attributes | *o* | *o* | *o* |
| d) View-dependent attributes | *o* | *o* | *o* |
| e) Time-varying | - | P | P |

(‘P’ = Required, ‘*o*’ = Optional ‘-’ = Not applicable)

# AI-based Point Cloud Compression requirements per category.

Table 2 defines per category, the AI-based Point Cloud Compression requirements for AI-GC as per their definition in sections 3.2 and 3.3.

Table 2 –Requirements of Point Cloud Compression per Category

|  |  |  |  |
| --- | --- | --- | --- |
| Requirements | Category 1 | Category 2 | Category 3.A |
| a) Lossy compression | P | P | P |
| b) Lossless geometry compression | P | P | P |
| c) Lossless attribute compression | P | P | P |
| d) Near-lossless geometry compression | *o* | *o* | P |
| e) Near-lossless attribute compression | *o* | *o* | *o* |
| f) Temporal variations | - | P | P |
| g) Low latency | P | P | P |
| h) Low complexity | P | P | P |
| i) Temporal scalability | - | P | P |
| j) Spatial scalability | *o* | *o* | *o* |
| k) Region-based scalability | *o* | *o* | *o* |
| l) Quality scalability | *o* | *o* | *o* |
| m) Spatial random access | *o* | *o* | *o* |
| n) Temporal random access | - | P | P |
| o) Error resilience | *o* | P | *o* |
| p) Parallel encoding and decoding | *o* | *o* | *o* |
| q-1) geometry only coding | P | P | P |
| q-2) multiple attribute coding | P | P | P |
| r) geometry precision | At least Up to 20 | At least Up to 12 | At least Up to 18 |
| s) Model architecture | P | P | P |
| t-1) on the fly Model Update | P | P | P |
| t-2) on demand Model Update & download | P | P | P |
| u)Inference Reproducibility | P | P | P |

(‘P’ = Required ‘*o*’ = Optional ‘-’ = Not applicable)

# References

[1] n16331 Use Cases for Point Cloud Compression ISO/JCT SC29 WG11 Geneva June 2016

[2] n00331 Use Case Scenarios for AI-GC ISO/IEC JTC1 SC29 WG2 Hannover October 2023