FINAL
DRAFT

INTERNATIONAL
STANDARD

**ISO/IEC
FDIS
23090-18**

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

Secretariat: **JISC**

Voting begins on:
2023-10-19

Voting terminates on:
2023-12-14

Information technology — Coded representation of immersive media —

Part 18:
Carriage of geometry-based point cloud compression data

Reference number
ISO/IEC FDIS 23090-18:2023(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives or www.iec.ch/members_experts/refdocs).

ISO and IEC draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC had received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and https://patents.iec.ch. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

A list of all parts in the ISO/IEC 23090 series can be found on the ISO and IEC websites.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

Introduction

Advances in 3D capturing and rendering technologies have unleashed a new wave of innovation in Virtual/Augmented/Mixed reality (VR/AR/MR) content creation and communication. Point clouds have arisen as one of the main representations for such applications. Geometry-based point cloud compression data is used for representing sparse dynamically varying point clouds such as those used in vehicular LiDAR or 3D mapping, as well as dense static point clouds used in cultural heritage, and industrial applications.

This document addresses technologies defining the carriage of geometry-based point cloud compression data for storage and delivery purposes. This document includes(but is not limited to):

— Storage of geometry-based point cloud compression data and the associated metadata using the ISO Base Media File Format (ISOBMFF) as specified in ISO/IEC 14496-12;

— Storage of non-timed geometry-based point cloud compression data and the associated metadata using HEVC Image File Format (HEIF) as specified in ISO/IEC 23008-12;

— Encapsulation, signalling, and streaming of geometry-based compression data in a media streaming system, e.g., dynamic adaptive streaming over HTTP (DASH) as specified in ISO/IEC 23009-1 or MPEG media transport (MMT) as specified in ISO/IEC 23008-1.

Information technology — Coded representation of immersive media —

Part 18:
Carriage of geometry-based point cloud compression data

1 Scope

This document specifies a media format that enables the storage and delivery of geometry-based point cloud compression data. The geometry-based point cloud compression data can be timed or non-timed. It supports flexible extraction of geometry-based point cloud compression data at delivery or decoding time.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

IEEE 754-2019, *IEEE Standard for Floating-Point Arithmetic.*

ISO/IEC 9834-1, *(Rec. ITU-T X.660), Information technology — Procedures for the operation of object identifier registration authorities — Part 1: General procedures and top arcs of the international object identifier tree*

ISO/IEC 9834-8, *(Rec. ITU-T X.667), Information technology — Procedures for the operation of object identifier registration authorities — Part 8: Generation of universally unique identifiers (UUIDs) and their use in object identifiers*

ISO/IEC 14496-12, *Information technology — Coding of audio-visual objects — Part 12: ISO base media file format*

ISO/IEC 23008-12:2022, *Information technology — MPEG systems technologies — Part 12: Image File Format*

ISO/IEC 23009-1:2022, *Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 1: Media presentation description and segment formats*

ISO/IEC 23090-9:2023, *Information technology — Coded representation of immersive media (MPEG-I) — Part 9: Geometry-based point cloud compression*

W3C Recommendation, *XML schema part 1: Structures*

W3C Recommendation, *XML schema part 2: Datatypes*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 23090-9 and the following apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

3.1
attribute track
volumetric visual track which carries ADUs of one instance of a particular attribute component of coded point cloud frames

3.2
attribute tile track
G-PCC tile track (3.10) which carries ADUs of one instance of a particular attribute component corresponding to one or more G-PCC tiles

3.3
G-PCC bitstream
sequence of bits of the coded point cloud sequence

Note 1 to entry: G-PCC bitstream is specified in ISO/IEC 23090-9.

3.4
G-PCC bitstream track
volumetric visual track which carries the entire coded point cloud sequence

3.5
G-PCC content
volumetric visual media that that consists of one or more point cloud frames

Note 1 to entry: Each point cloud frame includes a number of points, identified by their positions in 3D space, and the associated attributes (e.g. color) at a particular time instance.

3.6
G-PCC component track
volumetric visual track which carries DUs of one instance of a particular G-PCC component

Note 1 to entry: There are two types of G-PCC component tracks: one is *geometry track* (3.7) and the other is *attribute track* (3.1).

3.7
geometry track
volumetric visual track which carries GDUs of the coded point cloud frames

3.8
geometry tile track
G-PCC tile track (3.10) which carries GDUs of one or more G-PCC tiles

3.9
G-PCC player
application responsible for receiving files/segments or accessing files locally, decapsulating files/segments, decoding the G-PCC bitstream, reconstructing point cloud frames from the decoded G-PCC bitstream, and rendering the point cloud frames

3.10
G-PCC tile track
volumetric visual track which carries either any of DUs corresponding to one or more G-PCC tiles

3.11
G-PCC tile base track
volumetric visual track which carries DUs which can be applied across the associated *G-PCC tile tracks* (3.10)

4 Abbreviated terms

ADU	attribute data unit (specified in ISO/IEC 23090-9)
APS	attribute parameter set (specified in ISO/IEC 23090-9)
DASH	dynamic adaptive streaming over HTTP (specified in ISO/IEC 23009-1)
DU	data unit (specified in ISO/IEC 23090-9)
FBDU	frame boundary marker data unit (specified in ISO/IEC 23090-9)
FSAP	frame-specific attribute properties (specified in ISO/IEC 23090-9)
GDU	geometry data unit (specified in ISO/IEC 23090-9)
GPS	geometry parameter set (specified in ISO/IEC 23090-9)
HTTP	Hyper-text transfer protocol
HEIF	HEVC image format (specified in ISO/IEC 23008-12)
ISOBMFF	ISO base media file format (specified in ISO/IEC 14496-12)
MMT	MPEG media transport (specified in ISO/IEC 23008-1)
SPS	sequence parameter set (specified in ISO/IEC 23090-9)

5 Overview

5.1 Overall architecture for carriage of geometry-based point cloud compression data

Geometry-based point cloud compression (G-PCC) provides the method for efficiently compressing the point cloud sequence which consists of one or more point cloud frames. Each point cloud frame consists of a number of points and each point is a tuple of a three-dimensional position and attribute values for every attribute present in the point cloud.

The coded point cloud sequence forms a G-PCC bitstream comprising of data represents a volumetric encoding of point clouds consisting of a sequence of point cloud frames. Each point cloud frame includes a number of points, identified by their positions in 3D space, and their associated attributes at a particular time instance. The number of points can vary from one frame to another.

parameter sets and slices of coded point cloud frames. Every slice includes a GDU which codes the slice geometry and ADUs or defaulted attribute DUs which code the slice attributes. The group of slices may be associated with spatial regions in a point cloud to aid spatial access.

Figure 1 shows the overall architecture for a typical content flow process for carriage of G-PCC data and it is applicable to both live and on-demand use cases.

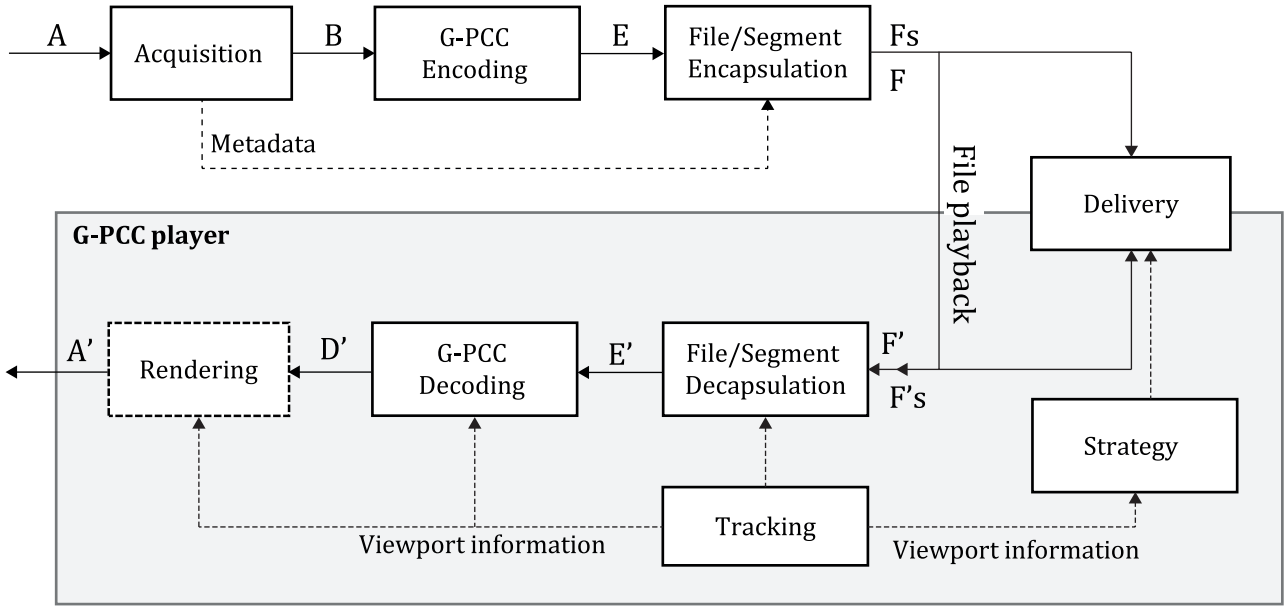

Figure 1 — Overall architecture flow process of for carriage of G-PCC data

A real-world visual scene (A) is captured by a set of cameras or a camera device with multiple lenses and sensors. A virtual visual scene (A) is also captured by virtual camera. The acquisition results in a point cloud sequence comprising of one or more point cloud frames (B). The point cloud sequence can be timed or non-timed. Each point cloud frame includes a set of points, identified by their positions in a three-dimensional Cartesian coordinate system, and associated attributes at a particular time instance. All points in the same point cloud frame have the same number of attributes. The number of points may vary from one frame to another. Each point cloud frame is coded as a sequence of slices and each slice comprises a sequence of encapsulated DUs. One or multiple point cloud frames are multiplexed into a G-PCC bitstream(E). The G-PCC bitstream are then encapsulated into a media file for file playback (F) or a sequence of an initialization segment and media segments for streaming (F_s), according to a particular media container file format. The metadata which can contribute to interpret and to consume the point cloud frames is encapsulated into the file or the segments. The point cloud metadata can describe, for example, the mapping between points to spatial regions within a point cloud. The segments F_s are delivered using a delivery mechanism to a G-PCC player.

The file that the file encapsulator outputs (F) is identical to the file that the file decapsulator inputs (F').

The G-PCC player processes the file (F') or the received segments (F'_s) and extracts the G-PCC bitstream (E') and parses the metadata. The G-PCC bitstream is then decoded into one or multiple point cloud frames (D') and the point clouds are reconstructed from the decoded point cloud frames (D'). The reconstructed point clouds are rendered and displayed onto the screen of a head-mounted display or any other display device based on the viewport information, such as the current viewing position, viewing orientation, or the field of view information, which is determined by various types of sensors. Besides being used by the G-PCC player to access the appropriate part of the point clouds, the viewport information can also be used for determining which tracks are extracted from the file. In viewport-dependent delivery, the viewport information is also passed to the strategy module, which determines the segments to be received based on the current viewport.

This process is applicable to both live and on-demand use cases.

The following interfaces are specified in this document:

— F/F': media file including the specification of the track formats in Clause 7 for timed G-PCC data, in Clause 8 for non-timed G-PCC data, and in Clause 9 for metadata in ISOBMFF.

— Clause 10 specifies the delivery related interfaces for DASH delivery.

— Clause 11 specifies the delivery related interfaces for MMT delivery.

The other interfaces in Figure 1 are not specified in this document.

5.2 Referenceable code points

5.2.1 Brands

The brands are used in this document to indicate conformance points to an encapsulation mode and a specific set of tools that are defined in this document. It may be indicated in the `FileTypeBox`.

The brands specified in this document are listed in Table 1 and defined in Annex A.

Table 1 — Brands specified in this document

Brand identifier	Subclause in this document	Description
gpst	A.2.1	Single track encapsulation
gpmt	A.2.2	Multiple track encapsulation
gppa	A.2.3	Encapsulation with partial access support
gpci	A.3	Non-timed G-PCC encapsulation

5.2.2 Sample entry type

The sample entry type specified in this document are listed in Table 2.

Table 2 — Sample entry types specified in this document

Sample entry type	Subclause in this document	Description
gpe1	7.3.2	For use with the single track encapsulation with all parameter set data units carrying SPS, GPS, and APS carried in decoder configuration record
gpeg	7.3.2	For use with the single track encapsulation with all parameter set data units carrying SPS, GPS, and APS carried in decoder configuration record and in track samples
gpc1	7.4.2	For use with the multiple track encapsulation with all parameter set data units carrying SPS, GPS, and APS carried in decoder configuration record
gpcg	7.4.2	For use with the multiple track encapsulation with all parameter set data units carrying SPS, GPS, and APS carried in decoder configuration record and in track samples
gpeb	7.5.2.1	For use with a tile base track with G-PCC tile track(s) containing DUs of all components
gpcb	7.5.2.1	For use with a tile base track with G-PCC tile track(s) containing DUs of one instance of a particular component
gpt1	7.5.3.1	For use with a G-PCC tile track
gpdr	9.1.3.2	For use with a timed metadata track indicating the dynamic spatial regions that are dynamically changing over time
gpdv	9.2.4.2	For use with a timed metadata track indicating viewport information that are dynamically changing over time

5.2.3 Box types

The box types specified in this document are listed in bold in Table 3. Mandatory boxes are marked with an asterisk. Box types without a four-character code are marked with '-' in the structure.

Table 3 — Box types specified in this document and their relation to boxes not specified in this document

Box types, structure, and cross-reference												
moov										*	ISOBMFF	*container for all the metadata*
	trak									*	ISOBMFF	*container for an individual track or stream*
		mdia								*	ISOBMFF	*container for the media information in a track*
			minf							*	ISOBMFF	*media information container*
				stbl						*	ISOBMFF	*sample table box, container for the time/space map*
					stsd					*	ISOBMFF	*sample descriptions (codec types, initialization etc.)*
						-					ISOBMFF	*visual sample entry*
						-					6.1.3	*volumetric visual sample entry*
							gpcC				7.2.2	*G-PCC decoder configuration box*
							ginf				7.2.3	*G-PCC component information box*
							gptC				7.5.3.1	*G-PCC tile configuration box*
							gpsr				8.3.3	*Static spatial region information box*
							gvpC				9.2.3	*Viewport information Configuration Box*
						gpdr					9.1.3.2	*Dynamic spatial region timed metadata sample entry*
							gpsr				8.3.3	*Static spatial region information box*
						gpdv					9.2.4.2	*Viewport information timed metadata sample entry*
							gvpC				9.2.3	*Viewport information Configuration Box*
meta											ISOBMFF	*Metadata*
	grpl										ISOBMFF	*group list box*
		vpta									8.4.1	*Viewport association box*
	iprp										ISOBMFF	*item properties box*
		ipco									ISOBMFF	*item property container box*
			gpcC								8.3.1	*G-PCC configuration item property*
			ginf								8.3.2	*G-PCC component information item property*
			gpsr								8.3.3	*G-PCC spatial region item property*
			gpti								8.3.4	*G-PCC tile information item property*

5.2.4 Track reference types

The track reference types specified in this document are listed in Table 4.

Table 4 — Track reference types specified in this document

Track reference type	Subclause in this document	Description
gpca	7.4.4	Referenced track is an attribute track
gpbt	7.5.5	Referenced track is a G-PCC tile track

5.2.5 Entity grouping types

The entity grouping types specified in this document are listed in Table 5.

Table 5 — Entity grouping types specified in this document

Entity groping type	Subclause in this document	Description
vpta	8.4.1	Viewport association between G-PCC items and the viewport information timed metadata track

5.2.6 Sample grouping types

The sample grouping types specified in this document are listed in Table 6.

Table 6 — sample grouping types specified in this document

Sample grouping type	Subclause in this document	Description
gtii	7.2.4	Tile inventory sample group

5.2.7 Uniform resource names

The URNs specified in this document are listed in Table 7.

Table 7 — URNs specified in this document

URN	Subclause in this document	Description
urn:mpeg:mpegI:gpcc:2023	10.2.2.1	Namespace for the XML elements and attributes specified in this document
urn:mpeg:mpegI:gpcc:2023:component	10.2.2.2	Scheme identifier for the G-PCC component DASH MPD descriptor
urn:mpeg:mpegI:gpcc:2023:gpc	10.2.2.3	Scheme identifier for the G-PCC content DASH MPD descriptor
urn:mpeg:mpegI:gpcc:2023:gpsr	10.3.1	Scheme identifier for the G-PCC static spatial region DASH MPD descriptor
urn:mpeg:mpegI:gpcc:2023:tileID	10.3.3.1	Scheme identifier for the G-PCC tile id DASH MPD descriptor
urn:mpeg:mpegI:gpcc:2023:rv	10.4.1	Scheme identifier for the static recommended viewports DASH MPD descriptor
urn:mpeg:mmt:app:gpcc:2023	11.3.1	Scheme identifier for the MMT G-PCC specific signalling messages

6 Volumetric media

6.1 Volumetric visual media

6.1.1 General

A volumetric visual track shall be identified by the volumetric visual media handler type `'volv'` in the `HandlerBox` of the `MediaBox`, as defined in ISO/IEC 14496-12, and by a volumetric visual media header as defined in subclause 6.1.2.

Multiple volumetric visual tracks can be present in the file.

6.1.2 Volumetric visual media header

6.1.2.1 Definition

Box Type: `'vvhd'`
Container: `MediaInformationBox`
Mandatory: Yes
Quantity: Exactly one

Volumetric visual tracks shall use a `VolumetricVisualMediaHeaderBox` in the `MediaInformationBox` as defined in ISO/IEC 14496-12.

6.1.2.2 Syntax

```
aligned(8) class VolumetricVisualMediaHeaderBox
   extends FullBox('vvhd', version = 0, 1) {
}
```

6.1.2.3 Semantics

`version` is an integer that specifies the version of this box.

6.1.3 Volumetric visual sample entry

6.1.3.1 Definition

Volumetric visual tracks shall use a `VolumetricVisualSampleEntry`.

6.1.3.2 Syntax

```
class VolumetricVisualSampleEntry(codingname)
   extends SampleEntry (codingname){
   unsigned int(8)[32] compressorname;
   // other boxes from derived specifications
}
```


6.1.3.3 Semantics

`compressorname` is a name, for informative purposes. It is formatted in a fixed 32-byte field, with the first byte set to the number of bytes to be displayed, followed by that number of bytes of displayable data encoded using UTF-8, and then padding to complete 32 bytes total (including the size byte). The field may be set to 0.

6.1.4 Volumetric visual sample group entry

```
abstract class VolumetricVisualSampleGroupEntry (unsigned int(32) grouping_type)
extends SampleGroupDescriptionEntry (grouping_type)
{
}
```

6.1.5 Volumetric visual samples

The format of a volumetric visual sample is defined by the coding system.

7 Timed G-PCC data storage in ISOBMFF

7.1 General

This clause specifies the below encapsulation of G-PCC bitstream in tracks within a file, and only one of the below encapsulations shall be used at the same time.

— Single track encapsulation of G-PCC bitstream, where one track carries the entire coded point cloud sequence, as specified in subclause 7.3.

— Multiple track encapsulation, where the coded bitstream of single G-PCC component of the point cloud sequence is encapsulated into a separate track, as specified in subclause 7.4.