**ISO/IEC JTC 1/SC 29/WG 11  
Coding of moving pictures and audio  
Convenorship: Japan (JISC)**

**Document type: Approved WG 11 document**

**Title: Text of ISO/IEC CD 23090-18 Carriage of Geometry-based Point Cloud Compression Data**

**Status: Approved**

**Date of document: 2020-07-30**

**Source: Systems**

**Expected action:**

**No. of pages: 30**

**Email of convenor:** ostermann@tnt.uni-hannover.de

**Committee URL:** <http://isotc.iso.org/livelink/livelink/open/jtc1sc29>

**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

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

**CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC 1/SC 29/WG 11 N19442**

**OnLine Meeting – July 2020**

**Source: Systems**

**Status : Committee Draft**

**Title: Text of ISO/IEC CD 23090-18 Carriage of Geometry-based Point Cloud Compression Data**

**Editors:** [**Sejin Oh**](mailto:sjin.oh@lge.com)**, Ryohei Takahashi, Youngkwon Lim**

**ISO/IEC 23090-18:2020 (E)**

Date : 2020-07-30

ISO/IEC JTC 1/SC 29/WG 11

Secretariat: JISC

**Information technology** — **Coded representation of immersive media (MPEG-I) — Part 18: Carriage of Geometry-based Point Cloud Compression Data**

CD stage

**Warning for WDs and CDs**

This document is not an ISO International Standard. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an International Standard.

Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

© ISO 20XX

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO’s member body in the country of the requester.

ISO copyright office

CP 401 • Ch. de Blandonnet 8

CH-1214 Vernier, Geneva

Phone: +41 22 749 01 11

Fax: +41 22 749 09 47

Email: copyright@iso.org

Website: www.iso.org

Published in Switzerland

Contents

[Foreword vii](#_Toc46906017)

[Introduction viii](#_Toc46906018)

[1 Scope 1](#_Toc46906019)

[2 Normative references 1](#_Toc46906020)

[3 Terms and definitions 1](#_Toc46906021)

[4 Abbreviations 2](#_Toc46906022)

[5 Overview 2](#_Toc46906023)

[5.1 Organization of this document 2](#_Toc46906024)

[5.2 Overall architecture for carriage of geometry-based point cloud compression data 2](#_Toc46906025)

[6 Volumetric media 3](#_Toc46906026)

[6.1 Volumetric visual media 3](#_Toc46906027)

[6.1.1 General 3](#_Toc46906028)

[6.1.2 Volumetric visual media header 4](#_Toc46906029)

[6.1.3 Volumetric visual sample entry 4](#_Toc46906030)

[6.1.4 Volumetric visual samples 4](#_Toc46906031)

[7 Carriage of geometry-based point cloud compression data in ISOBMFF 5](#_Toc46906032)

[7.1 General 5](#_Toc46906033)

[7.2 Common boxes and data structures 5](#_Toc46906034)

[7.3 Encapsulation of G-PCC data in ISOBMFF 8](#_Toc46906039)

[7.4 Partial Access of G-PCC data 12](#_Toc46906045)

[8 Encapsulation and signalling in DASH 14](#_Toc46906057)

[8.1 General 14](#_Toc46906059)

[8.2 DASH MPD descriptors for G-PCC media 14](#_Toc46906060)

[8.3 GPCC Preselection 15](#_Toc46906063)

[8.4 Supporting multiple versions of GPCC media 16](#_Toc46906064)

[Annex A (normative) File format toolsets and brands 17](#_Toc46906065)

[Annex B (normative) GPCC DASH Schema 18](#_Toc46906066)

[Annex C (informative) DASH MPD examples 19](#_Toc46906067)

[C.1. Grouping of G-PCC Components 19](#_Toc46906068)

[C.2. Supporting multiple versions of GPCC media 19](#_Toc46906069)



Editor’s notes: The change marks are relative to N19286 (WD of ISO/IEC 23090-18) The integrated changes include all the agreed changes as noted in M54774-v3 (MPEG#131 Notes of Systems V3C BoG session)

**List of agreements integrated:**

**MPEG#129(as per M52893-v3)**

* M52289 : agreed as base text for WD
* M52242 : Editorial comments are accepted except more than one contiguous G-PCC TLV with same type are considered as a single sub-sample.

**MPEG#130**

* M53465 : agreed to add the architecture and changes by merging between single track encapsulation and multiple track encapsulation. (<http://mpegx.int-evry.fr/software/MPEG/Systems/PCC-SYS/G-PCC/issues/1>)
* M53466-v2 : agreed to add the below for the support of spatial access of G-PCC data (<http://mpegx.int-evry.fr/software/MPEG/Systems/PCC-SYS/G-PCC/issues/5>)
  + 3D spatial region information structure
  + signalling of static spatial region information
  + a new timed-metadata track to indicate dynamically changed 3D spatial region information
  + 3D spatial region track grouping.

**MPEG#131 (as per M54774-v3)**

* M54294 (<http://mpegx.int-evry.fr/software/MPEG/Systems/PCC-SYS/G-PCC/issues/8>)
  + agreed to update GPCCSpatialRegionInfoBox and timed metadata track for dynamic spatial region information
  + agreed to remove clause 7.4.3 3D spatial region track grouping in N19286 (WD of ISO/IEC 23090-18) and include it into N19416 Technologies under consideration on Carriage of G-PCC data.
* M54357 (<http://mpegx.int-evry.fr/software/MPEG/Systems/PCC-SYS/G-PCC/issues/10>)
  + agreed to add a tile-based subsample definition and to update the syntax and semantics of codec\_specific\_parameters
* M54463 (<http://mpegx.int-evry.fr/software/MPEG/Systems/PCC-SYS/G-PCC/issues/7>)
  + agreed to add MPEG DASH signalling and to add examples in Annex

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

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 ISO documents 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](https://www.iso.org/directives-and-policies.html)).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see [www.iso.org/patents](https://www.iso.org/iso-standards-and-patents.html)).

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](https://www.iso.org/foreword-supplementary-information.html).

This document was prepared by 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 website.

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. The 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. What is specified in this document includes (but is not limited to):

1. 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,
2. 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 (MPEG-I) — Part 18: Carriage of Geometry-based Point Cloud Compression Data

# Scope

This document specifies a media format that enables to store and deliver geometry-based point cloud compression data. Another important aspect considered by this document is supporting flexible extraction of geometry-based point cloud compression data at delivery and/or decoding time.

# 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.

|  |  |
| --- | --- |
| [ISOBMFF] | ISO/IEC 14496-12, Information technology — Coding of audio-visual objects — Part 12: ISO base media file format |
| [GPCC] | ISO/IEC 23090-9, Information technology — Coded representation of immersive media (MPEG-I) — Part 9: Geometry-based point cloud compression |
| [DASH] | ISO/IEC 23009-1, Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 1: Media presentation description and segment formats |
|  |  |
|  |  |

# Terms and definitions

For the purposes of this document, the following terms and definitions 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 <http://www.electropedia.org/>

— The terms and definitions specified in ISO/IEC 23090-9 [GPCC] apply unless they are overridden by this clause.

3.1

G-PCC content

volumetric media that is encoded using ISO/IEC 23090-9 [GPCC].

3.2.

point cloud frame

set of 3D points specified by their cartesian coordinates (x,y,z) and zero or more corresponding attributes at a particular time instance

3.3

G-PCC tile

a rectangular cuboid inside a bounding box of point cloud frame, which consists of group of slices.

3.4

G-PCC track

volumetric visual track which carries either the coded geometry bitstream or the coded attribute bitstream or both of them.

3.5

G-PCC geometry track

volumetric visual track which carries the coded geometry bitstream.

3.6

G-PCC attribute track

volumetric visual track which carries the coded attribute bitstream

3.7

Bounding box

rectangular cuboid in which the source point cloud frame is included

# Abbreviations

For the purposes of this International Standard, the following abbreviations apply:

|  |  |
| --- | --- |
| APS | Attribute Parameter Set (specified in ISO/IEC 23090-9) |
| G-PCC | Geometry-based Point Cloud Compression |
| GPS | Geometry Parameter Set (specified in ISO/IEC 23090-9) |
| ISOBMFF | ISO Base Media File Format (specified in ISO/IEC 14496-12) |
| PCC | Point Cloud Compression |
| SPS | Sequence Parameter Set (specified in ISO/IEC 23090-9) |
| TLV | Type-length-value bytestream format (specified in ISO/IEC 23090-9) |
|  |  |
|  |  |

# Overview

## Organization of this document

This document is organized as follows:

* Clause 5.2 provides the overall architecture for storing and streaming geometry-based point cloud compression data.
* Clause 6 specifies syntax and semantic of volumetric visual media track.
* Clause 7 specifies extensions to the ISO Base Media File Format for geometry-based point cloud compression data as well as for the associated timed metadata.
* Clause 8 specifies extensions to DASH for geometry-based point cloud compression data.
* Annex A specifies file format toolsets and brands to store G-PCC data in file format.
* Annex B contains the DASH XML schema to support the streaming of G-PCC data.
* Annex C contains some informative DASH MPD examples

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

The below Figure shows a typical content flow process for G-PCC data.

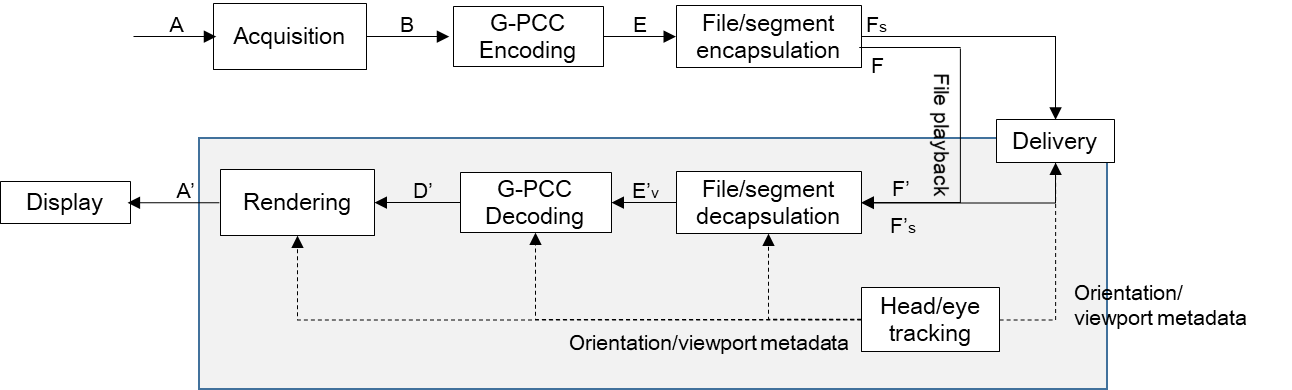


Figure 5.1 Content Flow process for G-PCC data

A real-world visual scene (A) is captured by a set of cameras or a camera device with multiple lenses and sensors. The acquisition results in source point cloud data (Bi). one or multiple point cloud frames are encoded as coded G-PCC bitstream including the coded geometry bitstream and attribute bitstream (E). one or more coded bitstream are then composed into a media file for file playback (F) or a sequence of an initialization segment and media segments for streaming (Fs), according to a particular media container file format. In this document, the media container file format is the ISO Base Media File Format specified in ISO/IEC 14496-12 [ISOBMFF]. The file encapsulator also includes metadata into the file or the segments. The segments Fs are delivered using a delivery mechanism to a player.

The file that the file encapsulator outputs (F) is identical to the file that the file decapsulator inputs (F'). A file decapsulator processes the file (F') or the received segments (F's) and extracts the coded bitstreams (E'v) and parses the metadata. The G-PCC bitstream are then decoded into decoded signals (D'). The decoded point cloud data (D') are rendered and displayed onto the screen of a head-mounted display or any other display device based on the current viewing orientation or viewport. The current viewing orientation is determined by the head tracking and possibly also eye tracking functionality. Besides being used by the renderer to render the appropriate part of decoded point cloud data, the current viewing orientation may also be used by the audio decoders for decoding optimization. In viewport-dependent delivery, the current viewing orientation is also passed to the strategy module, which determines the tracks to be received based on the viewing orientation.

The process described above is applicable to both live and on-demand use cases.

The following interfaces are normatively specified in this document:

* E'v,: coded G-PCC bitstream
* F/F': media file including the specification of the track formats, which may contain constraints on the elementary streams contained within the samples of the tracks; see clause 7.
* Clause 8 specifies the delivery related interfaces for DASH delivery.

The other interfaces in Figure 5.1 are not specified normatively in this document.

# Volumetric media

## Volumetric visual media

[Ed (SOH): Will we use the same syntax and semantics of volumetric media header, handler and volumetric sample entry? If the same one is used both in here and in ISO/IEC 23090-10, it should be specified in ISO/IEC 14496-12 [ISOBMFF] and referenced by this specification]

### 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 [ISOBMFF], and by a volumetric visual media header as defined in this specification.

Multiple volumetric visual tracks may be present in the file.

### Volumetric visual media header

#### 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 [ISOBMFF].

#### Syntax

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

#### Semantics

version is an integer that specifies the version of this box

### Volumetric visual sample entry

#### Definition

Volumetric visual tracks shall use a VolumetricVisualSampleEntry

#### Syntax

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

#### 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.

### Volumetric visual samples

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

# Carriage of geometry-based point cloud compression data in ISOBMFF

## General

Geometry-based point cloud compression data represents a volumetric encoding of point clouds consisting of a sequence of point cloud frames. Each point cloud frame includes the number of points, their positions, and their attributes and it may vary from one frame to another.

This clause specifies the encapsulation of G-PCC bitstream in tracks within a file. A G-PCC bitstream is composed of type-length-value encapsulation structures carrying parameter sets, a coded geometry bitstream and zero or more coded attribute bitstreams. This G-PCC bitstream is stored either in one or more tracks.

## Common boxes and data structures

### G-PCC decoder configuration record

#### Definition

This subclause specifies the G-PCC decoder configuration information for geometry-based point cloud content.

This G-PCC decoder configuration record contains a version field. This version of the specification defines version 1 of this record. Incompatible changes to the record will be indicated by a change of version number. Readers shall not attempt to decode this record or the streams to which it applies if the version number is unrecognised. Compatible extensions to this record will extend it and will not change the configuration version code. Readers should be prepared to ignore unrecognised data beyond the definition of the data they understand.

The values for profile\_idc,profile\_compatibility\_flags,level\_idc shall be valid for all parameter sets that are activated when the stream described by this record is decoded (referred to as "all the parameter sets" in the following sentences in this paragraph). Specifically, the following restrictions apply:

* The profile indication profile\_idc shall indicate a profile to which the stream associated with this configuration record conforms.
* Each bit in profile\_compatibility\_flags may only be set if all the parameter sets set that bit.
* The level indication level\_idc shall indicate a level of capability equal to or greater than the highest level indicated for the highest tier in all the parameter sets.

The setupUnit array shall include G-PCC TLV encapsulation structures that are constant for the stream referred to by the sample entry in which the decoder configuration record is present. The type of G-PCC encapsulation structures is restricted to indicate SPS, GPS, APS, tile inventory as defined in ISO/IEC 23090-9 [GPCC].

#### Syntax

aligned(8) class GPCCDecoderConfigurationRecord {  
 unsigned int(8) configurationVersion = 1;  
 unsigned int(8) profile\_idc;  
 unsigned int(24) profile\_compatibility\_flags;  
 unsigned int(8) level\_idc;

unsigned int(8) numOfSetupUnits;  
 for (i=0; i<numOfSetupUnits; i++) {

tlv\_encapsulation setupUnit; //as defined in ISO/IEC 23090-9  
 }  
 // additional fields  
}

#### Semantics

configurationVersion is a version field. Incompatible changes to the record are indicated by a change of version number.

profile\_idc contains the profile code as defined in ISO/IEC 23090-9 [GPCC].

profile\_compatibility\_flags equal to 1, indicates that the bitstream conforms to the profile indicated by profile\_idc as defined in ISO/IEC 23090-9 [GPCC].

level\_idc contains the level code as defined in ISO/IEC 23090-9 [GPCC].

numOfSetupUnits specifies the number of G-PCC setup units in the decoder configuration record.

setupUnit is an instance of TLV encapsulation structure carrying SPS, GPS, APS, and tile inventory as defined in ISO/IEC 23090-9 [GPCC]

### G-PCC decoder configuration box

A G-PCC decoder configuration box includes the GPCCDecoderConfigurationRecord as defined in 7.2.1

#### Syntax

class GPCCConfigurationBox extends Box('gpcC') {  
 GPCCDecoderConfigurationRecord() GPCCConfig;  
}

#### Semantics

GPCCDecoderConfigurationRecord is defined in ‎7.2.1

### G-PCC component type box

#### Definition

This box indicates the type of G-PCC components, e.g., geometry, attribute. When this box is present in a sample entry of tracks carrying G-PCC component bitstream, it indicates the type of G-PCC components carried by the respective track.

When the G-PCC bitstream stored in a single track, it shall not be present in the sample entry.

#### Syntax

aligned(8) class GPCCComponentTypeBox   
 extends FullBox('gtyp', version = 0, 0) {  
 unsigned int(8) gpcc\_type;  
}

#### Semantics

gpcc\_type identifies the type of G-PCC component as specified in Table 7.1.

Table 7.1 G-PCC Component Type

|  |  |
| --- | --- |
| gpcc\_type value | Description |
| 1 | Reserved |
| 2 | Geometry Data |
| 3 | Reserved |
| 4 | Attribute Data |
| 5..31 | Reserved. |

### 3D Spatial Region information structure

#### Definition

3DSpatialRegionStruct provides the 3D spatial region information including the anchor point, and the size of the 3D spatial region in the cartesian coordinate along to X, Y, Z axes relative to the anchor point.

#### Syntax

aligned(8) class 3DSpatialRegionStruct(anchor\_included, dimension\_included) {  
 unsigned int(16) 3d\_region\_id;  
 if(anchor\_included)

{

unsinged int(16) anchor\_x;

unsinged int(16) anchor\_y;

unsinged int(16) anchor\_z;

}  
if (dimension\_included)

{  
 unsinged int(16) region\_dx;

unsinged int(16) region\_dy;

unsinged int(16) region\_dz;  
}

}

#### Semantics

anchor\_included indicates whether the X,Y, Z coordinate values of the origin position of the 3D spatial region is included in the structure.

anchor\_x, anchor\_y, and anchor\_z indicate the x, y , and z offsets of the anchor point of the 3D spatial region, respectively, in cartesian coordinates. When it is absent in the structure, the anchor point shall be inferred to be equal to (0, 0, 0).

dimension\_included equal 1 indicates the dimension of the 3D spatial region is signalled in the structure. 3d\_dimension\_included equal 0 indicates the dimension of the 3D spatial region is not signalled in the structure.

region\_dx, region\_dy, and region\_dz indicate the size of the 3D spatial region in the Cartesian coordinates along the x, y, and z axes, respectively, relative to the origin position. It indicates the width, height, and depth of the 3D spatial region in the Cartesian coordinates.

## Encapsulation of G-PCC data in ISOBMFF

### General

When the G-PCC bitstream is carried in a single track, it requires the G-PCC encoded bitstream to be represented by a single-track declaration. Single-track encapsulation of G-PCC data could utilize the simple ISOBMFF encapsulation by storing the G-PCC bitstream in a single track without further processing.

Each sample in this track contains one or more G-PCC components. That is, each sample is compose of one or more TLV encapsulation structures. The Figure 7.1 depicts an example of the sample structure when the G-PCC geometry and attribute bitstream are stored in a single track.

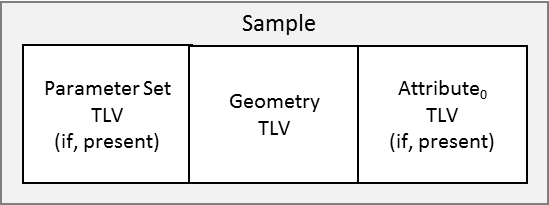


Figure 7.1 Sample structure when coded G-PCC bitstream is stored in a single track.

When the coded G-PCC geometry bitstream and the coded G-PCC attribute bitstream is stored in separate tracks, each sample in a track contains at least one TLV encapsulation structure carrying a single G-PCC component data, not both of geometry and attribute data. The general layout of this case is shown in Figure 7.2.

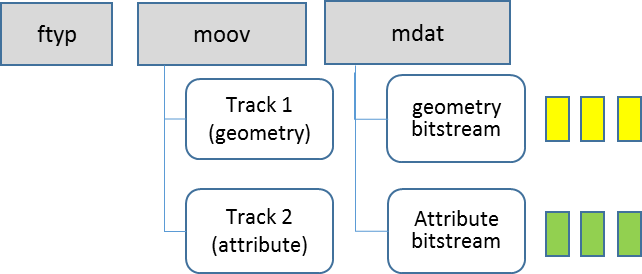


Figure 7.2 Multitrack container of G-PCC bitstream

Since the G-PCC geometry bitstream should be decoded first and decoding of G-PCC attribute bitstream depends on the decoded geometry, storing different G-PCC component bitstreams in separate tracks enable the player to access the track carrying the geometry bitstream before attribute bitstreams. An example of the sample structure in the track carrying only the coded G-PCC geometry bitstream is depicted in Figure 7. 3.

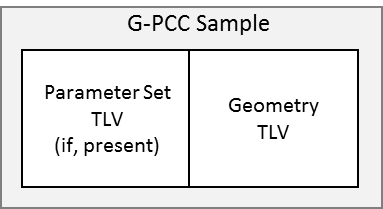


Figure 7.3 Sample structure of a track carrying G-PCC geometry bitstream only

The following conditions shall be satisfied:

1. When the G-PCC bitstream is carried in multiple tracks, a track carrying G-PCC geometry bitstream shall be an entry point.
2. In the sample entry, a new box is added which indicates the role of the stream contained in this track
3. A track reference is introduced from the track carrying only the G-PCC geometry bitstream to tracks carrying the G-PCC attribute bitstream.

### Sample entry

Sample Entry Type: 'gpe1', 'gpeg', 'gpc1'or 'gpcg'  
Container: SampleDescriptionBox  
Mandatory: A 'gpe1' , 'gpeg', 'gpc1'or 'gpcg' sample entry is mandatory  
Quantity: One or more sample entries may be present

G-PCC tracks shall use VolumetricVisualSampleEntry with a sample entry type of 'gpe1' , 'gpeg', 'gpc1'or 'gpcg'.

A G-PCC sample entry shall contain a GPCCConfigurationBox as defined in [7.2.2](#_G-PCC_decoder_configuration) and optionally a GPCCComponentTypeBox, as defined in 7.2.3..

Under the 'gpe1' sample entry, all SPS, GPS, APS, tile inventory (as defined in ISO/IEC 23090-9 [GPCC]) shall be in the setupUnit array. Under the 'gpeg' sample entry, the parameter sets may be present in this array, or in the stream. Under the ‘gpe1’or the 'gpeg' sample entry, GPCCComponentTypeBox shall not present.

Under the ‘gpc1’ sample entry, all SPS , GPS , and tile inventory (as defined in ISO/IEC 23090-9 [GPCC]) shall be in the setupUnit array of tracks carrying the G-PCC geometry bitstream. All associated APS shall be in the setupUnit array of tracks carrying the G-PCC attribute bitstream. Under the 'gpcg' sample entry, SPS, GPS, APS, or tile inventory may be present in this array, or in the stream.

Under the ‘gpc1’or the 'gpcg' sample entry, GPCCComponentTypeBox shall present.

NOTE: It is recommended that when several parameter sets are used and parameter set updating is desired, the parameter sets are included in the samples of the stream

#### Syntax

aligned(8) class GPCCSampleEntry()

extends VolumetricVisualSampleEntry (codingname) {  
 GPCCConfigurationBox config; //mandatory  
 GPCCComponentTypeBox type; // optional   
}

#### Semantics

compressorname in the base class VolumetricVisualSampleEntry indicates the name of the compressor used with the value "\013GPCC Coding" being recommended; the first byte is a count of the remaining bytes, here represented by \013, which (being octal 13) is 11 (decimal), the number of bytes in the rest of the string.

config includes G-PCC decoder configuration record information, as defined in [7.2.2](#_G-PCC_decoder_configuration).

type indicates the type of G-PCC component carried in the respective track, as defined in 7.2.3.

### Sample format

#### Definition

Each G-PCC sample corresponds to a single point cloud frame and shall be comprised of one or more TLV encapsulation structures which belong to the same presentation time. Each TLV encapsulation structure contains a single type of G-PCC payload, e.g., geometry data unit, attribute data unit, or parameter set.

#### Syntax

aligned(8) class GPCCSample  
{  
 unsigned int GPCCLength = sample\_size; //Size of Sample

for (i=0; i< GPCCLength; ) // to end of the sample  
 {  
 tlv\_encapsulation gpcc\_unit; //as defined in ISO/IEC 23090-9

i += (1+4)+ gpcc\_unit.tlv\_num\_payload\_bytes;  
 }  
}

#### Semantics

gpcc\_unit contain an instance of G-PCC TLV encapsulation structure containing a single G-PCC data unit. The syntax of tlv\_encapsulation is defined in Annex B of ISO/IEC 23090-9 [GPCC].

#### Sub-sample

For the use of SubSampleInformationBox in a G-PCC bitstream, a sub-sample is defined on the basis of the value of the flags field of the sub-sample information box as specified below. The flags specifies the type of sub-sample information given in this box as follows:

* 0: G-PCC TLV encapsulation structure based sub-samples. A sub-sample contains only one G-PCC TLV encapsulation structure as defined in ISO/IEC 23090-9 [GPCC].
* 1: Tile-based sub-samples. A sub-sample either contains one or more contiguous TLV encapsulation structure(s) that corresponds to one G-PCC tile or contains one or more contiguous TLV encapsulation structure(s) which contain either each parameter set, tile inventory or frame boundary marker.
* Other values of flags are reserved.

When the G-PCC geometry bitstream and the G-PCC attribute bitstream are carried in a same track, exactly one SubSampleInformationBox with flags equal to 0 or 1 in SampleTableBox, or in the TrackFragmentBox of each of its MovieFragmentBoxes shall be present.

If SubSampleInformationBox with flags equal to 0 is present, the 8-bit type value of TLV encapsulation structure and if the TLV encapsulation structure contains attribute data unit, the 6-bit value of attribute index, shall be included to the 32-bit codec\_specific\_parameters field of the sub-sample entry in the SubSampleInformationBox. The type of each sub-sample shall be identified by parsing the codec\_specific\_parameters field of the sub-sample entry in SubSampleInformationBox.

The codec\_specific\_parameters field of the SubsampleInformationBox is defined as below:

if (flags == 0) {

unsigned int(8) payloadType;

if (PayloadType == 4) { // attribute payload

unsigned int(6) attrIdx;

bit(18) reserved = 0;

}

else

bit(24) reserved = 0;

} else if (flags == 1) {

unsigned int(1) tile\_data;

bit(7) reserved = 0;

if (tile\_data)

unsigned int(24) tile\_id;

else

bit(24) reserved = 0;

}

PayloadType indicates the tlv\_type of the TLV encapsulation structure in the sub-sample.

AttrIdx indicates the ash\_attr\_sps\_attr\_idx of the TLV encapsulation structure containing attribute data unit in the sub-sample.

tile\_data indicates whether the sub-sample contains one tile or others. tile\_data equal to 1 indicates that the sub-sample contains TLV encapsulation structure(s) which contains either geometry data unit or attribute data unit that corresponds to one G-PCC tile. tile\_data equal to 0 indicates that the sub-sample contains TLV encapsulation structure(s) which contains either each parameter set, tile inventory or frame boundary marker.

tile\_id indicates the index of the G-PCC tile which the sub-sample is associated with within the tile inventory.

### Referencing G-PCC tracks

To link between tracks when G-PCC geometry bitstream and attribute bitstream are carried in separate tracks, the track reference tool of ISO/IEC 14496-12 [ISOBMFF] shall be used. One TrackReferenceTypeBoxes shall be added to a TrackReferenceBox within the TrackBox of the G-PCC geometry track. The TrackReferenceTypeBox shall contain an array of track\_IDs designating the tracks which the G-PCC track references. To link the G-PCC geometry track to the G-PCC attribute track, reference\_type of a TrackReferenceTypeBox in the G-PCC geometry track identifies the associated attribute tracks. The 4CCs of these track reference types shall be:

* 'gpca': the referenced track(s) contain the coded bitstream of G-PCC attribute data

## Partial Access of G-PCC data

### Static 3D Spatial Region Information.

#### Definition

GPCCSpatialRegionInfoBox provides the associated 3D spatial region information including the X, Y, Z coordinate values of the anchor point and the size of the spatial region along to X, Y, Z axes relative to the anchor point.

When GPCCSpatialRegionInfoBox presents in the sample entry of track carrying both G-PCC geometry bitstream and G-PCC attribute bitstream, it indicates the static 3D spatial region information of point cloud data carried in track. [Ed (SOH): need further study how this box is present in the sample entry when G-PCC geometry bitstream and G-PCC attribute bitstream are carried in separate tracks.]

#### Syntax

aligned(8) class GPCCSpatialRegionInfoBox extends FullBox('gpsr',0,0){

unsigned int(15) num\_regions;  
 unsigned int(1) tile\_id\_present;

for (int i=0; i < num\_regions; i++) {

3DSpatialRegionStruct(1,1);  
 if(tile\_id\_present){  
 unsigned int(8) num\_tiles[i];  
 for(int j=0; j < num\_tiles; j++)  
 unsigned int(16) tile\_id[j];  
 }

}

}

#### Semantics

num\_regions indicates the number of 3D spatial regions of the point cloud.

tile\_id\_present equal to 0 that the identifier of tiles associated with this spatial region is not present. tile\_id\_present equal to 1 that the identifier of tiles associated with this spatial region presents.

3DSpatialRegionStruct provides the 3D spatial region information of the part or all point cloud data indicated by the anchor point and the size of the spatial region along to X, Y, Z axes relative to the anchor point.

num\_tiles[i] indicates the number of G-PCC tiles associated with the i-th 3D spatial region

tile\_id [j] identifies the j-th G-PCC tile associated with the i-th 3D spatial region.

### Dynamic Spatial Region Information.

#### General

This metadata track indicates the dynamically changed 3D spatial region information and the association between the regions and associated G-PCC tiles over time. It shall contain a 'cdsc' track reference to the track carrying partial or all G-PCC geometry bitstream.

If a track carrying partial or all G-PCC geometry bitstream has an associated timed-metadata track with a sample entry type 'gpdr', the associated 3D spatial region information of the point cloud data carried by the track is considered as dynamic.

#### Sample entry

aligned(8) class Dynamic3DSpatialRegionSampleEntry

extends MetaDataSampleEntry('gpdr’)

{

GPCCSpatialRegionInfoBox();  
 bit(6) reserved=0;  
 unsigned int(1) dynamic\_dimension\_flag;  
 unsigned int(1) dynamic\_tile\_id\_flag;  
}

GPCCSpatialRegionInfoBox indicates one or more initial 3D spatial region information.

num\_regions indicates the number of 3D spatial regions signalled in the sample entry.

tile\_id\_present equal to 0 indicates that the identifier of tiles associated with 3D spatial regions is not present in the sample entry. tile\_id\_present equal to 1 indicates that the identifier of tiles associated with 3D spatial regions presents in the sample entry.

3DSpatialRegionStruct provides the initial information of the i-th 3D spatial region including the anchor point and the size along to X, Y, Z axes relative to the anchor point in Cartesian coordinate in the sample entry.

num\_tiles[i] indicates the number of G-PCC tiles associated with the i-th 3D spatial region in the sample entry

tile\_id [j] identifies the j-th G-PCC tile associated with the i-th 3D spatial region in the sample entry.

dynamic\_dimension\_flag equal to 0 specifies that the dimension of the 3D spatial region remains unchanged in all samples referring to this sample entry. dynamic\_dimension\_flag equal to 1 specifies that the dimension of the 3D spatial region is indicated in each sample.

dynamic\_tile\_id\_flag equal to 0 specifies that identifier of tiles associated with the 3D spatial region remains unchanged in all samples referring to this sample entry. dynamic\_tile\_id\_flag equal to 1 specifies that identifier of tiles associated with the 3D spatial region presents in each sample.

#### Sample format

The sample syntax of this sample entry type 'gpdr’ is specified as follows:

aligned(8) Dynamic3DSpatialRegionSample() {  
 unsigned int(16) num\_regions;  
 for (int i=0; i < num\_regions; i++) {

3DSpatialRegionStruct(1,dynamic\_dimension\_flag[i]);  
if(dynamic\_tile\_id\_flag){  
 unsigned int(8) num\_tiles[i];  
 for(int j=0; j < num\_tiles; j++)  
 unsigned int(16) tile\_id[j];  
}

}

}

num\_regions indicates the number of 3D spatial regions signalled in the sample.

3DSpatialRegionStruct provides the 3D spatial region information of G-PCC data when this sample is applied

num\_tiles[i] indicates the number of G-PCC tiles associated with the i-th 3D spatial region when this sample is applied

tile\_id[j] identifies a particular G-PCC tile associated with the i-th 3D spatial region when this sample is applied.

Let the target be the point cloud data associated with samples in the referenced tracks with composition times greater than or equal to the composition time of this sample and less than the composition time of the next sample.

[Ed (RT): How to associate partial point clouds that contribute to the same entire point cloud needs further study.]

# Encapsulation and signalling in DASH

## General

Each G-PCC component shall be represented in the DASH manifest (MPD) file as a separate Adaptation Set, referred to as a Component Adaptation Set. The Adaptation Set containing the geometry information is also the Main Adaptation Set which serves as the access point for the G-PCC content. The Main Adaptation Set shall contain either a single Initialization Segment at adaptation set level or multiple Initialization Segments at representation level (one for each Representation). Initialization Segments shall contain the G-PCC parameter sets as defined in [1], which are needed to initialize the G-PCC decoder.

## DASH MPD descriptors for G-PCC media

A number of XML elements and attributes are defined. These XML elements are defined in a separate namespace "urn:mpeg:mpegI:gpcc:2020". The namespace designator "gpcc:" is used to refer to this name space.

### GPCCComponent Descriptor

To identify the type of the G-PCC component in a Component Adaptation Set, an **EssentialProperty** descriptor can be used with the @schemeIdUri attribute equal to "urn:mpeg:mpegI:gpcc:2020:component".

This descriptor is referred to as a **GPCCComponent** descriptor. At adaptation set level, one **GPCCComponent** descriptor shall be signalled for each point cloud component that is present in the Representations of the adaptation set. The **GPCCComponent** descriptor shall include the attributes defined in Table 8.1.

Table 8.1 Elements and attributes for the GPCCComponent descriptor

|  |  |  |  |
| --- | --- | --- | --- |
| **Elements and Attributes for GPCCComponent descriptor** | **Use** | **Data type** | **Description** |
| **component** | 0..N | gpcc:gpccComponentType | An element whose attributes specify information for one of the Geometry point cloud components present in the representation(s) of the adaptation set. |
| **component**@component\_type | M | xs:string | Indicates the type of the point cloud component. Value ‘geom’ indicates a geometry component, and ‘attr’ indicates an attribute component. |
| **component**@attribute\_type | CM | xs:unsignedByte | Indicates the type of the attribute (refer to Table 8 in the G-PCC DIS 23090-9 [1]). Only values between 0 and 255, inclusive, are allowed. Shall be present only if the component is a geometry point cloud attribute (i.e., component\_type has the value ‘attr’).  Attribute type & value  0: Colour 1: Reflectance 2: Frame index 3: Material ID 4: Transparency 5: Normals 6 .. 255: Reserved |
| Legend:  For attributes: M=Mandatory, O=Optional, OD=Optional with Default Value, CM=Conditionally Mandatory.  For elements: <minOccurs>..<maxOccurs> (N=unbounded)  Elements are bold; attributes are non-bold and preceded with an @. | | | |

When more than one Representation is signaled in the Main Adaptation Set, the initialization segment of each Representation shall contain the G-PCC parameter sets for that Representation. The Representation(s) of a Component Adaptation Set shall list the corresponding Representation identifier from the Main Adaptation Set using the @dependencyId attribute defined in ISO/IEC 23009-1 [DASH].

### GPCC Descriptor

A streaming client can identify the type of point cloud component in an Adaptation Set or Representation by checking the **GPCCComponent** descriptor within the corresponding element. However, a streaming client also needs to distinguish between different geometry point cloud streams present in the MPD file.

A **SupplementalProperty** element with a @schemeIdUri attribute equal to "urn:mpeg:mpegI:gpcc:2020:gpc" is referred to as a GPCC descriptor. The attributes of **GPCC** descriptor is shown in Table 8.2.

At most one **GPCC** descriptor may be present at the adaptation set level for the Main Adaptation Set of the G-PCC media or at the preselection level.

Table 8.2 Attributes for the GPCC descriptor

|  |  |  |  |
| --- | --- | --- | --- |
| **Attributes for GPCC descriptor** | **Use** | **Data type** | **Description** |
| gpcc:@gpcId | CM | xs:string | An id for the volumetric media. This attribute shall be present if multiple versions of the same volumetric media are signalled in separate AdapatationSets. |
| Legend:  For attributes: M=Mandatory, O=Optional, OD=Optional with Default Value, CM=Conditionally Mandatory.  For elements: <minOccurs>..<maxOccurs> (N=unbounded)  Elements are **bold**; attributes are non-bold and preceded with an @. | | | |

## GPCC Preselection

A G-PCC Preselection is signalled in the MPD using a **Preselection** element (as defined in ISO/IEC 23009-1 [DASH]) with an id list for the @preselectionComponents attribute including the id of the Main Adaptation Set for the G-PCC media followed by the ids of other Component Adaptation Sets. The @codecs attribute for the Preselection shall be set to ‘gpc1’, indicating that the Preselection media is a geometry-based point cloud. The PreSelection may either be signalled using a **Preselection** element within the **Period** element or a Preselection descriptor at the adaptation set level in the Main Adaptation Set.

## Supporting multiple versions of GPCC media

Multiple versions of the same point cloud media shall be signalled using separate Preselections. Preselections that represent alternative versions of the same geometry-based point cloud media shall contain a GPCC descriptor with the same @gpcId value. At most one GPCC descriptor shall be present at the preselection level. These Preselections are alternatives to each other. The id of the Main Adaptation Set is the first id in the list of Adaptation Set ids for the Preselection followed by the ids of Component Adaptation Sets. Figure 2 illustrates an exemplary DASH configuration for grouping multiple versions of G-PCC components belonging to the same point cloud within an MPEG-DASH MPD file.

1. (normative)  
     
   File format toolsets and brands
2. (normative)  
     
   GPCC DASH Schema
3. (informative)  
     
   DASH MPD examples

## C.1. Grouping of G-PCC Components

The below Figure C.1 illustrates an exemplary DASH configuration for grouping of G-PCC components belonging to a single point cloud media within an MPEG-DASH MPD file.



**Figure C.1 Grouping GPCC components in an MPD using Preselection**

## C.2. Supporting multiple versions of GPCC media

In this example, two versions of the G-PCC components bitstreams are available. For each version of G-PCC component (geometry and attributes), two different bitstreams are available (encoded at different qualities). Each version of the geometry and attribute components can be signalled by separate Adaptation Sets, with two Representations one for each quality level in the Adaptation Set.

The Main Adaptation Set shall contain the Initialization Segment for the complete experience (G-PCC parameter sets, etc.). Compatible Component Adaptation Sets, along with the respective Main Adaptation Set, are grouped together in separate Preselections in the MPD. To indicate that these Adaptation Sets are referenced in at least one Preselection, a Preselection descriptor without the @value attribute is signaled in each Adaptation Set. Each Preselection shall include a GPCC descriptor that indicates at least the mandatory @gpcId attribute. The values assigned to the @gpcId attributes of the two Preselections are identical, indicating that both Preselections belong to the same point cloud content. When the Preselections represent different point clouds, the @gpcId attribute for each Preselection shall be different.



**Figure C.2. Grouping of multiple versions of GPCC components in an MPD using PreSelections**

A skeleton for the complete MPD file that signals the point cloud content described in the above example is given below.

|  |
| --- |
| <?xml version="1.0" encoding="UTF-8"?>  <MPD  <Period>   <!-- Main GPCC AdaptationSet -->   <AdaptationSet id="1" codecs="gpc1">   <EssentialProperty schemeIdUri="urn:mpeg:dash:preselection:2016" />   <EssentialProperty schemeIdUri="urn:mpeg:mpegI:gpcc:2020:component">   <gpcc:component component\_type="geom" />   </EssentialProperty>  <Representation>   ...  </Representation>   </AdaptationSet>  <AdaptationSet id="2" codecs="gpc1">   <EssentialProperty schemeIdUri="urn:mpeg:dash:preselection:2016" />  <EssentialProperty schemeIdUri="urn:mpeg:mpegI:gpcc:2020:component">   <gpcc:component component\_type="geom" />   </EssentialProperty>  <Representation>   ...  </Representation>   </AdaptationSet>  <!-- Attribute -->  <AdaptationSet id="3" codecs="gpc1">   <EssentialProperty schemeIdUri="urn:mpeg:dash:preselection:2016" />   <EssentialProperty schemeIdUri="urn:mpeg:mpegI:gpcc:2020:component">   <gpcc:component component\_type="attr" attribute\_type="0" />   </EssentialProperty>   <Representation>   ...   </Representation>   </AdaptationSet>    <AdaptationSet id="4" codecs="gpc1">   <EssentialProperty schemeIdUri="urn:mpeg:dash:preselection:2016" />   <EssentialProperty schemeIdUri="urn:mpeg:mpegI:gpcc:2020:component">   <gpcc:component component\_type="attr" attribute\_type="0" />   </EssentialProperty>   <Representation>   ...   </Representation>   </AdaptationSet>    <AdaptationSet id="5" codecs="gpc1">   <EssentialProperty schemeIdUri="urn:mpeg:dash:preselection:2016" />   <EssentialProperty schemeIdUri="urn:mpeg:mpegI:gpcc:2020:component">   <gpcc:component component\_type="attr" attribute\_type="1"/>   </EssentialProperty>   <Representation>   ...   </Representation>   </AdaptationSet>    <AdaptationSet id="6" codecs="gpc1">   <EssentialProperty schemeIdUri="urn:mpeg:dash:preselection:2016" />   <EssentialProperty schemeIdUri="urn:mpeg:mpegI:gpcc:2020:component">   <gpcc:component component\_type="attr" attribute\_type="1"/>   </EssentialProperty>   <Representation>   ...   </Representation>   </AdaptationSet>    <AdaptationSet id="7" codecs="gpc1">   <EssentialProperty schemeIdUri="urn:mpeg:dash:preselection:2016" />   <EssentialProperty schemeIdUri="urn:mpeg:mpegI:gpcc:2020:component">   <gpcc:component component\_type="attr" attribute\_type="4"/>   </EssentialProperty>   <Representation>   ...   </Representation>   </AdaptationSet>    <AdaptationSet id="8" codecs="gpc1">   <EssentialProperty schemeIdUri="urn:mpeg:dash:preselection:2016" />   <EssentialProperty schemeIdUri="urn:mpeg:mpegI:gpcc:2020:component">   <gpcc:component component\_type="attr" attribute\_type="4"/> />   </EssentialProperty>   <Representation>   ...   </Representation>   </AdaptationSet>    <!-- Preselections -->   <Preselection id="1" tag="1" preselectionComponents="1 3 5 7" codecs="gpc1">   <!—GPCC Descriptor -->   <EssentialProperty schemeIdUri="urn:mpeg:mpegI:gpcc:2020:gpc" gpcId="1" />   </Preselection>    <Preselection id="2" tag="2" preselectionComponents="2 4 6 8" codecs="gpc1">  <!—GPCC Descriptor -->   <EssentialProperty schemeIdUri="urn:mpeg:mpegI:gpcc:2020:gpc" gpcId="1" />   </Preselection>   </Period>  </MPD> |