**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC1/SC29/WG11**

**CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC1/SC29/WG11 m52971**

**April 2020, Electronic Meeting**

|  |  |
| --- | --- |
| **Source** | **InterDigital Communications, Inc.** |
| **Status** | **Input** |
| **Title** | **[PCC Systems] A Volumetric Metadata Track for Partial Access in V-PCC** |
| **Author** | Ahmed Hamza |
| **Agenda** | 33.1 |

# Introduction

A Core Experiment on partial access of point cloud data was established in the MPEG#127 meeting ‎[1]. The draft text of ISO/IEC DIS 23090-10 ‎[2] includes a clause on partial access support based on the responses to the CE that were adopted in the MPEG#128 and MPEG#129 meetings.

In the MPEG#129 meeting, a number of new SEI messages ‎[3]to annotate, label, and add properties to patches and volumetric rectangles were adopted in ISO/IEC 23090-5. These SEI messages replace the previously defined volumetric tiling information SEI message and provides extended functionality. Multiple patches in the atlas frame or 2D volumetric rectangles, which can contain themselves multiple patches, could be associated with a single object within the point cloud scene and associate a number of properties to that object.

While the newly adopted SEI messages enable partial access and rendering of a V-PCC sequence, the latest draft ISO/IEC DIS 23090-10 only supports spatial division of the 3D space within the bounding box of the point cloud and does not support the notion of objects. Therefore, in applications where object-level partial access is desirable, the client is unable to easily identify where an object is located in 3D space and, consequently, which set of tracks need to be accessed/retrieved to decode and render that particular object. Even if object bounding box information are signaled in SEI messages within the bitstream, it should not be expected that a player or streaming client must decode these SEI messages, to map the bounding box to spatial regions, as these messages are intended for the V-PCC decoder.

In this contribution, we propose a solution where a minimal set of object-related information needed to support object-level partial access is signaled in a timed-metadata track which maps those objects to previously defined spatial regions or directly to the tracks carrying patches associated with the objects.

# Proposed Solution

It is proposed to extend the 'dysr' timed-metadata track for spatial region information and have a more generic timed-matadata track that carries volumetric information including both spatial regions and different objects within a point cloud scene. We refer to this track as a volumetric metadata or dynamic volumetric metadata track. While we propose a new timed-metadata track that replaces the dynamic spatial region information in clause 8.6 of ISO/IEC 23090-10, this is mainly to provide a more meaningful name that takes into consideration a more general role for the track. The main difference between the proposed track and the dynamic spatial regions timed-metadata track is the additional syntax in the sample format.

Note that object annotation SEI messages are carried in the V-PCC track (or any track carrying samples for corresponding atlas sub-bitstream).

## Volumetric Metadata Track

A volumetric metadata track is a timed-metadata track that carries information about the different objects within the point cloud scene as well as any 3D spatial partitioning. Object information are carried in the samples of the track. This timed-metadata track is a track with a DynamicVolumetricMetadataSampleEntry) and is defined as follows, where VPCCSpatialRegionsBox is as defined in ‎[2]

aligned(8) class DynamicVolumetricMetadataSampleEntry extends MetaDataSampleEntry('dyvm') {  
 VPCCSpatialRegionsBox();  
}

The volumetric metadata track shall contain a 'cdsc' track reference to the V-PCC track.

## Sample Format for Volumetric Metadata Track

In addition to spatial region information, the samples of the volumetric metadata track include a table that maps object ids to one or more track groups carrying the V-PCC tile groups to which the corresponding object is mapped. The samples may also include a dependency list for each signaled object which includes the ids of other objects that the signaled object depends on.

Samples of the volumetric metadata track are defined as follows, where 3DSpatialRegionStruct is as defined in ISO/IEC CD 23090-10.

aligned(8) class VPCCVoumetricMetadataSample() {  
 bit(1) region\_updates\_flag;  
 bit(1) object\_updates\_flag;  
 bit(6) reserved = 0;  
 if (region\_updates\_flag) {  
 unsigned int(16) num\_regions;  
 for (i=0; i < num\_regions; i++) {  
 3DSpatialRegionStruct(dimensions\_included\_flag);  
 }  
 }  
 if (object\_updates\_flag) {  
 unsigned int(8) num\_obj\_updates;  
 for (i=0; i<num\_obj\_updates; i++) {  
 unsigned int(8) obj\_index\_length[i];  
 unsigned int(obj\_index\_length[i] \* 8) object\_id[i];  
 bit(1) obj\_cancel\_flag[i];  
 if (obj\_cancel\_flag[i]) {  
 bit(7) reserved = 0;  
 } else {  
 bit(4) reserved = 0;  
 bit(1) obj\_spatial\_region\_mapping\_flag[i];  
 bit(1) obj\_dependencies\_present\_flag[i];  
 bit(1) obj\_bounding\_box\_present\_flag[i];  
 if (obj\_bounding\_box\_present\_flag[i]) {  
 3DSpatialRegionStruct(1); // not needed if we are mapping each object to 3D region  
 }  
 if (obj\_spatial\_region\_mapping\_flag[i]) {  
 // mapping object to spatial regions  
 unsigned int(8) num\_spatial\_regions[i];  
 for (j=0; j < num\_spatial\_regions[i]; j++) {  
 region\_id[j][i];  
 }  
 } else {  
 // mapping object to tile group tracks  
 unsigned int(8) num\_track\_groups[i];  
 for (j=0; j < num\_track\_groups[i]; j++) {  
 unsigned int(32) track\_group\_id[j][i];  
 }  
 }  
 if (obj\_dependencies\_present\_flag[i]) {  
 unsigned int(8) num\_obj\_dependencies[i];  
 for (j=0; j < num\_obj\_depedendencies[i]; j++) {  
 unsigned int(8) obj\_dep\_index\_length[j][i];  
 unsigned int(obj\_dep\_index\_length[j][i] \* 8) obj\_index[j][i];  
 }  
 }  
 }  
 }  
 }  
}

The semantics of the fields in VPCCVolumetricMetadataSample are:

region\_updates\_flag indicates whether the sample includes updates to 3D spatial regions or not.

object\_updates\_flag indicates whether the sample includes updates to point cloud scene objects or not.

num\_obj\_updates number of point cloud scene objects updated in the sample.

obj\_index\_length[i] is the length of the object index, in number of bytes, for the i-th object in the object update list of the sample.

object\_index[i] is the index of the i-th object in the object update list of the sample.

obj\_cancel\_flag[i] indicates whether the i-th object in the object update list of the sample is cancelled.

obj\_spatial\_region\_mapping\_flag[i] indicates whether a mapping to a spatial region is signaled for the i-th object in the object update list of the sample.

obj\_depdendencies\_present\_flag[i] indicates whether object depdendency information is available for the i-th object in the object update list of the sample or not. Value 0 indicates that the object does not depend on other objects. Value 1 indicates that the object depends on one or more objects within the point cloud scene.

obj\_bounding\_box\_present\_flag[i] indicates whether 3D bounding boxing information is available for the i-th object in the object update list of the sample or not. Value 0 indicates that no bounding box information is given. Value 1 indicates that that 3D bounding box information for the i-th object are signaled in the sample.

num\_spatial\_regions[i] is the number of 3D spatial regions that the i-th object in the object update list of the sample is associated with.

region\_id[j][i] is the id of the j-th spatial region that the i-th object in the object update list of the sample is associated with.

num\_track\_groups[i] is the number of track groups that the i-th object in the object update list of the sample is associated with.

track\_group\_id[j][i] is the id of the j-th track group that the i-th object in the object update list of the sample is associated with.

num\_obj\_depedencies[i] is the number of objects that the i-th object in the object update list of the sample depdends on.

obj\_dep\_index\_length[j][i] is the length, in number of bytes, of the index of the j-th object that the i-th object in the object update list of the sample depends on.

obj\_index[j][i] is the index of the j-th object that the i-th object in the object update list of the sample depends on.

# References

1. N18663, “Description of Core Experiment on Partial Access of PC Data”, MPEG 127, Gothenburg, Sweden, July 2019
2. N19066, “Draft Text of ISO/IEC DIS 23090-10 Carriage of Video-based Point Cloud Compression Data”, MPEG 129, Brussels, Belgium, January 2020
3. m52705, “Object Annotation of Patches and Volumetric Rectangles in V-PCC”, MPEG 129, Brussels, Belgium, January 2020