ISO/IEC JTC 1/SC 29/WG 03 N0417

**ISO/IEC JTC 1/SC 29/WG 03  
MPEG Systems   
Convenorship: KATS (Korea, Republic of)**

**Document type:** Output Document

**Title:** WD of ISO/IEC 23090-7 AMD 1 Common Metadata for Immersive Media

**Status:** Approved

**Date of document:** 2021-10-15

**Source:** ISO/IEC JTC 1/SC 29/WG 03

**No. of pages:** 1 (with cover page)

**Email of Convenor:** young.L @ samsung . com

**Committee URL:** <https://isotc.iso.org/livelink/livelink/open/jtc1sc29wg3>

**INTERNATIONAL ORGANIZATION FOR STANDARDIZATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC 1/SC 29/WG 03 MPEG SYSTEMS**

**ISO/IEC JTC 1/SC 29/WG 03 N0417**

**October 2021, Virtual**

|  |  |
| --- | --- |
| **Title** | **WD of ISO/IEC 23090-7 AMD 1 Common Metadata for Immersive Media** |
| **Source** | **WG 03, MPEG Systems** |
| **Status** | **Approved** |
| **Serial Number** | **20894** |

Contents

[**1** **Introduction** 2](#_Toc85208088)

[**2** **Scope:** 2](#_Toc85208089)

[**3** **Guiding principles** 2](#_Toc85208090)

[**4** **Usage of Metadata in ISOBMFF** 2](#_Toc85208091)

[**5** **Metadata of individual features** 2](#_Toc85208092)

[**5.1** **Basic** 2](#_Toc85208093)

[**5.1.1** **Coordinate Systems** 2](#_Toc85208094)

[**5.1.2** **Dimensions, Positions (or Offsets, Translations, Locations), Sizes (or Ranges)** 2](#_Toc85208095)

[**5.1.3** **Rotations and orientations** 2](#_Toc85208096)

[**5.1.4** **Scaling** 3](#_Toc85208097)

[**5.2** **Viewing Spaces** 3](#_Toc85208098)

[**5.3** **Regions** 3](#_Toc85208099)

[**5.4** **Viewpoints** 3](#_Toc85208100)

[**5.5** **Viewports** 3](#_Toc85208101)

[**6** **Metadata of spatially related features** 3](#_Toc85208102)

[**6.1** **Localized coordinate systems (e.g., one is in another “global” coordinate system)** 3](#_Toc85208103)

[**6.2** **Localized viewing spaces (e.g., one is in another “larger” viewing space)** 4](#_Toc85208104)

[**6.3** **Sub-regions (e.g., one or more are parts of another “source” region)** 4](#_Toc85208105)

[**6.4** **Objects in Regions (e.g., one or more objects are in a “containing” region)** 4](#_Toc85208106)

[**6.5** **Overlay of Regions/Objects (e.g., one is in front a “background” region)** 4](#_Toc85208107)

[**7** **Dynamic metadata tracks** 4](#_Toc85208108)

[**8** **DASH descriptors** 4](#_Toc85208109)

1. **Introduction**

Many parts of ISO/IEC 23090 have developed metadata definition. Some of them are quite similar each other. For harmonization among the parts and reuse of already defined metadata for the future development, this amendment intends collect the metadata defined by the parts of ISO/IEC 23090, e.g. 23090-10, 23090-18 and so on, and study harmonization among them. The scope of the new amendment doesn’t expand the scope of the original project.

1. **Scope:**

* metadata collected from and applicable to all MPEG-I parts
* 23090: 5, 7, 9, 10, 12, 14, 18
* 23001: 10

1. **Guiding principles**
2. **Usage of Metadata in ISOBMFF**

* Static: extension of containing boxes
* Dynamic: timed metadata tracks

1. **Metadata of individual features**
   1. **Basic** 
      1. **Coordinate Systems**

* Coordinate system – Cartesian coordinate system
* Unit sphere coordinate system (OMAF specific)
* Object coordinate space – referring to object space, where manipulation is done relative to object origin
* World coordinate space – referring to scene space, where manipulation is done relative to scene origin
* Provide example of how to move between different spaces
  + 1. **Dimensions, Positions (or Offsets, Translations, Locations), Sizes (or Ranges)**
* Define syntax structures Vector3Uint, Vector3Int, Vector3Float
* Define translation processes whenever required
* Use them consistently when 3d positions, offsets, dimensions, translations or scaling is handled
  + 1. **Rotations and orientations**
* Define syntax structures and processes for rotation and orientation
* Make sure that other syntax structures use orientation and rotation correctly and efficiently
  + 1. **Scaling**
* TBD: How does differ from scaling defined in 4.1.2
  1. **Viewing Spaces**
* Dimensions: 3D and 2D
  + TBD: What is a 2D viewing space?
* Shapes: Cuboid (Bounding Box), Sphere, Cylinder, Ellipsoid, etc.
  1. **Regions**
* Dimensions: 2D and 3D

aligned(8) class BoundingVolume (

unsigned int(1) anchor\_included,

unsigned int(8) precision)

{

if (anchor\_included) { // anchor is not 0,0,0

Vector3 anchor(precision);

}

Vector3 dimension(precision);

}

* Shapes
* Sizes
  + TBD: Depending on the definition isn’t the size of the region or shape included in the definition itself?
* IDs
  + TBD: Not sure if ID can be considered generic enough. People can create other standard specific structures that use the common metadata here.
  1. **Viewpoints**
* Extrinsics
  + Orientations
  + Positions
  + IDs
* Intrinsics
  + Type
  + Type dependent parameters
* Objects/Components
  + TBD: we need to discuss, if this is generally used by other specifications.
  + Shape
  + Sizes
  + Attributes
  + IDs
  1. **Viewports**

1. **Metadata of spatially related features**
   1. **Localized coordinate systems (e.g., one is in another “global” coordinate system)**

* References to global coordinate systems
* Positions
* Rotations
* Scaling
  1. **Localized viewing spaces (e.g., one is in another “larger” viewing space)**
* References to global viewing spaces
* Positions
* Rotations
* Scaling
  1. **Sub-regions (e.g., one or more are parts of another “source” region)**
* Positions
* Rotations
* Scaling
  1. **Objects in Regions (e.g., one or more objects are in a “containing” region)**
* Positions
* Rotations
* Scaling
  1. **Overlay of Regions/Objects (e.g., one is in front a “background” region)**
* Positions
* Rotations
* Scaling
* Alpha blending

1. **Dynamic metadata tracks**

* TBD: needs to discuss, if we need to define sample and track design for some common dynamic metadata scenarios like viewport tracks. One possibility is to define them in MPEG-B part 10 (Carriage of Timed Metadata in ISOBMFF)

1. **DASH descriptors**

* TBD