ISO/IEC JTC 1/SC 29/WG 03 N1224

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

**Document type:** Output Document

**Title:** Technologies under Consideration for file format reference software and conformance

**Status:** Approved

**Date of document:** 2024-05-10

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

**No. of pages:** 11 (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 N1224**

**April 2024 – Rennes, FR**

|  |  |
| --- | --- |
| **Title** | **Technologies under Consideration for file format reference software and conformance** |
| **Source** | **WG 03, MPEG Systems** |
| **Status** | **Approved** |
| **Serial Number** | **23834** |

#### HEVC\_1440x960\_derived\_image\_SSTR.heic

HEIF Single Stream Item Property – A file with SingleStreamProperty associated with overlay derived image item.

#### HEVC\_8192x4096\_25fps\_NALM\_PPRR\_PIPM.mp4

Picture-in-Picture - A file with PicInPicRegionReplacementEntry when region\_id\_type is equal to 1.

#### HEIF extrinsic intrinsic properties

— ./**ex\_in\_trinsics/intrinsics/intrinsics.heic** A file with intrinsics properties with the following contents:

Item 1:

fx = 512

fy = 512

s = 0

cx = 256

cy = 96

Coded as:

flags: 6656

Values: [33554432, 16777216, 8388608]

Item 2:

fx = 512

fy = 501.5

s = -100.5

cx = 256

cy = 96

Coded as:

flags: 72193

Values: [33554432, 16777216, 8388608, 43821739, -201]

— ./**ex\_in\_trinsics/extrinsic\_position/position.heic** The file contains 6 image items with 5 different extrinsic positions:

Item 1:

Coordinate System ID: 0

Position: 0.0, 0.0, 0.0

Coded as flags: 0; No values

Item 2:

Coordinate System ID: 1

Position: 0.01, 0.0, 0.0

Coded as flags: 33; Values: [10000, 1]

Item 3:

Coordinate System ID: 0

Position: 0.0, 0.02, 0.0

Coded as flags: 2; Values: [20000]

Item 4:

Coordinate System ID: 1

Position: 0.0, 0.0, -0.04

Coded as flags: 36; Values: [-40000, 1]

Item 5:

Coordinate System ID: 1

Position: -0.03, 0.05, -0.08

Coded as flags: 39; Values: [-30000, 50000, -80000, 1]

Item 6:

Coordinate System ID: 0

Position: 0.0, 0.0, 0.0

Coded as flags: 0; No values

Coordinate System ID: 1

Position: -0.03, 0.05, -0.08

Coded as flags: 36; Values: [-30000, 50000, -80000, 1]

— **./ex\_in\_trinsics/extrinsic\_rotation/ rotation\_[00,...,15].heic** The files with extrinsics properties with rotation signaling. All files have extrinsics position set to 0.1, 0.1, -0.1 (in meters).

rotation\_00.heic:

Quaternion: 0.987965 0.0940609 -0.0789265

Input angles: X:170, Y:10, Z:10

Corresponding input matrix:

0.9698, 0.2007, -0.1383,

0.1710, -0.9646, -0.2007,

-0.1736, 0.1710, -0.9698,

rotation\_01.heic:

Quaternion: 0.363338 0.665546 -0.522402

Input angles: X:-110, Y:64, Z:170

Corresponding input matrix:

-0.4317, 0.8912, 0.1396,

0.0761, 0.1902, -0.9788,

-0.8988, -0.4119, -0.1499,

rotation\_02.heic:

Quaternion: -0.0469889 0.386794 0.913955

Input angles: X:45, Y:10, Z:170

Corresponding input matrix:

-0.9698, -0.2437, 0.0019,

0.1710, -0.6750, 0.7177,

-0.1736, 0.6964, 0.6964,

rotation\_03.heic:

Quaternion: 0.0789265 0.0940609 0.0789265

Input angles: X:10, Y:10, Z:10

Corresponding input matrix:

0.9698, -0.1413, 0.1986,

0.1710, 0.9751, -0.1413,

-0.1736, 0.1710, 0.9698,

rotation\_04.heic: // This creates a quaternion with w < 0, which should trigger a sign change

Quaternion: 0.242726 0.241691 -0.789943

Input angles: X:170, Y:141, Z:62

Corresponding input matrix:

-0.3648, 0.9208, -0.1376,

-0.6862, -0.3659, -0.6287,

-0.6293, -0.1349, 0.7653,

rotation\_05.heic: // This is the worst case wrt precision when running the precision tests

Quaternion: -0.99981 0.00872521 0.0174517

Input angles: X:0, Y:178, Z:179

Corresponding input matrix:

0.9992, -0.0175, -0.0349,

-0.0174, -0.9998, 0.0006,

-0.0349, -0.0000, -0.9994,

rotation\_06.heic: // Testing setting various angles to zero

Quaternion: 0.707107 0 0

Input angles: X:90, Y:0, Z:0

Corresponding input matrix:

1.0000, 0.0000, 0.0000,

0.0000, 0.0000, -1.0000,

-0.0000, 1.0000, 0.0000,

rotation\_07.heic:

Quaternion: 0 0.707107 0

Input angles: X:0, Y:90, Z:0

Corresponding input matrix:

0.0000, 0.0000, 1.0000,

0.0000, 1.0000, 0.0000,

-1.0000, 0.0000, 0.0000,

rotation\_08.heic:

Quaternion: 0 0 0.707107

Input angles: X:0, Y:0, Z:90

Corresponding input matrix:

0.0000, -1.0000, 0.0000,

1.0000, 0.0000, 0.0000,

-0.0000, 0.0000, 1.0000,

rotation\_09.heic:

Quaternion: 0.130526 0 0

Input angles: X:15, Y:0, Z:0

Corresponding input matrix:

1.0000, 0.0000, 0.0000,

0.0000, 0.9659, -0.2588,

-0.0000, 0.2588, 0.9659,

rotation\_10.heic:

Quaternion: 0.12941 0.12941 -0.0170371

Input angles: X:15, Y:15, Z:0

Corresponding input matrix:

0.9659, 0.0670, 0.2500,

0.0000, 0.9659, -0.2588,

-0.2588, 0.2500, 0.9330,

rotation\_11.heic:

Quaternion: 0.111411 0.145194 0.111411

Input angles: X:15, Y:15, Z:15

Corresponding input matrix:

0.9330, -0.1853, 0.3085,

0.2500, 0.9504, -0.1853,

-0.2588, 0.2500, 0.9330,

rotation\_12.heic:

Quaternion: -0.0170371 0.12941 0.12941

Input angles: X:0, Y:15, Z:15

Corresponding input matrix:

0.9330, -0.2588, 0.2500,

0.2500, 0.9659, 0.0670,

-0.2588, 0.0000, 0.9659,

rotation\_13.heic:

Quaternion: 0 0 0.130526

Input angles: X:0, Y:0, Z:15

Corresponding input matrix:

0.9659, -0.2588, 0.0000,

0.2588, 0.9659, 0.0000,

-0.0000, 0.0000, 1.0000,

rotation\_14.heic:

Quaternion: 0 0.130526 0

Input angles: X:0, Y:15, Z:0

Corresponding input matrix:

0.9659, 0.0000, 0.2588,

0.0000, 1.0000, 0.0000,

-0.2588, 0.0000, 0.9659,

— **camera\_parameters\_1\_v1.heic** An image item with CameraIntrinsicMatrixProperty ('cmin') and CameraExtrinsicMatrixProperty ('cmex'). CameraExtrinsicMatrixProperty is with version equal to 0 and flags field is equal to 63:

— pos\_x/y/z are present

— orientation is present using quaternions and in large format

— ID is present

CameraIntrinsicMatrixProperty is with version equal to 0 and flags field equal 65793:

— denominatorShiftOperand is equal to 1

— skewDenominatorShiftOperand is equal to 1

— focal\_length\_y and skew\_factor are present

— **camera\_parameters\_2\_v1.heic** An image item with CameraIntrinsicMatrixProperty ('cmin') and CameraExtrinsicMatrixProperty ('cmex'). CameraExtrinsicMatrixProperty is with version is equal to 1 and flags field is equal 47:

— pos\_x/y/z are present

— orientation is present using ViewpointGlobalCoordinateSysRotationStruct

— ID is present

CameraIntrinsicMatrixProperty is with version equal to 0 and flags field equal 65793

— denominatorShiftOperand is equal to 1

— skewDenominatorShiftOperand is equal to 1

— focal\_length\_y and skew\_factor are present

#### text\_item\_v1.heic

A text item with ImageSpatialExtentsProperty (ispe) and TextLayoutProperty (txlo). Text item is associated with image item by an item reference of type 'text'. TextLayoutProperty has the following fields

— reference\_width = 960

— reference\_height = 960

— x = 128

— y = 128

#### text\_and\_font\_items.heic

Text and Font items with txlo and fnch descriptive item properties.

[Editors note] No description was provided in [m67810](https://dms.mpeg.expert/doc_end_user/current_document.php?id=92989).

#### progressive-diagonal.heic

The progressive-diagonal.heic conformance sample is an example where a base quality image is progressively refined starting from the top-left and ending at the bottom-right. In this figure, for each cell, the large number indicates the item\_id of the corresponding input image, and the small number at the bottom left indicates the progressive rendering step for this cell.

A grid of numbers and a long arrow

Description automatically generated

Figure 1: Progressive Diagonal Display Order

#### progressive-focus.heic

The progressive-focus.heic file, according to the ‘prgr’ entity group and the ‘prdi’ item properties stores the data with the thumbnail first, followed by the input images corresponding to the in-focus areas, and finishing with the remaining input images. The ‘prgr’ entity group and the ‘prdi’ item property indicate to a player that when realizing a progressive rendering of the grid, two intermediate reconstructions are of particular interest: a first preview of the primary item (the grid image item) using the thumbnail and an intermediate reconstruction (as shown on Figure 2) of the grid where the most interesting areas are displayed in high quality.

A black grid with white squares

Description automatically generated

Figure 2: Progressive Focus Display Order (numbers correspond to item\_IDs)

#### Conformance Samples for Region Items

9 conformance samples for region items, are described in table below. There are 7 conformance samples for the different types of geometry that can be used in a region item. There are 2 conformance samples containing a list of regions inside a single region item. One is using 16 bits for defining the characteristics of the regions, the other is using 32 bits (i.e., in the first case, the value flags & 1 for the ‘rgan’ item is 0, in the second case it is 1).

|  |  |  |
| --- | --- | --- |
| Number | Filename | Feature |
| 1 | region-point.heic | Point region with ‘udes’ annotation. |
| 2 | region-rectangle.heic | Rectangle region with ‘udes’ annotation. |
| 3 | region-ellipse.heic | Ellipse region with ‘udes’ annotation. |
| 4 | region-polygon.heic | Polygon region with ‘udes’ annotation. |
| 5 | region-polyline.heic | Polyline region with ‘udes’ annotation. |
| 6 | region-mask.heic | Mask region with ‘udes’ annotation, where the mask is defined as an item. |
| 7 | region-inline-mask.heic | Mask region with ‘udes’ annotation, where the mask is defined inline in the region item. |
| 8 | region-list.heic | One rectangle and two points regions in a single region item, with ‘udes’ annotation. |
| 9 | region-list-large.heic | One rectangle and two points regions in a single region item, with ‘udes’ annotation, where the regions are defined using a field\_size of 32. |

3 conformance samples for derived region items are described in the table below. Each of these conformance samples contains an image item and a derived image item. A region item is associated with the image item inside which the regions are defined using an item reference of type ‘cdsc’ from the region item to the image item. A derived region item, includes a ‘drgn’ item reference to the region item. The derived region item is associated with the derived image item using an item reference of type ‘cdsc’ from the derived region item to the derived image item. Each region item contains a rectangle region and two point regions and is annotated with a ‘udes’ item property.

|  |  |  |
| --- | --- | --- |
| Number | Filename | Feature |
| 1 | region-derived-crop.heic | Derived region item using a ‘clap’ transformative item property. |
| 2 | region-derived-mirror.heic | Derived region item using an ‘imir’ transformative item property. |
| 3 | region-derived-rotation.heic | Derived region item using an ‘irot’ transformative item property. |

#### region-rectangle.mp4

This conformance sample contains a video track of 5 frames corresponding to a panning over a still image. This video track is annotated with a region track containing a rectangle region corresponding to the same area in the original still image. Therefore, the rectangle region is moving to follow the panning. This rectangle region is annotated with a ‘udes’ description stored in an ‘sbgp’ box.

#### essg\_sample1.mp4

A file with a ‘essg’ restricted track, with a sample group description with version 3, and an ‘esgh’ sample group listing the v3 sample group.

A file with the ‘esgh’ sample group with ‘stsd’ and no ‘cenc’ (a synthetic VVC track with a SAP sample group marked as essential).

A file with the ‘esgh’ sample group with num\_grouping > 1.

#### essg\_sample2.mp4

A file with a ‘essg’ restricted track, with a sample group description with version 3, and an ‘esgh’ sample group listing the v3 sample group.

A file with the ‘esgh’ sample group with ‘cenc’ (an encrypted synthetic VVC track).

#### Conforming files for picture-in-picture

**NALFF\_PiP\_FourPeople\_1280x720\_BQSquare\_416x240\_Offset\_0x480.mp4**

**NALFF\_PiP\_FourPeople\_1280x720\_BQSquare\_416x240\_Offset\_864x480.mp4**

The two conforming files use the following file format features:

1. The supplementary track reference ('supm'), indicating the relationship between a main video track and a supplementary video track,
2. The matrix fields of the TrackHeaderBoxes of the supplementary video track and the main video track, indicating the position and size of the overlaying window in the main video for overlaying the supplementary video, and
3. The layer fields of the TrackHeaderBox of the supplementary video track and the main video track, indicating that the supplementary video is overlaid on top of the main video.

There is only one difference between the two conforming files, on the position of the overlaying window, one at the bottom-right corner (as illustrated by Fig. 1), the other at the bottom-left corner.

A blue and green arrows pointing to a black background

Description automatically generated

Figure 3: A picture-in-picture example by overlaying a supplemental video on top of a main video at the bottom-right corner

#### EDRAP

— **vvc\_edrap\_sap1.mp4:** This file contains two video tracks, a main track, containing a VVC single-layer bitstream with EDRAP samples, and an associated external stream track.The main track has sample entry ‘vvc1’, and the EDRAP samples are grouped by the EDRAP sample group ‘edrp’ with edrap\_type equal to 1. The main track has an aest track reference that references the associated external stream track.The associated external stream track has the restricted sample entry ‘resv’ with the ‘rinf’ box, which contains the ‘schm’ box with scheme\_type equal to ‘spkt’.

— **vvc\_edrap\_sap2.mp4:** Same as above, except that the value of edrap\_type is equal to 2.

— **vvc\_edrap\_sap3.mp4:** Same as above, except that the value of edrap\_type is equal to 3.

#### vvc\_nnsei\_1.mp4

A Single layer coded bitstream in VVC track with sample entry 'vvc1'. Neural-network post-filter characteristics sample group ('nfcs') and neural-network post-filter activation sample group ('nfas').

#### Conformance for Efficient Sub-samples

The video sequence used in the conformance files from this clause is a synthetic sequence of 3 pictures, with size of 832x480 pixels, encoded in intra mode. Each picture contains 4 subpictures (SP1 to SP4) as illustrated in Figure 4:



Figure 4: Subpicture organisation

The VVC bitstream sequence is composed of 3 frames, as illustrated in Figure 5, with:

* The first frame comprising 1 PH (Picture Header) and 4 IDR VCL NAL units,
* All other frames comprising 1 PH and 4 CRA VCL NAL units.



Figure 5: VVC bitstream sequence

All the conformance samples of the VVC sequence use single track encapsulation, with ‘vvc1’ sample entry.

— **efficientsubs\_subpictures\_vvc\_v1.mp4:** *‘subs’* version 2 syntax and *SubSampleReferenceTableBox* describing VVC subsamples corresponding to Subpicture-based sub-samples, as defined in section 11.3.6 of [4].The values for the properties of the sub-samples of the first picture are not using references to the SubSampleReferenceTableBox (*has\_reference* = 0) and are provided in the subs box.The values for the properties of the sub-samples for following frames are identical to the properties in entries of the SubSampleReferenceTableBox (has\_reference=1).

— **efficientsubs\_subsamples\_vvc\_v1.mp4:** *‘subs’* version 2 syntax and *SubSampleReferenceTableBox* describing VVC subsamples corresponding to NAL-unit-based sub-samples, as defined in section 11.3.6. of [4].The values for the properties of all sub-samples are identical to the properties in entries of the SubSampleReferenceTableBox (*has\_reference* = 1)

#### mhm1\_preselection.mp4

This conformance file contains 2 preselections for MPEG-H 3D Audio (mhm1) signaled using the EntityToGroup box. The signaled audio scene contains a channel bed as well as 2 languages (English and Spanish) with the following 2 preselections:

— “Default”: Contains the channel bed and a language selection containing English and Spanish. The default prominence of the language track is set to 0.

— “Dialog+”: Contains the channel bed and a language selection containing English and Spanish. The default prominence of the language track is set to 9 for increased intelligibility.

Additionally, both preselections contain prominence as well as position interactivity.

#### mhm2\_preselection.mp4

This conformance file contains ISOBMFF preselections for multistream MPEG-H 3D Audio (mhm2). The described audio scene contains a channel bed as well as 3 languages (English, German and Korean) that result in 3 preselections:

— “English”: Contains channel bed and English commentary.

— “German”: Contains channel bed and German commentary.

— “Korean”: Contains channel bed and Korean commentary.