|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  28th Meeting: Torino, IT, 15–21 July 2017 | Document: JCTVC-AB0026-v3 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Omnidirectional media format SEI messages** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | **Ye-Kui Wang** Qualcomm Incorporated 5775 Morehouse Drive San Diego, CA 92130, USA | Tel: Email: | +1 858 651 8345 [yekuiw@qti.qualcomm.com](mailto:yekuiw@qti.qualcomm.com) |
| *Source:* | Qualcomm Incorporated | | |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Abstract

This contribution proposes the following items related to omnidirectional video:

1. A new SEI message for signalling of region-wise packing (RWP), and
2. A new SEI message for signalling of sphere rotation.

It is asserted that the two new SEI messages are necessary for the first version of the Omnidirectional MediA Format (OMAF) to work properly.

The latest OMAF draft specification text in MPEG document M40849 includes processes for conversion of the packed picture (i.e., the decoded picture), via the projected picture, to the sphere locations. The processes include equations for RWP and rotation. If the SEI messages proposed above are adopted, similar processes should also be included in the semantics of these SEI messages, or maybe the overall process for conversion of the packed picture (i.e., the decoded picture), via the projected picture, to the sphere locations is specified separately from the semantics of the different SEI messages, including the projection indication SEI message.

It is suggested that the following topics are discussed at a joint meeting between JCT-VC and MPEG Systems (including at least experts of the OMAF AHG), and consistent decisions should be made regarding these for both OMAF and the omnidirectional video related SEI messages:

* The need of separate rotation signalling in addition to the coverage signalling.
  + If it is concluded that separate rotation is not needed because of the existence of the coverage signalling, then
    1. Change the omnidirectional projection indication SEI message syntax such that the syntax elements omni\_projection\_yaw\_center, omni\_projection\_pitch\_center, and omni\_projection\_roll\_center are present regardless of the values of sub\_geometry\_flag.
    2. Apply the same conclusion to OMAF, including removal of the projection orientation box, removal of the distinction between global and local axels, and alignment of other places (e.g., the coverage boxes) as needed.
  + Otherwise, add an SEI message for rotation signalling.
* Correction of the rotation equations, taking into account at least Section 3.2.3 of JCTVC-AB0026 (i.e., Section 2 of m41008) and JCTVC-AB0023.
* The need of equations for both rotation and coverage.

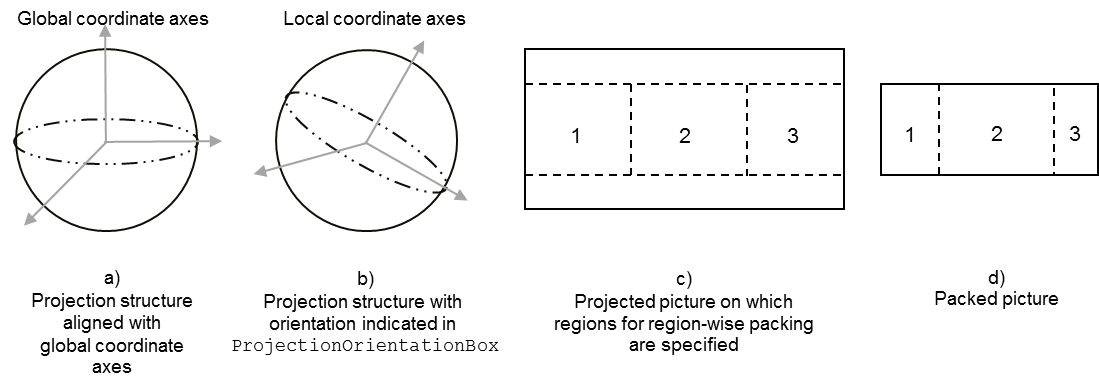
# Background

## OMAF processing steps for generation of pictures before encoding in content production

Clause 7.2.2.1 of the latest OMAF draft specification text in MPEG document M40849 is as follows:

**7.2.2.1 Relation of decoded pictures to global coordinate axes (informative)**

Figure 7‑1 illustrates the conversions from a spherical picture to a packed picture that can be used in content authoring and the corresponding conversions from a packed picture to a spherical picture to be rendered that can be used in a player.



**Figure** 7‑1 – **Example of processing stages to derive a packed picture from a spherical image or vice versa**

The projection structure is along the global coordinate axes as illustrated in Figure 7‑1a, when

* the equator of the equirectangular panorama picture is aligned with the X axis of the global coordinate axes,
* the Y axis of the equirectangular panorama picture is aligned with the Y axis of the global coordinate axes, and
* the Z axis of the global coordinate axes passes through the middle point of the equirectangular panorama picture.

The content authoring can include the following steps:

* The input images are ***stitched*** to generate a sphere picture on the unit sphere per the global coordinate axes as indicated in Figure 7‑1a.
* The unit sphere is then ***rotated*** relative to the global coordinate axes, as indicated in Figure 7‑1b. The amount of rotation is specified by the orientation angles indicated in the ProjectionOrientationBox. The local coordinate axes of the projection structure are the axes of the coordinate system that has been rotated as specified by ProjectionOrientationBox, when present. The absence of ProjectionOrientationBox indicates that the local coordinate axes are the same as the global coordinate axes.
* As illustrated in Figure 7‑1c, the spherical picture on the rotated unit sphere is then converted to a two-dimensional ***projected*** picture, for example using the equirectangular projection specified in clause 5.2.1.
* Rectangular RWP can be applied to obtain a ***packed*** picture from the projected picture. One example of packing is depicted in Figure 7‑1c and Figure 7‑1d. The dashed rectangles in Figure 7‑1c indicate the projected regions on a projected picture, and the respective areas in Figure 7‑1d indicate the corresponding packed regions. In this example, projected regions 1 and 3 are horizontally downsampled, while projected region 2 is kept at its original resolution.

GlobalCoverageInformationBox can be used to indicate which part of the sphere is covered by the packed picture.

In order to map a sample location of a packed picture (such as that in Figure 7‑1d) to a projection structure used in rendering (Figure 7‑1a), the player perform sequential mappings in reverse order from Figure 7‑1d to Figure 7‑1a.

As can be seen, OMAF content production consists of the ordered steps of stitching, rotation, projection and RWP for generation of video pictures that are sent to the encoder for encoding.

RWP enables manipulations (resize, reposition, rotation, and mirroring) of any rectangular region of the projected picture. RWP can be used to generate an emphasis on a specific viewport orientation or circumvent weaknesses of projections such as oversampling towards the poles in ERP. The latter is depicted in the top of Figure 1 where the areas near the poles of the sphere video are reduced in resolution. The bottom of Figure 1 depicts an example for an emphasized viewport orientation.



Figure 1: Two RWP examples based on the ERP format.  
*(credit: Robert Skupin of HHI)*

Rotation may be applied to improve coding efficiency or to move an object or a region to a desired position on the sphere, e.g., as discussed in JCTVC-Z0025.

## OMAF requirement of carrying rendering information in video bitstreams

In the OMAF HEVC media profiles under consideration in MPEG document N16826, the use of HEVC omnidirectional projection indication (OPI) SEI message and HEVC frame packing arrangement SEI message are mandated. The reason is to enable OMAF systems that require to access information that is needed for rendering, e.g., projection and frame packing arrangement, directly from the video elementary bitstream rather than from file format or DASH MPD, for security and content protection considerations.

However, when RWP and rotation are applied for generation of video pictures before encoding, information on RWP and rotation would also be needed for such OMAF systems to access directly from the video elementary bitstream. The most straightforward way of carrying such information in a video elementary bitstream is to use SEI messages, similarly as projection and frame packing arrangement.

## Discussion on rotation

On rotation, one may argue that this is already covered by the syntax elements omni\_projection\_yaw\_center, omni\_projection\_pitch\_center, and omni\_projection\_roll\_center of the OPI SEI message. However, these syntax elements plus omni\_projection\_yaw\_range and omni\_projection\_pitch\_range of the OPI SEI message indicate the coverage (i.e., the position and size of the region covered by the bitstream) on the sphere, they do not cover the rotation. For example, if sub\_geometry\_flag of the OPI SEI message is equal to 0, none of the coverage syntax elements are present, but rotation may still be performed. Another issue is that, while the coverage may be fixed while the rotation can be different in different CVSs (a counter argument for this point is that it is possible to make coverage different in different CVS too).

In the context of OMAF, one may also argue that the rotation can anyway be merged into the coverage, and everything is based on one set of coordinate axes (i.e., the local coordinate axes are considered as the global coordinate axes while the original global coordinate axes can be forgotten about; the initial viewpoint metadata can be used for correct rendering anyway). However, this way, when certain rotation is applied for video, then the same amount of rotation needs to be also applied to audio and other media types, if present. ***Discuss***: does this actually matter, since Clause 5.1 of OMAF mandates that the coordination systems for different media types were made aligned during content production?

# Proposal

## Joint meeting discussion on rotation signalling

It is suggested that the need of a separate rotation signalling is needed in addition to the coverage signalling, preferably at a joint meeting between JCT-VC and MPEG Systems (including at least experts of the OMAF AHG), and based on the conclusion, discuss and decide the following suggestion:

* If it is concluded that separate rotation is not needed now that we have the coverage signalling, then
  + In the OPI SEI message, change the syntax such that the syntax elements omni\_projection\_yaw\_center, omni\_projection\_pitch\_center, and omni\_projection\_roll\_center are present regardless of the values of sub\_geometry\_flag. Note that the omni\_projection\_yaw\_range and omni\_projection\_pitch\_range would still be absent when sub\_geometry\_flag is equal to 0.
  + Apply the same conclusion to OMAF, including removal of the projection orientation box, removal of the distinction between global and local axels, and alignment of other places (e.g., the coverage boxes) as needed.
* Otherwise, add an SEI message for rotation signalling.

## Omnidirectional region-wise packing SEI message

### Syntax

|  |  |
| --- | --- |
| omni\_region\_wise\_packing( payloadSize ) { | **Descriptor** |
| **omni\_region\_wise\_packing\_cancel\_flag** | u(1) |
| if( !omni\_region\_wise\_packing\_cancel\_flag ) { |  |
| **omni\_region\_wise\_packing\_persistence\_flag** | u(1) |
| **rwp\_reserved\_zero\_6bits** | u(6) |
| **num\_packed\_regions** | u(8) |
| **proj\_picture\_width** | u(16) |
| **proj\_picture\_height** | u(16) |
| for( i = 0; i < num\_packed\_regions; i++ ) { |  |
| **rwp\_reserved\_zero\_4bits**[ i ] | u(4) |
| **packing\_type**[ i ] | u(4) |
| if( packing\_type[ i ] = = 0 ) { |  |
| **proj\_reg\_width**[ i ] | u(16) |
| **proj\_reg\_height**[ i ] | u(16) |
| **proj\_reg\_top**[ i ] | u(16) |
| **proj\_reg\_left**[ i ] | u(16) |
| **transform\_type**[ i ] | u(3) |
| **rwp\_reserved\_zero\_5bits**[ i ] | u(5) |
| **packed\_reg\_width**[ i ] | u(16) |
| **packed\_reg\_height**[ i ] | u(16) |
| **packed\_reg\_top**[ i ] | u(16) |
| **packed\_reg\_left**[ i ] | u(16) |
| } |  |
| } |  |
| } |  |
| } |  |

### Semantics

The omnidirectional region-wise packing SEI message provides information to enable remapping of the colour samples of the output decoded pictures onto projected pictures. [Ed. (YK): Add definitions of "projected picture" and "packed picture" similarly as defined in the OMAF draft spec, and change "to enable remapping of the colour samples of the output decoded pictures " in the first sentence of the semantics of the OPI SEI message to "to enable remapping of the colour samples of the projected pictures".]

**omni\_region\_wise\_packing\_cancel\_flag** equal to 1 indicates that the SEI message cancels the persistence of any previous omnidirectional region-wise packing SEI message in output order. omni\_region\_wise\_packing\_cancel\_flag equal to 0 indicates that omnidirectional region-wise packing information follows.

**omni\_region\_wise\_packing\_persistence\_flag** specifies the persistence of the omnidirectional region-wise packing SEI message for the current layer.

omni\_region\_wise\_packing\_persistence\_flag equal to 0 specifies that the omnidirectional region-wise packing SEI message applies to the current decoded picture only.

Let picA be the current picture. omni\_region\_wise\_packing\_persistence\_flag equal to 1 specifies that the omnidirectional region-wise packing SEI message persists for the current layer in output order until one or more of the following conditions are true:

– A new CLVS of the current layer begins.

– The bitstream ends.

– A picture picB in the current layer in an access unit containing an omnidirectional region-wise packing SEI message that is applicable to the current layer is output for which PicOrderCnt( picB ) is greater than PicOrderCnt( picA ), where PicOrderCnt( picB ) and PicOrderCnt( picA ) are the PicOrderCntVal values of picB and picA, respectively, immediately after the invocation of the decoding process for picture order count for picB.

When an omnidirectional projection indication SEI message with omni\_projection\_information\_cancel\_flag equal to 0 is not present in the CLVS that applies to the current picture and precedes the omnidirectional region-wise packing SEI message in decoding order, an omnidirectional region-wise packing SEI message with omni\_region\_wise\_packing\_persistence\_flag equal to 0 shall not be present in the CLVS that applies to the current picture. Decoders shall ignore omnidirectional region-wise packing SEI messages with omni\_region\_wise\_packing\_persistence\_flag equal to 0 that do not follow, in decoding order, an omnidirectional projection indication SEI message with omni\_projection\_information\_cancel\_flag equal to 0 in the CLVS that applies to the current picture.

**rwp\_reserved\_zero\_6bits** shall be equal to 0 in bitstreams conforming to this version of this Specification. Other values for rwp\_reserved\_zero\_6bits[ i ] are reserved for future use by ITU-T | ISO/IEC. Decoders shall ignore the value of rwp\_reserved\_zero\_6bits[ i ].

**num\_packed\_regions** specifies the number of packed regions. The value of num\_packed\_regions shall be greater than 0.

**proj\_picture\_width** and **proj\_picture\_height** specify the width and height, respectively, of the projected picture. The value of proj\_picture\_width and proj\_picture\_height shall be both greater than 0.

**rwp\_reserved\_zero\_4bits** shall be equal to 0 in bitstreams conforming to this version of this Specification. Other values for rwp\_reserved\_zero\_4bits[ i ] are reserved for future use by ITU-T | ISO/IEC. Decoders shall ignore the value of rwp\_reserved\_zero\_4bits[ i ].

**packing\_type**[ i ] specifies the type of region-wise packing. packing\_type[ i ] equal to 0 indicates rectangular region-wise packing. Other values are reserved. The value of packing\_type[ i ] shall be equal to 0 in this version of this Specification. Decoders shall allow values of packing\_type[ i ] greater than 0 and shall ignore all omnidirectional region-wise packing SEI messages with packing\_type[ i ] greater than 0 for any value of i.

**proj\_reg\_width**[ i ], **proj\_reg\_height**[ i ], **proj\_reg\_top**[ i ], and **proj\_reg\_left**[ i ] are indicated in units of luma samples in a projected picture with width and height equal to proj\_picture\_width and proj\_picture\_height, respectively.

**proj\_reg\_width**[ i ] specifies the width of the i-th projected region. proj\_reg\_width[ i ] shall be greater than 0.

**proj\_reg\_height**[ i ] specifies the height of the i-th projected region. proj\_reg\_height[ i ] shall be greater than 0.

**proj\_reg\_top**[ i ] and **proj\_reg\_left**[ i ] specify the top luma sample row and the left-most luma sample column, respectively, in the projected picture. The values of proj\_reg\_top[ i ] and proj\_reg\_left[ i ], shall be in the range from 0, inclusive, indicating the top-left corner of the projected picture, to proj\_picture\_height − 1, inclusive, and proj\_picture\_width − 1, inclusive, respectively.

The sum proj\_reg\_width[ i ] and proj\_reg\_left[ i ] shall be less than proj\_picture\_width. The sum of proj\_reg\_height[ i ] and proj\_reg\_top[ i ] shall be less than proj\_picture\_height.

When the projected picture is stereoscopic, proj\_reg\_width[ i ], proj\_reg\_height[ i ], proj\_reg\_top[ i ] and proj\_reg\_left[ i ] shall be such that the projected region identified by these fields is within a single constituent picture of the projected picture.

**transform\_type**[ i ] specifies the rotation and mirroring that have been applied to the i-th projected region to map it to the packed picture before encoding. When transform\_type[ i ] specifies both rotation and mirroring, rotation has been applied after mirroring in the region-wise packing from the projected picture to the packed picture before encoding. The values of transform\_type[ i ] are specified in Table D. X:

Table D.X – transform\_type[ i ] values

|  |  |
| --- | --- |
| **Value** | **Description** |
| 0 | no transform |
| 1 | mirroring horizontally |
| 2 | rotation by 180 degrees (counter-clockwise) |
| 3 | rotation by 180 degrees (counter-clockwise) after mirroring horizontally |
| 4 | rotation by 90 degrees (counter-clockwise) after mirroring horizontally |
| 5 | rotation by 90 degrees (counter-clockwise) |
| 6 | rotation by 270 degrees (counter-clockwise) after mirroring horizontally |
| 7 | rotation by 270 degrees (counter-clockwise) |

**rwp\_reserved\_zero\_5bits** shall be equal to 0 in bitstreams conforming to this version of this Specification. Other values for rwp\_reserved\_zero\_5bits[ i ] are reserved for future use by ITU-T | ISO/IEC. Decoders shall ignore the value of rwp\_reserved\_zero\_5bits[ i ].

**packed\_reg\_width**[ i ], **packed\_reg\_height**[ i ], **packed\_reg\_top**[ i ], and **packed\_reg\_left**[ i ] specify the width, height, the top luma sample row, and the left-most luma sample column, respectively, of the packed region in the packed picture.

Let packedPicWidth and packedPicHeight be the width and height of the packed picture, which has the same size as the conformance cropping window. The values of packed\_reg\_width[ i ], packed\_reg\_height[ i ], packed\_reg\_top[ i ], and packed\_reg\_left[ i ] are constrained as follows:

* packed\_reg\_width[ i ] and packed\_reg\_height[ i ] shall both be greater than 0.
* The values of packed\_reg\_top[ i ] and packed\_reg\_left[ i ] shall in the range from 0, inclusive, indicating the top-left corner luma sample of the packed picture, to packedPicHeight − 1, inclusive, and packedPicWidth − 1, inclusive, respectively.
* The sum of packed\_reg\_width[ i ] and packed\_reg\_left[ i ] shall be less than packedPicWidth.
* The sum of packed\_reg\_height[ i ] and packed\_reg\_top[ i ] shall be less than packedPicHeight.
* The rectangle specified by packed\_reg\_width[ i ], packed\_reg\_height[ i ], packed\_reg\_top[ i ], and packed\_reg\_left[ i ] shall be non-overlapping with the rectangle specified by packed\_reg\_width[ j ], packed\_reg\_height[ j ], packed\_reg\_top[ j ], and packed\_reg\_left[ j ] for any value of j in the range of 0 to i – 1, inclusive.

## Omnidirectional sphere rotation SEI message

### Syntax

|  |  |
| --- | --- |
| omni\_sphere\_rotation( payloadSize ) { | **Descriptor** |
| **omni\_sphere\_rotation\_cancel\_flag** | u(1) |
| if( !omni\_sphere\_rotation\_cancel\_flag ) { |  |
| **omni\_sphere\_rotation\_persistence\_flag** | u(1) |
| **opr\_reserved\_zero\_6bits** | u(6) |
| **omni\_rotation\_yaw** | i(32) |
| **omni\_rotation\_pitch** | i(32) |
| **omni\_rotation\_roll** | i(32) |
| } |  |
| } |  |

### Semantics

The omnidirectional sphere rotation SEI message provides information to enable conversion from the local coordinate axes to the global coordinate axes. [Ed. (YK): Add definitions of "local sphere axes " and "global coordinate axes" similarly as defined in the OMAF draft spec.]

**omni\_sphere\_rotation\_cancel\_flag** equal to 1 indicates that the SEI message cancels the persistence of any previous omnidirectional sphere rotation SEI message in output order. omni\_sphere\_rotation\_cancel\_flag equal to 0 indicates that omnidirectional sphere rotation information follows.

**omni\_region\_wise\_packing\_persistence\_flag** specifies the persistence of the omnidirectional sphere rotation SEI message for the current layer.

omni\_sphere\_rotation\_persistence\_flag equal to 0 specifies that the omnidirectional sphere rotation SEI message applies to the current decoded picture only.

Let picA be the current picture. omni\_sphere\_rotation\_persistence\_flag equal to 1 specifies that the omnidirectional sphere rotation SEI message persists for the current layer in output order until one or more of the following conditions are true:

– A new CLVS of the current layer begins.

– The bitstream ends.

– A picture picB in the current layer in an access unit containing an omnidirectional sphere rotation SEI message that is applicable to the current layer is output for which PicOrderCnt( picB ) is greater than PicOrderCnt( picA ), where PicOrderCnt( picB ) and PicOrderCnt( picA ) are the PicOrderCntVal values of picB and picA, respectively, immediately after the invocation of the decoding process for picture order count for picB.

When an omnidirectional projection indication SEI message with omni\_projection\_information\_cancel\_flag equal to 0 is not present in the CLVS that applies to the current picture and precedes the omnidirectional sphere rotation SEI message in decoding order, an omnidirectional sphere rotation SEI message with omni\_sphere\_rotation\_persistence\_flag equal to 0 shall not be present in the CLVS that applies to the current picture. Decoders shall ignore omnidirectional sphere rotation SEI messages with omni\_sphere\_rotation\_persistence\_flag equal to 0 that do not follow, in decoding order, an omnidirectional projection indication SEI message with omni\_projection\_information\_cancel\_flag equal to 0 in the CLVS that applies to the current picture.

**opr\_reserved\_zero\_6bits** shall be equal to 0 in bitstreams conforming to this version of this Specification. Other values for rwp\_reserved\_zero\_6bits[ i ] are reserved for future use by ITU-T | ISO/IEC. Decoders shall ignore the value of rwp\_reserved\_zero\_6bits[ i ].

**omni\_rotation\_yaw**, **omni\_rotation\_pitch**, and **omni\_rotation\_roll** specify the yaw, pitch, and roll angles, respectively, of the rotation that has been applied to the unit sphere at the content production side to convert the global coordinate axes to the local coordinate axes.

The value of omni\_rotation\_yaw shall be in the range of −180 \* 216 to 180 \*216 − 1, inclusive.

The value of omni\_rotation\_pitch shall be in the range of −90 \* 216 to 90 \* 216, inclusive.

The value of omni\_rotation\_roll shall be in the range of −180 \* 216 to 180 \* 216 − 1, inclusive.

# Further discussions and proposals

## Equations for RWP and rotation

The latest OMAF draft specification text in MPEG document M40849 includes processes for conversion of the packed picture (i.e., the decoded picture), via the projected picture, to the sphere locations. The processes include equations for RWP and rotation. If the SEI messages proposed above are adopted, similar processes should also be included in the semantics of these SEI messages, or maybe the overall process for conversion of the packed picture (i.e., the decoded picture), via the projected picture, to the sphere locations is specified separately from the semantics of the different SEI messages, including the projection indication SEI message.

## On rotation and coverage

### General

MPEG document M41008 includes discussions and suggestions on 1) the need of separate rotation in addition to coverage, 2) the correctness of the rotation equations, and 3) the need of equations for both rotation and coverage. The first was also briefly discussed above in Section 1.3, but the one in M41008 includes more details that the author thinks would be useful. The other two discussions are also closely related to the SEI message proposed here, and it was suggested therein to have joint meeting discussions. Therefore, the discussions are also provided below in this document such that they can be reviewed by the JCT-VC before a potential joint meeting with MPEG systems.

### On the need of separate rotation in addition to coverage

As can be seen from clause 7.2.2.1 of the latest OMAF draft text, OMAF content production consists of the ordered steps of stitching, rotation, projection and region-wise packing (RWP) for generation of video pictures that are sent to the encoder for encoding.

In OMAF, rotation is signalled in the ProjectionOrientationBox. There are different understandings on the relationship between rotation and coverage. The JCT-VC meeting minutes publicly available here <http://wftp3.itu.int/av-arch/jctvc-site/2017_01_Z_Geneva/JCTVC-Z_Notes_dA.doc> include the following notes under JCTVC-Z0025, which proposed an SEI message for rotation signalling:

In the discussion, it was commented that the “sub-geometry” description previously discussed is adequate to cover the yaw and pitch aspects of this, but not the roll.

…

It was suggested that simply adding a roll angle to the other parameters that describe the coded content may be sufficient.

Decision: Include roll and persistence specification with other aspects as recorded in response to Z0036.

The author's understanding from the above notes is that, in JCT-VC rotation has been considered what is signalled by the following syntax elements of the omnidirectional projection indication SEI message in [JCTVC-AA1005](http://phenix.int-evry.fr/jct/doc_end_user/current_document.php?id=10734):

* omni\_projection\_yaw\_center
* omni\_projection\_pitch\_center
* omni\_projection\_roll\_center

These syntax elements correspond to the following syntax elements of the GlobalCoverageInformationBox in OMAF (for simplicity, let's assume that sub-picture composition tracks are not used):

* center\_yaw
* center\_pitch
* center\_roll

Not the following syntax elements of the ProjectionOrientationBox in OMAF, which are for rotation:

* orientation\_yaw
* orientation\_pitch
* orientation\_roll

Therefore, there are the following two understandings or different design choices in this context:

1. OMAF: The coverage indicates the position and size of the sphere region covered by the content on the sphere *after rotation*, i.e., the sphere per the *local coordination axes* defined in OMAF. In this case, rotation is a separate from coverage and needs a separate signalling and its own conversion process.
2. JCT-VC: The coverage indicates the position and size of the sphere region covered by the content on the sphere *before rotation*, i.e., the sphere per the *global coordination axes* defined in OMAF. In this case, rotation is part of coverage and does not need a separate signalling or its own conversion process.

So which one makes more sense?

In the JCT-VC approach, essentially everything is based on one set of coordinate axes, i.e., the local coordinate axes can be considered as the (new) global coordinate axes, while the original global coordinate axes can be forgotten about. For correct rendering, the initial viewpoint metadata can be used. In other words, this could be understood as treating rotation as a non-normative pre-processing step.

However, this way, when certain rotation is applied for video, then the same amount of rotation needs also to be applied to audio and other media types, if present. However again, does this really matter, since clause 5.1 of OMAF mandates that the coordination systems for different media types were made aligned during content production?

### On the rotation equations

Clause 5.3 of the latest OMAF draft text is as follows:

**5.3 Conversion between spherical coordinate systems of different orientations**

Inputs to this clause are:

* orientation change yaw\_center (in the range of −180, inclusive, to 180, exclusive), pitch\_center (in the range of −90, inclusive, to 90, inclusive), roll\_center (in the range of −180, inclusive, to 180, exclusive), all in units of degrees, and
* angular coordinates (φ, θ) relative to the coordinate axes that have been rotated as specified by the above orientation change.

Outputs of this clause are:

* angular coordinates (φ', θ') relative to the coordinate system specified in 5.1.

The outputs are derived as follows:

α = ( Clipyaw( φ + yaw\_center ) ) \* π ÷ 180  
β = ( Clippitch( θ + pitch\_center ) ) \* π ÷ 180  
ω = −roll\_center \* π ÷ 180  
φ' = ( Cos( ω ) \* α − Sin( ω ) \* β ) \* 180 ÷ π  
θ' = ( Sin( ω ) \* α + Cos( ω ) \* β ) \* 180 ÷ π

[Ed. (YK): The handling of yaw and pitch in the above equations might have issues, as the values of yaw could be affected by non-zero pitch and roll values. All the three angular values (yaw, pitch, and roll) may need to be converted to 3D Cartesian coordinates (x, y, z) first, then apply the 3D Cartesian rotation math, and then convert the updated values (x', y', z') back to angular values. Ed. (BD): The conversion process should be projection-agnostic for future extensions.]

These equations were generated based on the equations in the semantics of the omnidirectional projection indication SEI message in an earlier working-in-progress version of JCTVC-AA1005.

As commented in the editor's note, the math might have some issues. The author has been particularly worrying about the last two equations, which essentially does the rotation of the roll part after yaw and pitch rotations. The issue was confirmed by [JCTVC-AB0023](http://phenix.int-evry.fr/jct/doc_end_user/current_document.php?id=10742). However, at the time of writing, the uploaded version 1 of JCTVC-AB0023 did not yet include the updated equations for its preferred proposal alternative 1.

As described in JCTVC-AB0023, the rotation operations in the semantics of the omnidirectional projection indication SEI message are as follows:

Though the semantics of the SEI message describe the position of each decoded output samples in relation to the sphere, the mapping can alternatively be conceptualized as a series of rotations from a starting zero spherical surface coordinates ϕ, θ = (0, 0):

(1) first rotate the sphere along the equator by "a clockwise rotation around the up vector when viewed from the origin looking upwards", this changing longitude of the cropped output patch, by several fractional precision degrees (change in ϕ) derived from the bitstream element omni\_projection\_yaw\_center.

(2) then rotate the sphere along a meridian "by a clockwise rotation around the omni\_projection\_pitch\_center when viewed from the origin looking outwards (anticlockwise when looking toward the origin), i.e., the right vector after yaw rotation" by several fractional precision degrees (change in θ) derived from the bitstream element omni\_projection\_pitch\_center

(3) finally, the step of "a clockwise rotation around the pitch\_center when viewed from the origin looking outwards (anticlockwise when looking toward the origin), i.e., the forward vector after yaw and pitch rotation," according to several fractional precision degrees (change in ω) derived from the bitstream element omni\_projection\_roll\_center.

As can be seen from the above three-step descriptions with the help of Figure D.12 (copied below for convenience) in JCTVC-AA1005, the yaw rotation is always performed around the up vector/axis, while the pitch and roll rotations are performed around the axes dependent on the location of the center point of the coverage region after each state of the rotation. Considering only the center point of the coverage region it can be seen that the intent is correct. However, this is problematic for the pitch and roll rotations for other points of the coverage region. For example, think about a pitch rotation of 45 degrees (assuming the yaw and roll components are both 0). After the rotation the delta pitch of the point (ϕ, θ) = (0, 0) would be 45 degrees, but the delta pitch for other points on the sphere would be different. The delta pitch of the point (ϕ, θ) = (−180, 0) would be −45 degrees, while the delta pitch of the points (ϕ, θ) = (−90, 0) and (ϕ, θ) = (90, 0) would both be 0 degrees.

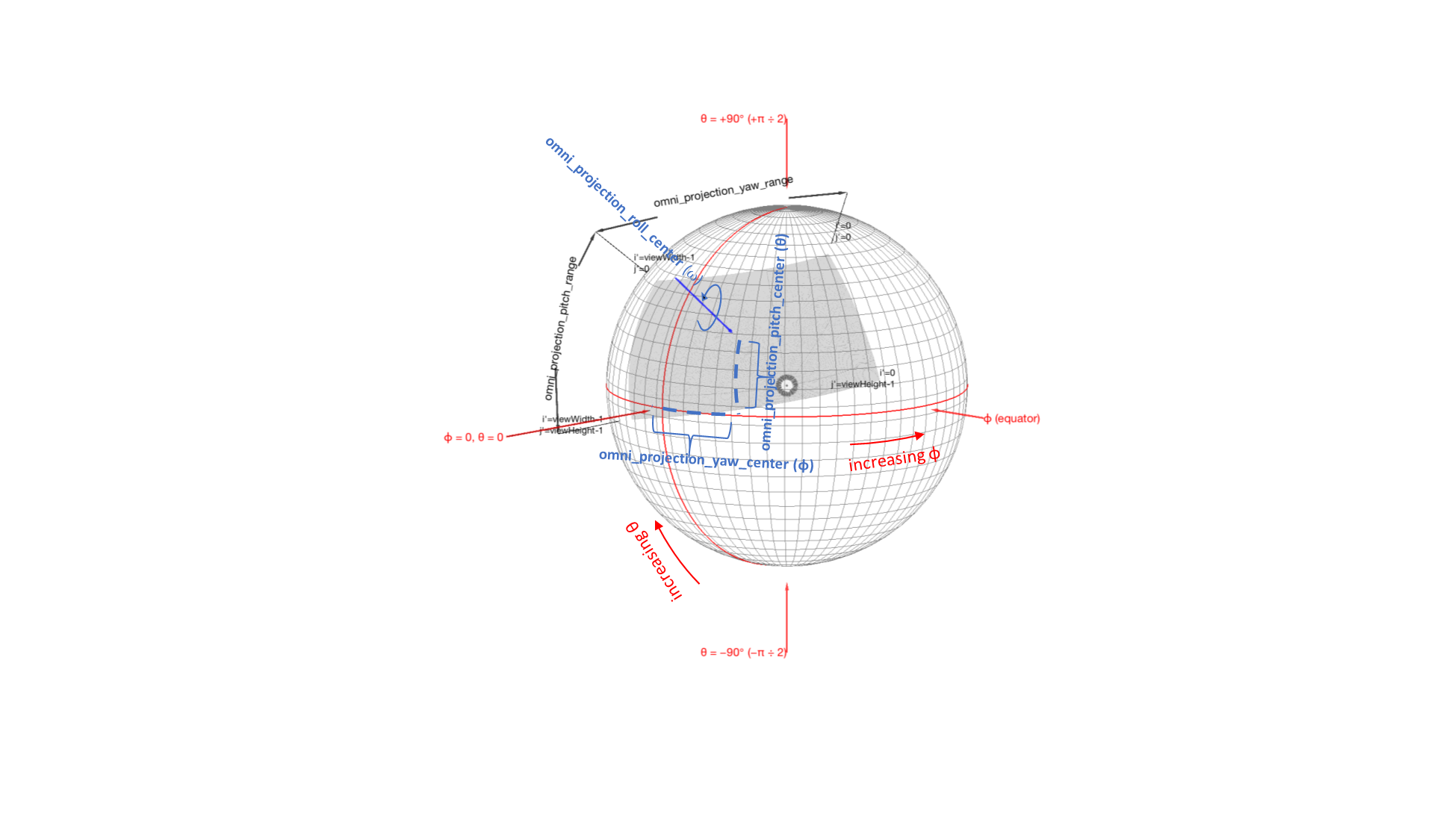


Figure D.12 – Spherical surface coordinates ϕ, θ with yaw, pitch, and roll of the region of a sphere covered by the cropped output picture relative to the equator and 0 meridian

The JVET output document [JVET-F1003](http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=3149) includes the following text at the beginning of its clause 2:

The (X, Y, Z) coordinates on the unit sphere can be evaluated from (ϕ, θ) using (1) (2) (3).

|  |  |
| --- | --- |
| X = cos(θ) cos(ϕ) | (1) |
| Y = sin(θ) | (2) |
| Z = −cos(θ) sin(ϕ) | (3) |

Inversely, the longitude and latitude (ϕ, θ) can be evaluated from (X, Y, Z) coordinates using (4)(5).

|  |  |
| --- | --- |
| ϕ = tan−1(−Z/X) | (4) |
| θ = sin−1(Y/(X2+Y2+Z2)1/2) | (5) |

And the beginning of clause 3 of JVET-F1003 includes the following text:

(Step 2) If needed, rotate the 3D point (X, Y, Z) along the three axes according to the three rotation angles (yaw, pitch, roll) to (X′, Y′, Z′) using the rotation matrix RXYZ defined in Equation (69):

|  |  |
| --- | --- |
|  | (69) |

Where the three rotation matrices around the Z, Y, and X axes, , , respectively, are defined as follows:

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |

Unless the expected updated equations for proposal alternative 1 of JCTVC-AB0023 are confirmed to be correct, both the OMAF group and JCT-VC may need to consider the approach suggested in the editor's note above, for each a point on the sphere:

* Firstly, convert the angular coordinates (yaw, pitch) to 3D Cartesian coordinates (x, y, z), by applying Equations (1)-(3) of JVET-F1003.
* Secondly, apply the 3D Cartesian rotation math to obtain the updated 3D Cartesian coordinates (x', y', z'), by applying Equation (69) of JVET-F1003.
* Lastly, convert the updated 3D Cartesian coordinates (x', y', z') back to angular values, by applying Equations (4)-(5) of JVET-F1003.

It is suggested that the rotation equations are discussed at a joint meeting between JCT-VC and MPEG Systems (including at least experts of the OMAF AHG), taking into account at least both Section 2 of m41008 and JCTVC-AB0023. A consistent decision should be made regarding this for both OMAF and the omnidirectional video related SEI messages.

### On the need of equations for both rotation and coverage

The latest OMAF draft text includes equations for project, region-wise packing (RWP), and rotation, but not for coverage. JCTVC-AA1005 currently does not support RWP and separate rotation from coverage. In the semantics of the projection indication SEI message, which includes signalling of both projection type and coverage, equations for both projection and coverage are provided.

At the Berlin OMAF AHG meeting, it was decided that the projected picture always covers the entire sphere. At the time of wring, the author realized that that decision brought a hidden functionality of the RWP signalling: among other information it now also provides the coverage information, but in the 2D Cartesian picture domain, not on the sphere domain.

This essentially makes the sphere-domain coverage information derivable. However, signalling of sphere-domain coverage information is still useful, as figuring out the sphere-domain coverage from RWP is not easy or straightforward, as the entire conversion process from the packed picture (i.e., the decoded picture), via the projected picture, to the sphere locations would be needed.

Although this observation does not invalid the need of sphere-domain coverage signalling, it does provide a good reasoning of not needing equations for coverage in addition to RWP equations.

One related side note is that, the semantics of the sphere-domain coverage should be understood and clarified as follows. Each covered sphere location shall have a corresponding sample in the decoded picture. However, there may be some sphere locations that do have corresponding samples in the decoded picture but are outside of the sphere-domain coverage region.

It is suggested that the need of equations for both rotation and coverage is discussed at a joint meeting between JCT-VC and MPEG Systems (including at least experts of the OMAF AHG), and a consistent decision regarding this should be made for both OMAF and the omnidirectional video related SEI messages.

# Patent rights declaration(s)

**Qualcomm Incorporated does not have any current or pending patent rights relating to the technology described in this contribution (to the extent of the personal awareness of the authors).**