|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  29th Meeting: Macao, CN, 19–24 Oct. 2017 | Document: JCTVC-AC0023-v4 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **On CMP padding and region-wise packing** | | |
| *Status:* | Input document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Ye-Kui Wang 5775 Morehouse Drive San Diego, CA 92122, USA  Minhua Zhou 16340 West Bernardo Drive San Diego, CA 92127, USA | Tel: Email: | +1-858-651-8345 [yekuiw@qti.qualcomm.com](mailto:yekuiw@qti.qualcomm.com)  +1-858-521-5838 [minhua.zhou@broadcom.com](mailto:minhua.zhou@broadcom.com) |
| *Source:* | Qualcomm Incorporated, Broadcom Limited | | |

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

# Abstract

This contribution proposes to align the semantics of the region-wise packing (RWP) SEI message with the RWP signalling in the Omnidirectional MediA Format (OMAF) that is being developed by the MPEG Systems subgroup, and to fix asserted issues associated with the padding signalling in the cubemap projection (CMP) SEI message.

The proposal includes the following five parts:

1. Support of an optional CMP padding along the four picture boundaries and in the middle of the picture between the boundaries of the upper three cubemap faces and the lower three cubemap faces. When the optional CMP padding exists, the RWP signalling shall not be present.
2. Text changes to the sample location remapping process for addressing the cases where the optional CMP padding exists.
3. Added a constraint that each CMP face shall be a square.
4. Alignment of the syntax and semantics of the RWP SEI message with the RWP signalling in OMAF, including addition of guard band signalling, specifying the sizes of projected/packed pictures and regions in relative units (signalling of the width and height of the packed picture is part of this), and refinement of constraints of syntax element values (allowing overlapping packed regions is part of this).
5. Some minor bug fixes related to the presence of the RWP SEI message.

Marked text changes relative to JCTVC-AB1005-v1 are provided in this document (JCTVC-AC0023). The corresponding text changes to the draft OMAF specification for alignment are provided in MPEG input document m41459.

It is claimed that the proposal resolves existing issues of padding for CMP, and makes the designs of RWP and guard band padding aligned between OMAF-related SEI messages and OMAF.

It is further claimed that the proposal resolves issues #6 and #7 in JCTVC-AC0021, and fixes some minor bug fixes related to the presence of the RWP SEI message.

# Introduction

## On CMP padding

On guard band signalling, OMAF relies on the signalling that is part of the region-wise packing (RWP) syntax, while in JCTVC-AB1005-v1, the syntax of the RWP SEI message does not include guard band signalling, and the cubemap projection (CMP) SEI message includes a guard band syntax (although it is called padding) that is different from that in OMAF.

Also, it should be noted that there some are issues in JCTVC-AB1005-v1 regarding padding or guard band signalling for CMP, as described below:

* The pictureWdith and pictureHeight (of the monoscopic projected luma picture) in the CMP equations should not count any padded samples. Therefore, when the RWP signalling is not present, in which case the size of the projected picture is not signalled, the size of the projected picture needs to be derived based on the CMP padding syntax. The width of the projected picture should be set equal to the width of the cropped output picture minus the total number of columns of padded samples, and the height should be set equal to the height of the cropped output picture minus the total number of row of padded samples, and there needs to be a constraint to require that all the remaining samples of the cropped output picture (i.e., excluding all the padded samples) shall exactly form a rectangle, which is the projected picture.
* When the RWP signalling is present for CMP, there needs to be a constraint that no packed region shall contain any padded sample.
* The semantics of the CMP padding parameters are not clear, e.g., when cmp\_padding\_type is equal to 2 or 3, and the entire semantics of cmp\_padding\_chroma\_sample\_range\_minus1 (e.g., regarding the position of the padded samples). Also the naming of the syntax element cmp\_padding\_chroma\_sample\_range\_minus1 is a bit strange. Why chroma? So the padding here has nothing to do with luma?

Specifications for the above derivation of the size of the projected picture, the constraints, and clear semantics of the CMP padding parameters are currently missing.

The above issues can be resolved by adding guard band padding signalling into the RWP syntax, same as in OMAF, and relying on that for providing support of guard band padding, i.e., remove the padding signalling from the CMP SEI message syntax. However, on the other hand, it is believed that it'd be beneficial to allow support of simple padding for CMP without the need of supporting the RWP signalling.

Therefore, a simple padding scheme for CMP is proposed that can be applied when RWP signalling is not present. Specification texts for resolving the above-mentioned issues are also provided.

## On units of sizes in RWP

In the semantics of the RWP SEI message, currently (same as in the OMAF draft text before the Torino MPEG meeting in July 2017), the unit of the size of the projected picture and the size of projected and packed regions are either unspecified or specified as luma samples. In the latest OMAF draft text, these sizes are specified in relative units, to allow the use of the same RWP syntax for multiple bitstreams representing the same source video content. For example, multiple bitstreams representing the same source video content may be generated for adaptive streaming purpose.

It is therefore suggested to align the syntax and semantics of the RWP SEI message with the RWP syntax and semantics in the latest OMAF draft text, to signal packed picture sizes and the sizes of the projected and packed regions in relative units.

# Proposal

## Introduction

The proposal includes the following five parts:

1. Support of an optional CMP padding along the four picture boundaries and in the middle of the picture between the boundaries of the upper three cubemap faces and the lower three cubemap faces. When the optional CMP padding exists, the RWP signalling shall not be present.
2. Text changes to the sample location remapping process for addressing the cases where the optional CMP padding exists.
3. Added a constraint that each CMP face shall be a square.
4. Alignment of the syntax and semantics of the RWP SEI message with the RWP signalling in OMAF, including addition of guard band signalling, specifying the sizes of projected/packed pictures and regions in relative units (signalling of the width and height of the packed picture is part of this), and refinement of the constraints on syntax element values (allowing overlapping packed regions is part of this).
5. Some minor bug fixes related to the presence of the RWP SEI message.

Changes relative to JCTVC-AB1005-v1 are marked. The corresponding text changes to the draft OMAF specification for alignment are provided in MPEG input document m41459.

## Changes to the CMP SEI message syntax and semantics

|  |  |
| --- | --- |
| cubemap\_projection( payloadSize ) { | **Descriptor** |
| **cmp\_cancel\_flag** | u(1) |
| if( !cmp\_cancel\_flag ) { |  |
| **cmp\_persistence\_flag** | u(1) |
| **cmp\_rotation\_flag** | u(1) |
| **cmp\_padding\_flag** | u(1) |
| **cmp\_reserved\_zero\_4bits** | u(4) |
| if( cmp\_padding\_flag  = =  1 ) { |  |
| **cmp\_padding\_size\_idc** | u(3) |
| **cmp\_reserved\_zero\_5bits** | u(5) |
| } |  |
| if( cmp\_rotation\_flag  = =  1 ) { |  |
| **cmp\_yaw\_rotation** | i(32) |
| **cmp\_pitch\_rotation** | i(32) |
| **cmp\_roll\_rotation** | i(32) |
| } |  |
| } |  |
| } |  |

...

**cmp\_rotation\_flag** equal to 1 indicates that a rotation for conversion between the global and the local coordinate systems applies. cmp\_rotation\_flag equal to 0 indicates that no rotation is applied and the global and local coordinate systems are identical.

**cmp\_padding\_flag** equal to 1 indicates that the cropped decoded picture contains padded areas for which the size is specified by the syntax element cmp\_padding\_size\_idc. cmp\_padding\_flag equal to 0 indicates that specifies that cropped decoded picture does not contain padded areas.

[Ed. (YK): The semantics of cmp\_padding\_flag is missing in JCTVC-AB1005-v1.]

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

**cmp\_padding\_size\_idc** specifies the size of the padded areas.

Let the variable CmpPaddingSize be set equal to 2cmp\_padding\_size\_idc + 2. When cmp\_padding\_flag is equal to 1, within the cropped decoded picture, the left-hand side CmpPaddingSize columns of samples, the right-hand side CmpPaddingSize columns of samples, the top CmpPaddingSize rows of samples, the bottom CmpPaddingSize rows of samples, and the middle CmpPaddingSize \* 2 rows of samples are all padded samples, and no other samples are padded samples.

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

**cmp\_yaw\_rotation** specifies the value of the yaw rotation angle, in units of 2−16 degrees. The value of cmp\_yaw\_rotation shall be in the range of −180 \* 216 (i.e., −11796480) to 180 \* 216 − 1 (i.e., 11796479), inclusive. When not present, the value of cmp\_yaw\_rotation is inferred to be equal to 0.

...

## Changes to the sample location remapping process

**D.3.41.5 Sample location remapping process**

***D.3.41.5.1 General***

To remap colour sample locations of a region-wise packed picture to a unit sphere, the following ordered steps are applied:

– A region-wise packed picture is obtained as the cropped output picture by decoding a coded picture. For purposes of interpretation of chroma samples, the input to the indicated remapping process is the set of decoded sample values after applying an (unspecified) upsampling conversion process to the 4:4:4 colour sampling format as necessary when chroma\_format\_idc is equal to 1 (4:2:0 chroma format) or 2 (4:2:2 chroma format). This (unspecified) upsampling process should account for the relative positioning relationship between the luma and chroma samples as indicated by chroma\_sample\_loc\_type\_top\_field and chroma\_sample\_loc\_type\_bottom\_field, when present.

– If region-wise packing is indicated, the sample locations of the region-wise packed picture are converted to sample locations of the respective projected picture as specified in clause D.3.41.5.4. Otherwise, the projected picture is identical to the region-wise packed picture.

– If frame packing is indicated, the sample locations of the projected picture are converted to sample locations of the respective constituent picture of the projected picture, as specified in clause D.3.41.5.6. Otherwise, the constituent picture of the projected picture is identical to the projected picture.

– The sample locations of a constituent picture the projected picture are converted to sphere coordinates relative to the local coordinate axes, as specified in clause D.3.41.5.2.

– If rotation is indicated, the sphere coordinates relative to the local coordinate axes are converted to sphere coordinates relative to the global coordinate axes, as specified in clause D.3.41.5.3. Otherwise, the global coordinate axes are identical to the local coordinate axes.

The overall process for mapping of luma sample locations within a region-wise packed picture to sphere coordinates relative to the global coordinate axes is normatively specified in clause D.3.41.5.5.

For each cropped decoded picture, the following applies:

– If an equirectangular projection SEI message with erp\_cancel\_flag equal to 0 and erp\_rotation\_flag equal to 1 that applies to the picture is present, or a cubemap projection SEI message with cmp\_cancel\_flag equal to 0 and cmp\_rotation\_flag equal to 1 that applies to the picture is present, RotationFlag is set equal to 1, and the following applies.

– If equirectangular projection is indicated, RotationYaw, RotationPitch, and RotationRoll are set equal to erp\_yaw\_rotation ÷ 216, erp\_pitch\_rotation ÷ 216, and erp\_roll\_rotation ÷ 216, respectively.

– Otherwise, RotationYaw, RotationPitch, and RotationRoll are set equal to cmp\_yaw\_rotation ÷ 216, cmp\_pitch\_rotation ÷ 216, and cmp\_roll\_rotation ÷ 216, respectively.

– Otherwise, RotationFlag is set equal to 0.

– If a frame packing arrangement SEI message with frame\_packing\_arrangement\_cancel\_flag equal to 0 that applies to the picture is not present, StereoFlag, TopBottomFlag, and SideBySideFlag are all set equal to 0, HorDiv1 is set equal to 1, and VerDiv1 is set equal to 1.

– Otherwise, the following applies:

– StereoFlag is set equal to 1.

– If the value of frame\_packing\_arrangement\_type of the frame packing arrangement SEI message is equal to 3, TopBottomFlag is set equal to 0, SideBySideFlag is set equal to 1, HorDiv1 is set equal to 2 and VerDiv1 is set equal to 1.

– Otherwise, if the value of frame\_packing\_arrangement\_type of the frame packing arrangement SEI message is equal to 4, TopBottomFlag is set equal to 1, SideBySideFlag is set equal to 0, HorDiv1 is set equal to 1, and VerDiv1 is set equal to 2.

– Otherwise, TopBottomFlag is set equal to 0, SideBySideFlag is set equal to 0, HorDiv1 is set equal to 1, and VerDiv1 is set equal to 1.

– If a region-wise packing SEI message with rwp\_cancel\_flag equal to 0 that applies to the picture is not present, RegionWisePackingFlag is set equal to 0, and ConstituentPicWidth and ConstituentPicHeight are set to be equal to cropPicWidth / HorDiv1 and cropPicHeight / VerDiv1, respectively, where cropPicWidth and cropPicHeight are the width and height, respectively, of the cropped output picture.

– Otherwise, RegionWisePackingFlag is set equal to 1, and ConstituentPicWidth and ConstituentPicHeight are set equal to proj\_picture\_width / HorDiv1 and proj\_picture\_height / VerDiv1, respectively.

***D.3.41.5.2 Projection for a sample location***

Inputs to this clause are:

– pictureWidth and pictureHeight, which are the width and height, respectively, of a monoscopic projected luma picture (when RegionWisePackingFlag is equal to 1) or a cropped decoded picture (when RegionWisePackingFlag is equal to 0), in luma samples, and

– the centre point of a sample location (i, j) along the horizontal and vertical axes, respectively.

Outputs of this clause are:

– sphere coordinates (φ, θ) for the sample location in degrees relative to the coordinate axes specified in clause D.3.41.5.1.

The projection for a sample location is derived as follows:

– If equirectangular projection is indicated, the following applies:

φ = ( erp\_azimuth\_min + ( 0.5 − i ÷ pictureWidth ) \* ( erp\_azimuth\_max − erp\_azimuth\_min ) ) \* 2−16  
 (D‑XX)  
θ = ( erp\_elevation\_min + ( 0.5 − j ÷ pictureHeight ) \* ( erp\_elevation\_max − erp\_elevation\_min ) ) \* 2−16  
 (D‑XX)

– Otherwise (cubemap projection is indicated), the following applies:

– When RegionWisePackingFlag is equal to 0 and cmp\_padding\_flag is equal to 1, the following applies:

i = i − CmpPaddingSize  
if( j < pictureHeight / 2)  
 j = j − CmpPaddingSize  
else (D‑XX)  
 j = j – CmpPaddingSize \* 3  
pictureWidth = pictureWidth – CmpPaddingSize \* 2  
pictureHeight = pictureHeight – CmpPaddingSize \* 4

– It is a requirement of bitstream conformance that pictureWidth shall be a multiple of 3 and pictureHeight shall be a multiple of 2, and that pictureWidth / 3 shall be equal to pictureHeight / 2, and the following applies:

lw = pictureWidth / 3  
lh = pictureHeight / 2  
i′ = −( 2 \* ( i % lw ) ÷ lw ) + 1  
j′ = −( 2 \* ( j % lh ) ÷ lh ) + 1  
w = Floor( i ÷ lw )  
h = Floor( j ÷ lh )  
if( w = = 1 && h = = 0 ) { /\* front face \*/  
 x = 1.0  
 y = −i′  
 z = j′  
} else if( w = = 1 && h = = 1 ) { /\* back face \*/  
 x = −1.0  
 y = j′  
 z = −i′  
} else if( w = = 2 && h = = 1 ) { /\* top face \*/ (D‑XX)  
 x = −i′  
 y = j′  
 z = 1.0  
} else if( w = = 0 && h = = 1 ) { /\* bottom face \*/  
 x = i′  
 y = j′  
 z = −1.0′  
} else if( w = = 0 && h = = 0 ) { /\* right face \*/  
 x = −i′  
 y = −1.0  
 z = j′  
} else { /\* ( w = = 2 && h = = 0 ), left face \*/  
 x = i′  
 y = 1.0  
 z = j′  
}  
φ = Atan2( y, x ) \* 180 ÷ π  
θ =

***D.3.41.5.3 Conversion from the local coordinate axes to the global coordinate axes***

Inputs to this clause are:

– rotation\_yaw (α), rotation\_pitch (β), rotation\_roll (γ), all in units of degrees, and

– sphere coordinates (φ, θ) relative to the local coordinate axes.

Outputs of this clause are:

– sphere coordinates (φ′, θ′) relative to the global coordinate axes.

The outputs are derived as follows:

x1 = Cos( φ ) \* Cos( θ )  
y1 = Sin( φ ) \* Cos( θ )  
z1 = Sin( θ )  
x2 = Cos( β ) \* Cos ( γ ) \* x1 − Cos( β ) \* Sin( γ ) \* y1 + Sin( β ) \* z1  
y2 = ( Cos( α ) \* Sin( γ ) + Sin( α ) \* Sin( β ) \* Cos( γ ) ) \* x1 +  
 ( Cos( α ) \* Cos( γ ) − Sin( α ) \* Sin( β ) \* Sin( γ ) ) \* y1 −  
 Sin( α ) \* Cos( β ) \* z1 (D‑XX)  
z2 = ( Sin( α ) \* Sin( γ ) − Cos( α ) \* Sin( β ) \* Cos( γ ) ) \* x1 +  
 ( Sin( α ) \* Cos( γ ) + Cos( α ) \* Sin( β ) \* Sin( γ ) ) \* y1 +  
 Cos( α ) \* Cos( β ) \* z1  
φ′ = Atan2( y2, x2 ) \* 180 ÷ π  
θ′ = Asin( z2 ) \* 180 ÷ π

***D.3.41.5.4 Conversion of sample locations for rectangular region-wise packing***

Inputs to this clause are:

– sample location (x, y) within the packed region in integer luma sample units,

– the width and the height of the projected region in luma sample units (projRegWidth, projRegHeight),

– the width and the height of the packed region in sample units (packedRegWidth, packedRegHeight),

– transform type (transformType), and

– offset values for sampling position (offsetX, offsetY).

Outputs of this clause are:

– the centre point of the sample location (i, j) within the projected region in sample units.

The outputs are derived as follows:

if( transformType  = =  0  | |  transformType  = =  1  | |  transformType  = =  2  | |  transformType  = =  3 ) {  
 horRatio = projRegWidth ÷ packedRegWidth  
 verRatio = projRegHeight ÷ packedRegHeight  
} else if ( transformType  = =  4  | |  transformType  = =  5  | |  transformType  = =  6  | |  
 transformType  = =  7 ) {  
 horRatio = projRegWidth ÷ packedRegHeight  
 verRatio = projRegHeight ÷ packedRegWidth  
}  
if( transformType = = 0 ) {  
 i = horRatio \* ( x + offsetX )  
 j = verRatio \* ( y + offsetY )  
} else if ( transformType = = 1 ) {  
 i = horRatio \* ( packedRegWidth − x − offsetX )  
 j = verRatio \* ( y + offsetY )  
} else if ( transformType = = 2 ) {  
 i = horRatio \* ( packedRegWidth − x − offsetX )  
 j = verRatio \* ( packedRegHeight − y − offsetY ) (D‑XX)  
} else if ( transformType = = 3 ) {  
 i = horRatio \* ( x + offsetX )  
 j = verRatio \* ( packedRegHeight − y − offsetY )  
} else if ( transformType = = 4 ) {  
 i = horRatio \* ( y + offsetY )  
 j = verRatio \* ( x + offsetX )  
} else if ( transformType = = 5 ) {  
 i = horRatio \* ( y + offsetY )  
 j = verRatio \* ( packedRegWidth − x − offsetX )  
} else if ( transformType = = 6 ) {  
 i = horRatio \* ( packedRegHeight − y − offsetY )  
 j = verRatio \* ( packedRegWidth − x − offsetX )  
} else if ( transformType = = 7 ) {  
 i = horRatio \* ( packedRegHeight − y − offsetY )  
 j = verRatio \* ( x+ offsetX )  
}

***D.3.41.5.5 Mapping of luma sample locations within a cropped decoded picture to sphere coordinates relative to the global coordinate axes***

This clause specifies the semantics of luma sample locations within a cropped decoded picture to sphere coordinates relative to the global coordinate axes.

offsetX is set equal to 0.5 and offsetY is set equal to 0.5.

If RegionWisePackingFlag is equal to 1, the following applies for each packed region n in the range of 0 to num\_regions − 1, inclusive:

– For each sample location (xPackedPicture, yPackedPicture) belonging to the n-th packed region with packing\_type[ n ] equal to 0 (i.e., with rectangular region-wise packing), the following applies:

– The corresponding sample location (xProjPicture, yProjPicture) of the projected picture is derived as follows:

– x is set equal to xPackedPicture – packed\_region\_left[ n ].

– y is set equal to yPackedPicture – packed\_region\_top[ n ].

– Clause D.3.41.5.4 is invoked with x, y, packed\_region\_width[ n ], packed\_region\_height[ n ], proj\_region\_width[ n ], proj\_region\_height[ n ], transform\_type[ n ], offsetX and offsetY as inputs, and the output is assigned to sample location (i, j).

– xProjPicture is set equal to proj\_region\_left[ n ] + i.

– When StereoFlag is equal to 0 or TopBottomFlag is equal to 1, and when xProjPicture is greater than or equal to proj\_picture\_width, xProjPicture is set equal to xProjPicture − proj\_picture\_width.

– When SideBySideFlag is equal to 1, the following applies:

– When proj\_region\_left[ n ] is less than proj\_picture\_width / 2 and xProjPicture is greater than or equal to proj\_picture\_width / 2, xProjPicture is set equal to xProjPicture − proj\_picture\_width / 2.

– When proj\_region\_left[n] is greater than or equal to proj\_picture\_width / 2 and xProjPicture is greater than or equal to proj\_picture\_width, xProjPicture is set equal to xProjPicture − proj\_picture\_width / 2.

– yProjPicture is set equal to proj\_region\_top[ n ] + j.

– Clause D.3.41.5.6 is invoked with xProjPicture, yProjPicture, ConstituentPicWidth, and ConstituentPicHeight as inputs, and the outputs indicating the sphere coordinates and the constituent picture index (for frame-packed stereoscopic video) for the luma sample location (xPackedPicture, yPackedPicture) belonging to the n-th packed region in the cropped decoded picture.

Otherwise (RegionWisePackingFlag is equal 0), the following applies for each sample location (x, y) that is not a cubemap projection padded sample within the cropped decoded picture:

– xPicture is set equal to x + offsetX.

– yPicture is set equal to y + offsetY.

– Clause D.3.41.5.6 is invoked with xPicture, yPicture, ConstituentPicWidth, and ConstituentPicHeight as inputs, and the outputs indicating the sphere coordinates and the constituent picture index (for frame-packed stereoscopic video) for the sample location (x, y) within the cropped decoded picture.

***D.3.41.5.6 Conversion from a sample location in a projected picture or a cropped decoded picture to sphere coordinates relative to the global coordinate axes***

Inputs to this clause are

– the centre point of a sample location (xPicture, yPicture) within a projected picture (when RegionWisePackingFlag is equal to 1) or a cropped decoded picture (when RegionWisePackingFlag is equal to 0), and

– pictureWidth and pictureHeight, which are the width and height, respectively, of a monoscopic projected luma picture (when RegionWisePackingFlag is equal to 1) or a monoscopic cropped decoded picture (when RegionWisePackingFlag is equal to 0), in luma samples.

Outputs of this clause are:

– sphere coordinates (azimuthGlobal, elevationGlobal), in units of degrees relative to the global coordinate axes, and

– when StereoFlag is equal to 1, the index of the constituent picture (constituentPicture) equal to 0 or 1.

The outputs are derived with the following ordered steps:

– If xPicture is greater than or equal to pictureWidth or yPicture is greater than or equal to pictureHeight, the following applies:

– constituentPicture is set equal to 1.

– If xPicture is greater than or equal to pictureWidth, xPicture is set to xPicture − pictureWidth.

– If yPicture is greater than or equal to pictureHeight, yPicture is set to yPicture − pictureHeight.

– Otherwise, constituentPicture is set equal to 0.

– Clause D.3.41.5.2 is invoked with pictureWidth, pictureHeight, xPicture, and yPicture as inputs, and the output is assigned to azimuthLocal, elevationLocal.

– If RotationFlag is equal to 1, clause D.3.41.5.3 is invoked with azimuthLocal, elevantionLocal, RotationYaw, RotationPitch, and RotationRoll as inputs, and the output is assigned to azimuthGlobal and elevationGlobal.

– Otherwise, azimuthGlobal is set equal to azimuthLocal and elevationGlobal is set equal to elevationLocal.

## Changes to RWP SEI message syntax and semantics

|  |  |
| --- | --- |
| regionwise\_packing( payloadSize ) { | **Descriptor** |
| **rwp\_cancel\_flag** | u(1) |
| if( !rwp\_cancel\_flag ) { |  |
| **rwp\_persistence\_flag** | u(1) |
| **constituent\_picture\_matching\_flag** | u(1) |
| **rwp\_reserved\_zero\_5bits** | u(5) |
| **num\_packed\_regions** | u(8) |
| **proj\_picture\_width** | u(16) |
| **proj\_picture\_height** | u(16) |
| **packed\_picture\_width** | u(16) |
| **packed\_picture\_height** | u(16) |
| for( i = 0; i < num\_packed\_regions; i++ ) { |  |
| **rwp\_reserved\_zero\_3bits**[ i ] | u(3) |
| **guard\_band\_flag**[ i ] | u(1) |
| **packing\_type**[ i ] | u(4) |
| if( packing\_type[ i ]  = =  0 ) { |  |
| **proj\_region\_width**[ i ] | u(16) |
| **proj\_region\_height**[ i ] | u(16) |
| **proj\_region\_top**[ i ] | u(16) |
| **proj\_region\_left**[ i ] | u(16) |
| **transform\_type**[ i ] | u(3) |
| **rwp\_reserved\_zero\_5bits**[ i ] | u(5) |
| **packed\_region\_width**[ i ] | u(16) |
| **packed\_region\_height**[ i ] | u(16) |
| **packed\_region\_top**[ i ] | u(16) |
| **packed\_region\_left**[ i ] | u(16) |
| if( guard\_band\_flag[ i ] ) { |  |
| **left\_gb\_width**[ i ] | u(8) |
| **right\_gb\_width**[ i ] | u(8) |
| **top\_gb\_height**[ i ] | u(8) |
| **bottom\_gb\_height**[ i ] | u(8) |
| **gb\_not\_used\_for\_pred\_flag**[ i ] | u(1) |
| for( j = 0; j < 4; j++ ) |  |
| **gb\_type**[ i ][ j ] | u(3) |
| **rwp\_gb\_reserved\_zero\_3bits**[ i ] | u(3) |
| } |  |
| } |  |
| } |  |
| } |  |
| } |  |

The region-wise packing SEI message provides information to enable remapping of the colour samples of the cropped decoded pictures onto projected pictures as well as information on the location and size of the guard bands, if any.

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

**rwp\_persistence\_flag** specifies the persistence of the region-wise packing SEI message for the current layer.

rwp\_persistence\_flag equal to 0 specifies that the region-wise packing SEI message applies to the current decoded picture only.

Let picA be the current picture. rwp\_persistence\_flag equal to 1 specifies that the 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 a 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 equirectangular projection SEI message with erp\_cancel\_flag equal to 0 or a cubemap projection SEI message with cmp\_cancel\_flag equal to 0 and cmp\_padding\_flag equal to 0 is not present in the CLVS that applies to the current picture and precedes the region-wise packing SEI message in decoding order, a region-wise packing SEI message with rwp\_cancel\_flag equal to 0 shall not be present in the CLVS that applies to the current picture. Decoders shall ignore region-wise packing SEI messages with rwp\_cancel\_flag equal to 0 that do not follow, in decoding order, an equirectangular projection SEI message with erp\_cancel\_flag equal to 0 or a cubemap projection SEI message with cmp\_cancel\_flag equal to 0 in the CLVS that applies to the current picture. [Ed. (YK): Note that the changes in this paragraph are bug fixes that should be made independent of the issues described in section 1 of this document.]

If a frame packing arrangement SEI message with frame\_packing\_arrangement\_cancel\_flag equal to 0, frame\_packing\_arrangement\_type equal to 3, 4, or 5, and quincunx\_sampling\_flag equal to 0 is not present that applies to the current picture, the variables StereoFlag, TopBottomFlag, SideBySideFlag, and TempInterleavingFlag are all set equal to 0, the variables HorDiv1 and VerDiv1 are both set equal to 1. Otherwise the following applies:

– StereoFlag is equal to 1.

– When the frame\_packing\_arrangement\_type is equal to 3, SideBySideFlag is set equal to 1, TopBottomFlag and TempInterleavingFlag are both set equal to 0, HorDiv1 is set equal to 2 and VerDiv1 is set equal to 1.

– When the frame\_packing\_arrangement\_type is equal to 4, TopBottomFlag is set equal to 1, SideBySideFlag and TempInterleavingFlag are both set equal to 0, HorDiv1 is set equal to 1 and VerDiv1 is set equal to 2.

– When the frame\_packing\_arrangement\_type is equal to 5, TempInterleavingFlag is set equal to 1, TopBottomFlag and TempInterleavingFlag are both set equal to 0, HorDiv1 and VerDiv1 are both set equal to 1.

**constituent\_picture\_matching\_flag** equal to 1 specifies that the projected region information, packed region information, and guard band region information in this SEI message apply individually to each constituent picture and that the packed picture and the projected picture have the same stereoscopic frame packing format indicated by the frame packing arrangement SEI message. constituent\_picture\_matching\_flag equal to 0 specifies that the projected region information, packed region information, and guard band region information in this SEI message apply to the projected picture.

When a frame packing arrangement SEI message with frame\_packing\_arrangement\_cancel\_flag equal to 0, frame\_packing\_arrangement\_type equal to 3, 4, or 5, and quincunx\_sampling\_flag equal to 0 is not present that applies to the current picture, the value of constituent\_picture\_matching\_flag shall be equal to 0.

**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 ].

**num\_packed\_regions** specifies the number of packed regions when constituent\_picture\_matching\_flag is equal to 0. The value of num\_packed\_regions shall be greater than 0. When constituent\_picture\_matching\_flag is equal to 1, the number of packed regions is equal to num\_regions \* 2, and the information in each entry of the loop of num\_packed\_regions entries applies to each constituent picture of the projected picture and the packed picture.

**proj\_picture\_width** and **proj\_picture\_height** specify the width and height, respectively, of the projected picture, in relative projected picture sample units.

The values of proj\_picture\_width and proj\_picture\_height shall both be greater than 0.

**packed\_picture\_width** and **packed\_picture\_height** specify the width and height, respectively, of the packed picture, in relative packed picture sample units.

The values of packed\_picture\_width and packed\_picture\_height shall both be greater than 0.

It is a requirement of bitstream conformance that packed\_picture\_width and packed\_picture\_height shall have such values that packed\_region\_width[ i ], packed\_region\_height[ i ], packed\_region\_top[ i ], and packed\_region\_left[ i ], represent integer horizontal and vertical coordinates of luma sample units within the cropped decoded pictures.

**rwp\_reserved\_zero\_3bits**[ i ] shall be equal to 0 in bitstreams conforming to this version of this Specification. Other values for rwp\_reserved\_zero\_3bits[ i ] are reserved for future use by ITU-T | ISO/IEC. Decoders shall ignore the value of rwp\_reserved\_zero\_3bits[ i ].

**guard\_band\_flag**[ i ] equal to 0 specifies that the i-th packed region does not have a guard band. guard\_band\_flag[ i ] equal to 1 specifies that the i-th packed region has a guard band.

**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 region-wise packing SEI messages with packing\_type[ i ] greater than 0 for any value of i.

**proj\_region\_width**[ i ], **proj\_region\_height**[ i ], **proj\_region\_top**[ i ] and **proj\_region\_left**[ i ] specify the width, height, top sample row, and the left-most sample column, respectively, of the i-th projected region, either within the projected picture (when constituent\_picture\_matching\_flag is equal to 0) or within the constituent picture of the projected picture (when constituent\_picture\_matching\_flag is equal to 1).

proj\_region\_width[ i ], proj\_region\_height[ i ], proj\_region\_top[ i ], and proj\_region\_left[ i ] are indicated in relative projected picture sample units.

NOTE 1 – Two projected regions may partially or entirely overlap with each other.

The values of proj\_region\_width[ i ], proj\_region\_height[ i ], proj\_region\_top[ i ] and proj\_region\_left[ i ] are constrained as follows:

– If SideBySideFlag is equal to 0, proj\_region\_width[ i ] shall be in the range of 1 to proj\_picture\_width, inclusive. Otherwise (SideBySideFlag is equal to 1), proj\_region\_width[ i ] shall be in the range of 1 to proj\_picture\_width / 2, inclusive.

– proj\_region\_height[ i ] shall be greater than 0.

– If constituent\_picture\_matching\_flag is equal to 0, the values of proj\_region\_top[ i ] and proj\_region\_left[ i ] shall be in the range from 0, indicating the top-left corner of the projected picture, to proj\_picture\_height − 1, inclusive, and proj\_picture\_width − 1, inclusive, respectively. Otherwise (constituent\_picture\_matching\_flag is equal to 1), the values of proj\_region\_top[ i ] and proj\_region\_left[ i ] shall be in the range from 0 to proj\_picture\_height / VerDiv1 − 1, inclusive, and proj\_picture\_width / HorDiv1 − 1, inclusive, respectively.

– When StereoFlag is equal to 0 or constituent\_picture\_matching\_flag is equal to 0, proj\_region\_height[ i ] and proj\_region\_top[ i ] shall be constrained such that proj\_region\_height[ i ] + proj\_region\_top[ i ] − 1 is less than proj\_picture\_height.

– When StereoFlag is equal to 1, the following applies:

– proj\_region\_width[ i ] shall be less than or equal to proj\_picture\_width / HorDiv1.

– proj\_region\_height[ i ] shall be less than or equal to proj\_picture\_height / VerDiv1.

– If proj\_region\_left[ i ] is less than proj\_picture\_width / HorDiv1, proj\_region\_width[ i ] and proj\_region\_left[ i ] shall be constrained such that proj\_region\_width[ i ] + proj\_region\_left[ i ] − 1 is less than proj\_picture\_width / HorDiv1. Otherwise (proj\_region\_left[ i ] is greater than or equal to proj\_picture\_width / HorDiv1), proj\_region\_width[ i ] and proj\_region\_left[ i ] shall be constrained such that proj\_region\_width[ i ] + proj\_region\_left[ i ] − proj\_picture\_width / HorDiv1 − 1 is less than proj\_picture\_width / HorDiv1.

– If proj\_region\_top[ i ] is less than proj\_picture\_height / VerDiv1, proj\_region\_height[ i ] and proj\_region\_top[ i ] shall be constrained such that proj\_region\_height[ i ] + proj\_region\_top[ i ] − 1 is less than proj\_picture\_height / VerDiv1. Otherwise (proj\_region\_top[ i ] is greater than or equal to proj\_picture\_height / VerDiv1), proj\_region\_height[ i ] and proj\_region\_top[ i ] shall be constrained such that proj\_region\_height[ i ] + proj\_region\_top[ i ] − proj\_picture\_height / VerDiv1 − 1 is less than proj\_picture\_height / VerDiv1.

– When constituent\_picture\_matching\_flag is equal to 1, the following applies:

– proj\_region\_width[ i ] and proj\_region\_left[ i ] shall be constrained such that proj\_region\_width[ i ] + proj\_region\_left[ i ] − 1 is less than proj\_picture\_width / HorDiv1.

– proj\_region\_height[ i ] and proj\_region\_top[ i ] shall be constrained such that proj\_region\_height[ i ] + proj\_region\_top[ i ] − 1 is less than proj\_picture\_height / VerDiv1.

**transform\_type**[ i ] specifies the rotation and mirroring to be applied to the i-th packed region to remap to the i-th projected region. When transform\_type[ i ] specifies both rotation and mirroring, rotation applies before mirroring. 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) before mirroring horizontally |
| 5 | rotation by 90 degrees (counter-clockwise) |
| 6 | rotation by 270 degrees (counter-clockwise) before 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\_region\_width**[ i ], **packed\_region\_height**[ i ], **packed\_region\_top**[ i ], and **packed\_region\_left**[ i ] specify the width, height, the top luma sample row, and the left-most luma sample column, respectively, of the packed region, either within the region-wise packed picture (when constituent\_picture\_matching\_flag is equal to 0) or within each constituent picture of the region-wise packed picture (when constituent\_picture\_matching\_flag is equal to 1).

packed\_region\_width[ i ], packed\_region\_height[ i ], packed\_region\_top[ i ], and packed\_region\_left[ i ] are indicated in relative region-wise packed picture sample units.

NOTE 2 – Two packed regions may partially or entirely overlap with each other.

The values of packed\_region\_width[ i ], packed\_region\_height[ i ], packed\_region\_top[ i ], and packed\_region\_left[ i ] are constrained as follows:

– packed\_region\_width[ i ] and packed\_region\_height[ i ] shall both be greater than 0.

– The values of packed\_region\_top[ i ] and packed\_region\_left[ i ] shall be in the range from 0, indicating the top-left corner luma sample of the region-wise packed picture, to packed\_picture\_height − 1, inclusive, and packed\_picture\_width − 1, inclusive, respectively.

– If constituent\_picture\_matching\_flag is equal to 0, the following applies:

– The sum of packed\_region\_width[ i ] and packed\_region\_left[ i ] minus 1 shall be less than packed\_picture\_width.

– The sum of packed\_region\_height[ i ] and packed\_region\_top[ i ] minus 1 shall be less than packed\_picture\_height.

– Otherwise (constituent\_picture\_matching\_flag is equal to 1), the following applies:

– The sum of packed\_region\_width[ i ] and packed\_region\_left[ i ] minus 1 shall be less than packed\_picture\_width / HorDiv1.

– The sum of packed\_region\_height[ i ] and packed\_region\_top[ i ] minus 1 shall be less than packed\_picture\_height / VerDiv1.

– When the decoded picture has 4:2:0 or 4:2:2 chroma format, packed\_region\_left[ i ] shall correspond to an even horizontal coordinate value of luma sample units, and packed\_region\_width[ i ] shall correspond to an even number of luma samples, both within the cropped decoded picture.

– When the decoded picture has 4:2:0 chroma format, packed\_region\_top[ i ] shall correspond to an even vertical coordinate value of luma sample units, and packed\_region\_height[ i ] shall correspond to an even number of luma samples, both within the cropped decoded picture.

**left\_gb\_width**[ i ] specifies the width of the guard band on the left side of the i-th packed region in relative region-wise packed picture sample units. When the decoded picture has 4:2:0 or 4:2:2 chroma format, left\_gb\_width[ i ] shall correspond to an even number of luma samples within the cropped decoded picture.

**right\_gb\_width**[ i ] specifies the width of the guard band on the right side of the i-th packed region in relative region-wise packed picture sample units. When the decoded picture has 4:2:0 or 4:2:2 chroma format, right\_gb\_width[ i ] shall correspond to an even number of luma samples within the cropped decoded picture.

**top\_gb\_height**[ i ] specifies the height of the guard band above the i-th packed region in relative region-wise packed picture sample units. When the decoded picture has 4:2:0 chroma format, top\_gb\_height[ i ] shall correspond to an even number of luma samples within the cropped decoded picture.

**bottom\_gb\_height**[ i ] specifies the height of the guard band below the i-th packed region in relative region-wise packed picture sample units. When the decoded picture has 4:2:0 chroma format, bottom\_gb\_height[ i ] shall correspond to an even number of luma samples within the cropped decoded picture.

When guard\_band\_flag[ i ] is equal to 1, left\_gb\_width[ i ], right\_gb\_width[ i ], top\_gb\_height[ i ], or bottom\_gb\_height[ i ] shall be greater than 0.

The i-th packed region as specified by this SEI message shall not overlap with any other packed region specified by the same SEI message or any guard band specified by the same SEI message.

The guard bands associated with the i-th packed region, if any, as specified by this SEI message shall not overlap with any packed region specified by the same SEI message or any other guard bands specified by the same SEI message.

**gb\_not\_used\_for\_pred\_flag**[ i ] equal to 0 specifies that the guard bands may or may not be used in the inter prediction process. gb\_not\_used\_for\_pred\_flag[ i ] equal to 1 specifies that the sample values of the guard bands are not used in the inter prediction process.

NOTE 3 – When gb\_not\_used\_for\_pred\_flag[ i ] is equal to 1, the sample values within guard bands in cropped decoded pictures can be rewritten even if the cropped decoded pictures were used as references for inter prediction of subsequent pictures to be decoded. For example, the content of a packed region can be seamlessly expanded to its guard band with decoded and re-projected samples of another packed region.

**gb\_type**[ i ][ j ] specifies the type of the guard bands for the i-th packed region as follows, with j equal to 0, 1, 2, or 3 indicating that the semantics below apply to the left, right, top, or bottom edge, respectively, of the packed region:

– gb\_type[ i ][ j ] equal to 0 specifies that the content of the guard bands in relation to the content of the packed regions is unspecified. When gb\_not\_used\_for\_pred\_flag[i] is equal to 0, gb\_type[ i ][ j ] shall not be equal to 0.

– gb\_type[ i ][ j ] equal to 1 specifies that the content of the guard bands suffices for interpolation of sample values at sub-pel sample fractional locations within the packed region and less than sample outside of the boundary of the packed region.

NOTE 4 – gb\_type[ i ][ j ] equal to 1 can be used when the boundary samples of a packed region have been copied horizontally or vertically to the guard band.

– gb\_type[ i ][ j ] equal to 2 specifies that the content of the guard bands represents actual picture content at quality that gradually changes from the picture quality of the packed region to that of the spherically adjacent packed region.

– gb\_type[ i ][ j ] equal to 3 specifies that the content of the guard bands represents actual picture content at the picture quality of the packed region.

– gb\_type[ i ][ j ] values greater than 3 are reserved. Decoders shall ignore the value of gb\_type[ i ][ j ] when the value is greater than 3.

**rwp\_gb\_reserved\_zero\_3bits**[ i ] shall be equal to 0 in bitstreams conforming to this version of this Specification. Other values for rwp\_gb\_reserved\_zero\_3bits[ i ] are reserved for future use by ITU-T | ISO/IEC. Decoders shall ignore the value of rwp\_gb\_reserved\_zero\_3bits[ i ].

# Patent rights declaration(s)

**Qualcomm Incorporated may have current or pending patent rights relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).**

**Broadcom Limited may have current or pending patent rights relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).**