|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  31st Meeting: San Diego, US, 13–20 Apr. 2018 | Document: JCTVC-AE0028-v1 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Fisheye video information SEI message** | | |
| *Status:* | Input document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Hyun-Mook Oh Sejin Oh | Email: | [hyunmook.oh@lge.com](mailto:hyunmook.oh@lge.com) |
| *Source:* | LG Electronics Inc. | | |

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

# Abstract

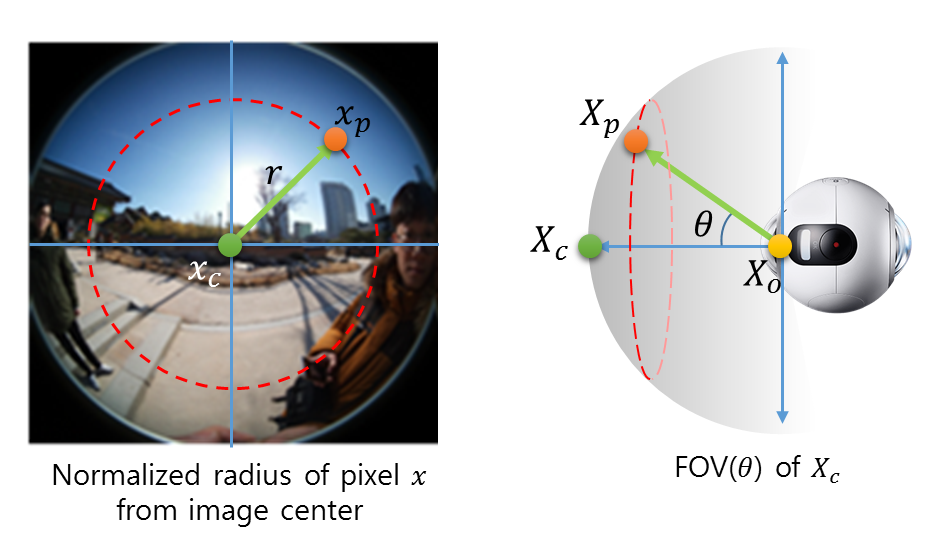
In this document, improved semantics of fisheye video information SEI message in JCTVC-AD1005 are proposed to resolve the comment from the last JCT-VC meeting and the editorial comments in the output document. Based on the definition of the polynomial coefficient in OMAF v1, the conversion equation using the polynomial coefficient is proposed. In addition, the corrections on the semantics of the ranges of rectangular region parameters, tilt angle, and field of view are proposed.

# Problem statements

Fisheye video information SEI message defines the parameter values those are essential in stitching process in 360 video applications[1]. To help the receivers to reconstruct an omnidirectional video from the raw images in a decoded picture, parameter values that describe the active region captured by a fisheye lens, field of view (FoV) and viewing orientation of the fisheye lens, center offset of the fisheye lens from the camera center, and the polynomial coefficients for luma sample to sphere conversion process are defined.

Based on the discussion in the 30th JCT-VC meeting in Gwangju, the SEI message was adopted in JCTVC-AD1005[2] with the comment in the meeting note that “F*urther work is needed to provide equations that explicitly account for the distortion model*.” To resolve the comment, the conversion equation using polynomial coefficient in OMAF v1 could be referenced[3], where the semantics are given as follows:

num\_polynomial\_coefs\_distortion and polynomial\_coef\_k\_distortion are lens distortion parameters that define the curve function between the normalized radius (*r*) of luma sample *xp* in the circular image and the angle () in the camera sphere where *Xp* and *Xc* are the corresponding points of *xp* and xc as shown in Figure 6‑2. The normalized radius (*r*) of *xp* is calculated as the distance (*d*) divided by full\_radius where the distance (*d*) is the distance between *xp* and the centre of the circular image, *xc* in units of luma samples.



**Figure** **6‑2 – Illustration of relation between radius and theta of distortion function**

num\_polynomial\_coefs\_distortion is an integer that specifies the number of polynomial coefficients corresponding to the circular image. This is the maximum order of the polynomial plus 1.

The instances of polynomial polynomial\_coef\_k\_distortion are fixed-point 8.24 polynomial coefficient values that describes the curve function from the normalized radius (*r*) of luma sample *xp* in the circular image to the angle (), in radians, using the following polynomial equation:



(6‑1)



where and are represented by polynomial\_coef\_k\_distortion and num\_polynomial\_coefs\_distortion, respectively.



Based on the definition in OMAF, the conversion equation of polynomial coefficient is proposed to be included in the fisheye video information SEI message. In addition, to resolve the editorial comments in the output document, corrections on the semantics of the SEI message are proposed with regard to the ranges of rectangular region parameters, tilt angle, and field of view.

# Proposed improvement on fisheye video information SEI message

The proposed improvements on the fisheye video information SEI message are:

* additional conversion equation for the polynomial coefficient (based on the definition in OMAF v1 text),
* clarification on the ranges of fisheye\_rect\_region\_top[ i ], fisheye\_rect\_region\_left[ i ], fisheye\_rect\_region\_width[ i ], and fisheye\_rect\_region\_height[ i ], and
* correction on the semantics of fisheye\_camera\_centre\_tilt[ i ] and fisheye\_field\_of\_view[ i ], respectively.

In the following, the additional texts, equations, and proposed corrections are pointed with green highlights.

**D.2.41.3 Fisheye video information SEI message syntax**

|  |  |
| --- | --- |
| fisheye\_video\_info( payloadSize ) { | **Descriptor** |
| **fisheye\_view\_dimension\_idc** | u(3) |
| **fisheye\_reserved\_zero\_5bits** | u(5) |
| **fisheye\_num\_active\_areas\_minus1** | u(8) |
| for( i = 0; i  <=  fisheye\_num\_active\_areas\_minus1; i++ ) { |  |
| **fisheye\_circular\_region\_centre\_x**[ i ] | u(32) |
| **fisheye\_circular\_region\_centre\_y**[ i ] | u(32) |
| **fisheye\_rect\_region\_top**[ i ] | u(32) |
| **fisheye\_rect\_region\_left**[ i ] | u(32) |
| **fisheye\_rect\_region\_width**[ i ] | u(32) |
| **fisheye\_rect\_region\_height**[ i ] | u(32) |
| **fisheye\_circular\_region\_radius**[ i ] | u(32) |
| **fisheye\_scene\_radius**[ i ] | u(32) |
| **fisheye\_camera\_centre\_azimuth**[ i ] | i(32) |
| **fisheye\_camera\_centre\_elevation**[ i ] | i(32) |
| **fisheye\_camera\_centre\_tilt**[ i ] | i(32) |
| **fisheye\_camera\_centre\_offset\_x**[ i ] | u(32) |
| **fisheye\_camera\_centre\_offset\_y**[ i ] | u(32) |
| **fisheye\_camera\_centre\_offset\_z**[ i ] | u(32) |
| **fisheye\_field\_of\_view**[ i ] | u(32) |
| **fisheye\_num\_polynomial\_coeffs**[ i ] | u(16) |
| for( j = 0; j < fisheye\_num\_polynomial\_coeffs[ i ]; j++ ) |  |
| **fisheye\_polynomial\_coeff**[ i ][ j ] | i(32) |
| } |  |
| } |  |

***D.3.41.1.1 Genaral***

……

If fisheye video is indicated, the sample locations in all of the active regions in the decoded picture are converted to locations of the sphere coordinates relative to the global coordinate axes, as specified in clause D.3.41.1.7.

***D.3.41.1.7 Conversion from a sample location of an active area to sphere coordinates relative to the global coordinate axes***

[Ed. (YK): When a fisheye video information SEI message that applies to a picture is present, the sample resampling process for fisheye video specified in clause D.3.41.1.7 needs to be invoked in clause D.3.41.1.1 that contains the general description of the overall sample location remapping process.]

Input to this clause are:

– the sample location (x, y) in units of luma samples,

– the centre location (xc, yc) and the radius (rc) of the circular region that contains the i-th active area, given by fisheye\_circular\_region\_centre\_x[ i ], fisheye\_circular\_region\_centre\_y[ i ], and fisheye\_circular\_region\_radius[ i ], respectively, all in units of 2−16 luma samples,

– the field of view (θv) of the lens corresponding to the i-th active area, given by fisheye\_field\_of\_view[ i ], in units of 2−16 degrees, and

– the rotation parameters (αc, βc, γc), given by fisheye\_camera\_centre\_azimuth[ i ], fisheye\_camera\_centre\_elevation[ i ], and fisheye\_camera\_centre\_tilt[ i ], respectively, all in units of 2−16 degrees.

– the number of polynomial coefficients (n) and the polynomial coefficients (pj) of the i-th active area, given by fisheye\_num\_polynomial\_coeffs[ i ] and fisheye\_polynomial\_coeff [ i ][ j ] in units of 2−24, respectively.

Output of this clause are:

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

The outputs are derived as follows:

If( n = = 0 )

ϕ′ = ( Sqrt( ( x − xc ÷ 216 )2 + ( y − yc ÷ 216 )2 ) ÷ ( rc ÷ 216 ) ) \* ( θv ÷ 216 \* π ÷ 180 ) ÷ 2

else

ϕ′ = ( ( pj \* 2−24 ) \* ( Sqrt( ( x – xc \* 2−16 )2 + ( y – yc \* 2−16 )2 ) ÷ ( rc \* 2−16 ) )j )

Given ϕ′

θ′ = Atan2( y − yc ÷ 216, x − xc ÷ 216 )  
x1 = Cos( ϕ′ )  
y1 = Sin( ϕ′ ) \* Cos( θ′ )  
z1 = Sin( ϕ′ ) \* Sin( θ′ )  
α = ( αc ÷ 216 ) \* π ÷ 180  
β = ( βc ÷ 216 ) \* π ÷ 180  
γ = ( γc ÷ 216 ) \* π ÷ 180  
x2 = Cos( β ) \* Cos ( γ ) \* x1 − Cos( β ) \* Sin( γ ) \* y1 + Sin( β ) \* z1 (D‑XX)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 ) \* z1z2 = ( 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 ) \* z1ϕ = Atan2( y2, x2 ) \* 180 ÷ π  
θ = Asin( z2 ) \* 180 ÷ π

**D.3.41.4 Fisheye video information SEI message semantics**

The presence of the fisheye video information SEI message in a CLVS indicates that each coded video picture in the CLVS is a fisheye video information video picture containing a number of active areas captured by fisheye camera lens. The information of the fisheye video information video carried in the fisheye video information SEI message can be used by a receiver to properly render the fisheye video information video.

The fisheye video information SEI message applies to the CLVS that contains the SEI message, also referred to as the current CLVS. When present in a CVLS the fisheye video information SEI message shall be present in the first access unit of the CLVS and may be present in other access units of the CLVS.

**fisheye\_view\_dimension\_idc** indicates the alignment and viewing direction of a fisheye lens, as follows:

– fisheye\_view\_dimension\_idc equal to 0 indicates that fisheye\_num\_active\_areas is equal to 2, and the values of fisheye\_camera\_centre\_azimuth, fisheye\_camera\_centre\_elevation, fisheye\_camera\_centre\_tilt, fisheye\_camera\_centre\_offset\_x, fisheye\_camera\_centre\_offset\_y, and fisheye\_camera\_centre\_offset\_z are such that the active areas have aligned optical axes and face opposite directions, and the sum of fisheye\_field\_of\_view values is greater than or equal to 360 \* 216.

– fisheye\_view\_dimension\_idc equal to 1 indicates that fisheye\_num\_active\_areas is equal to 2, and the values of fisheye\_camera\_centre\_azimuth, fisheye\_camera\_centre\_elevation, fisheye\_camera\_centre\_tilt, fisheye\_camera\_centre\_offset\_x, fisheye\_camera\_centre\_offset\_y, and fisheye\_camera\_centre\_offset\_z are such that the active areas have parallel optical axes that are orthogonal to the line intersecting the camera centre points, and the camera corresponding to i equal to 0 is the left view.

– fisheye\_view\_dimension\_idc equal to 2 indicates that fisheye\_num\_active\_areas is equal to 2, and the values of fisheye\_camera\_centre\_azimuth, fisheye\_camera\_centre\_elevation, fisheye\_camera\_centre\_tilt, fisheye\_camera\_centre\_offset\_x, fisheye\_camera\_centre\_offset\_y, and fisheye\_camera\_centre\_offset\_z are such that the active areas have parallel optical axes that are orthogonal to the line intersecting the camera centre points, and the camera corresponding to i equal to 0 is the right view.

– fisheye\_view\_dimension\_idc equal to 7 indicates that no additional constraints are implied for the syntax element values within the fisheye video information SEI message.

– Values of fisheye\_view\_dimension\_idc in the range of 3 to 6, inclusive, are reserved for future use by ITU-T | ISO/IEC. Decoders encountering a value of fisheye\_view\_dimension\_idc in the range of 3 to 6, inclusive, shall ignore it.

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

**fisheye\_num\_active\_areas\_minus1** plus 1 specifies the number of active areas in the coded picture. The value of fisheye\_num\_active\_areas\_minus1 shall be in the range of 0 to 3, inclusive. Values of fisheye\_num\_active\_areas\_minus1 greater than 3 are reserved for future use by ITU-T | ISO/IEC. Decoders encountering a fisheye video information SEI message with fisheye\_num\_active\_areas\_minus1 greater than 3 shall ignore the fisheye video information SEI message.

**fisheye\_circular\_region\_centre\_x**[ i ] and **fisheye\_circular\_region\_centre\_y**[ i ] specify the horizontal and vertical coordinates of the centre of the circular region that contains the i-th active area in the coded picture, respectively, in units of 2−16 luma samples. The value of fisheye\_circular\_region\_centre\_x[ i ] and fisheye\_circular\_region\_centre\_y[ i ] shall be in the range of 0 to 65 536 \* 216 − 1 (i.e., 4 294 967 295), inclusive.

**fisheye\_rect\_region\_top**[ i ], **fisheye\_rect\_region\_left**[ i ], **fisheye\_rect\_region\_width**[ i ], and **fisheye\_rect\_region\_height**[ i ] specify the coordinates of the top-left corner and the width and height of the i-th rectangular region that contains the i-th active area, in units of luma samples. The value of fisheye\_rect\_region\_top[ i ] shall be in the range of 0 to the height of the decoded picture minus 1 and the value of fisheye\_rect\_region\_height[ i ] shall be in the range of 0 to the height of the decoded picture minus 1, where the sum of fisheye\_rect\_region\_top[ i ] and fisheye\_rect\_region\_height[ i ] shall be smaller than the height of the decoded picture. The values of fisheye\_rect\_region\_left[ i ] shall be in the range of 0 to the width of the decoded picture minus 1 and the value of fisheye\_rect\_region\_width[ i ] shall be in the range of 0 to the width of the decoded picture minus 1 where the sum of fisheye\_rect\_region\_left[ i ] and fisheye\_rect\_region\_width[ i ] shall be smaller than the width of the decoded picture.

**fisheye\_circular\_region\_radius**[ i ] specifies the radius of the circular region that contains the i-th active area that is defined as a length from the centre of the circular region specified by fisheye\_circular\_region\_centre\_x[ i ] and fisheye\_circular\_region\_centre\_y[ i ] to the outermost pixel boundary of the circular region, in units of 2−16 luma samples, that corresponds to the maximum field of view of the i-th fisheye lens, specified by fisheye\_field\_of\_view[ i ]. The value of fisheye\_circular\_region\_radius[ i ] shall be in the range of 0 to 65 536 \* 216 − 1 (i.e., 4 294 967 295), inclusive.

The i-th active area is defined as the intersection of the i-th rectangular region, specified by fisheye\_rect\_region\_top[ i ], fisheye\_rect\_region\_left[ i ], fisheye\_rect\_region\_width[ i ], and fisheye\_rect\_region\_height[ i ], and the i-th circular region, specified by fisheye\_circular\_region\_centre\_x[ i ], fisheye\_circular\_region\_centre\_y[ i ], and fisheye\_circular\_region\_radius[ i ]. There shall not be overlapped regions between the i-th active area and the j-th active area when j is not equal to i.

[Ed. (JB): Is there any restriction between the relative sizes of the diameter of the circle and the length or width of the rect region? (AT): Can the specified parameters exceed the picture size? What happens when the top/left/width/height parameters, for example, exceed the picture size or overlap? Could I have regions with exactly the same region coordinates but have them described with different parameters? If yes, what should I do then? There is also no mention of the range of those values.]

**fisheye\_scene\_radius**[ i ] specifies the radius of a circular region within the i-th active area in units of 2−16 luma samples, where the obstruction, such as the camera body, is not included in the region specified by fisheye\_circular\_region\_centre\_x[ i ], fisheye\_circular\_region\_centre\_y[ i ], and fisheye\_scene\_radius[ i ]. The value of fisheye\_scene\_radius[ i ] shall be less than or equal to fisheye\_circular\_region\_radius[ i ], and shall be in the range of 0 to 65 536 \* 216 − 1 (i.e., 4 294 967 295), inclusive. The enclosed area is the suggested area for stitching as recommended by the encoder.

**fisheye\_camera\_centre\_azimuth**[ i ] and **fisheye\_camera\_centre\_elevation**[ i ] indicate the spherical coordinates that correspond to the centre of the circular region that contains the i-th active area in the cropped output picture, in units of 2−16 degrees. The value of fisheye\_camera\_centre\_azimuth[ i ] shall be in the range of −180 \* 216 (i.e., −11 796 480) to 180 \* 216 − 1 (i.e., 11 796 479), inclusive, and the value of fisheye\_camera\_centre\_elevation[ i ] shall be in the range of −90 \* 216 (i.e., −5 898 240) to 90 \* 216 (i.e., 5 898 240), inclusive.

**fisheye\_camera\_centre\_tilt**[ i ] indicates the tilt angle of the sphere region that correspond to the i-th active area of the cropped output picture, in units of 2−16 degrees. The value of fisheye\_camera\_centre\_tilt[ i ] shall be in the range of −180 \* 216 (i.e., −11 796 480) to 180 \* 216 − 1 (i.e., 11 796 479), inclusive.

[Ed. (JB): fisheye\_camera\_centre\_azimuth[ i ] and fisheye\_camera\_centre\_elevation[ i ] refer to the circular region center containing the active area, but fisheye\_camera\_centre\_tilt[ i ] refers to the active area. Is that intentional? (YK): I think some (rather significant) improvement is indeed needed here. From the definition of the term tilt angle, the tilt angle is defined as associated with a sphere region, not a region on a 2D picture, while in the context here the active area, regardless of its shape, is a region on a 2D picture.]

**fisheye\_camera\_centre\_offset\_x**[ i ], **fisheye\_camera\_centre\_offset\_y**[ i ] and **fisheye\_camera\_centre\_offset\_z**[ i ] indicate the XYZ offset values, in units of 2−16 millimeters, of the focal centre of the fisheye camera lens corresponding to the i-th active area from the focal centre origin of the overall fisheye camera configuration. The value of each of fisheye\_camera\_centre\_offset\_x[ i ], fisheye\_camera\_centre\_offset\_y[ i ], and fisheye\_camera\_centre\_offset\_z[ i ] shall be in the range of 0 to 65 536 \* 216 − 1 (i.e., 4 294 967 295), inclusive.

**fisheye\_field\_of\_view**[ i ] specifies the field of view of the lens which corresponds to the spherical domain coverage of the region on a sphere that corresponds to the i-th circular region, in units of 2−16 degrees. The value of fisheye\_field\_of\_view[ i ] shall be in the range of 0 to 360 \* 216, inclusive.

[Ed. (JB): fisheye\_field\_of\_view[ i ] isn’t separated by x and y dimensions and refers to active area. Is it really referring to the field of view of the circle rather than the active area? (YK): Good question. If just one component it needs at least say which of the two dimensions this is about. But why just one component anyway.]

**fisheye\_num\_polynomial\_coeffs**[ i ] specifies the number of polynomial coefficients for the circular region corresponding to the i-th active area. The value of fisheye\_num\_polynomial\_coeffs[ i ] shall be in the range of 0 to 8, inclusive. Values of fisheye\_num\_polynomial\_coeffs[ i ] greater than 8 are reserved for future use by ITU-T | ISO/IEC. Decoders encountering a fisheye video information SEI message with fisheye\_num\_polynomial\_coeffs[ i ] greater than 8 shall ignore the fisheye video information SEI message.

**fisheye\_polynomial\_coeff**[ i ][ j ] specifies the j-th polynomial coefficient value, in units of 2−24, of the curve function that maps the normalized distance of a luma sample from the centre of the circular region corresponding to the i-th active area to the angular value of a sphere coordinate from the normal vector of the i-th image plane. [Ed. (YK): What is an "image plane"? Check the need of clarifying this in the text.] The value of fisheye\_polynomial\_coeff[ i ][ j ] shall be in the range of −128 \* 224 (i.e., 2 147 483 648) to 128 \* 224 − 1 (i.e., 2 147 483 647), inclusive.

1. **References**
2. JCTVC-AD0023, “Omnidirectional fisheye video SEI message”, H.-M. Oh, S. Oh, Jan. 2018, Gwangju.
3. JCTVC-AD1005, “Additional Supplemental Enhancement Information for HEVC (Draft 1)”, J. Boyce, H.-M. Oh, G. J. Sullivan, A. Tourapis, Y.-K. Wang, Jan. 2018, Gwangju.
4. W17235, “Text of ISO/IEC FDIS 23090-2 Omnidirectional Media Format”, B. Choi, Y.-K. Wang, M. M. Hannuksela, Y. Lim, A. Murtaza, Oct. 2017, Macao.

# Patent rights declaration(s)

**LG Electronics Inc. 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).**