|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  30th Meeting: Gwangju, KR, 20–26 Jan. 2018 | Document: JCTVC-AD0023r2 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Omnidirectional fisheye video 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. | | |

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

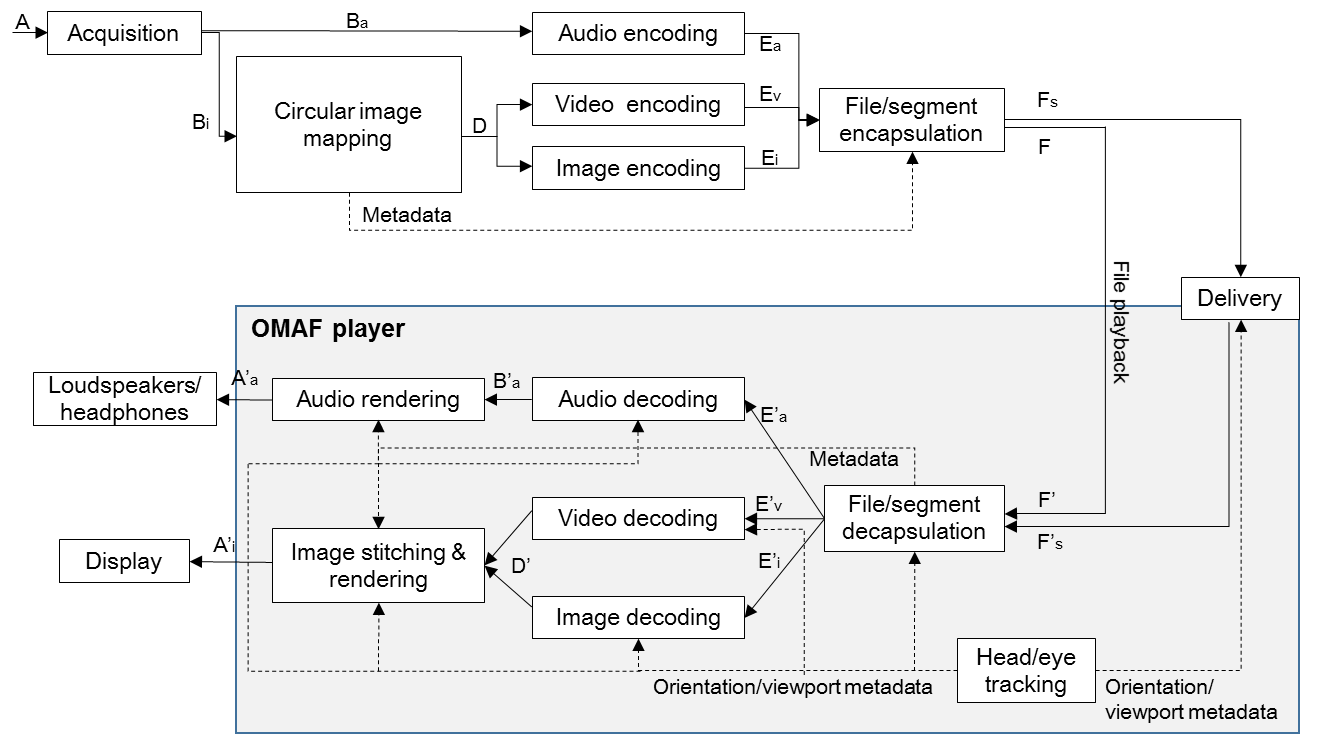
# Abstract

This contribution proposes to define an omnidirectional fisheye video SEI message that carries parameters used in the receiver to reconstruct an omnidirectional spherical video from circular fisheye camera outputs contained in a decoded picture. Based on the FisheyeVideoEssentialInfoStruct in OMAF FDIS, the essential fisheye video parameters, such as region information of circular images in the coded picture, field of view of fisheye lens, and fisheye camera parameters, are defined in the proposed SEI message.

The second revised contribution includes the conversion from a smaple location of a circular image to sphere mapping to resolve a comment addressed during the 30th JCT-VC meeting.

# Problem Statement

In omnidirectional video applications, two types of video formats are considered as inputs to video encoding: a projected picture and a fisheye circular picture. The projected picture is generated by stitching an omnidirectional camera output and then arranging the image on a sphere onto a two-dimensional picture, while the fisheye circular picture conveys raw omnidirectional camera outputs in a rectangular format. When the projected picture is delivered, the picture is decoded and rendered to display on the screen of head-mounted display or any other display devices. On the other hand, when the fisheye circular picture is delivered, additional steps, such as stitching, should be performed in the receiver devices as shown in Figure 1.



**Figure** 1 – **Content flow process for omnidirectional media with fisheye video [1]**

To support the 360 fisheye camera use case, omnidirectional fisheye video SEI message was proposed in the previous JCT-VC meetings [2][3]. In the 28th JCT-VC meeting, the OMAF DIS based SEI message [2] was reviewed and it was recommended to simplify the syntax elements. Based on the feedback from JCT-VC, the simplification of the parameters was discussed in OMAF. In the last OMAF meeting, it was agreed that some of the parameters were optional for receivers in stitching process, so that the fisheye-specific parameters were categorized into essential and supplementary parts, i.e., FisheyeVideoEssentialInfoStruct and FisheyeVideoSupplementalInfoStruct. Based on the result, OMAF made the following recommendation [4]: "Recommend JCT-VC to specify an SEI message to carry the essential fisheye video parameters.", where the essential parameters in OMAF FDIS [1] were

region information of circular images in the coded picture,

 field of view and camera parameters of fisheye lens, and

 lens distortion correction (LDC) parameters.

Following the definition in the FisheyeVideoEssentialInfoStruct in OMAF FDIS and the recommendation from OMAF, we propose to define omnidirectional fisheye video SEI message which contains essential parameters for omnidirectional video stitching as a receiver process.

Moreover, the precise equations for conversion from a sample location of a circular image to sphere mapping need to be specified. To achieve this, this contribution also describes the spherical coordinate mapping of a sample location of the circular image.

# Proposal

In this contribution, it is proposed to define an SEI message to carry on fisheye camera and image information that is necessary to construct a spherical video from a decoded picture. The syntax elements and its semantics are proposed based on the FisheyeVideoEssentialInfoStruct in OMAF FDIS, which defines region information of circular images in the coded picture, field of view and camera parameters of fisheye lens, and lens distortion correction parameters.

## Syntax

|  |  |
| --- | --- |
| omnidirectional\_fisheye\_video ( payloadSize ) { | **Descriptor** |
| **view\_dimension\_idc** | u(3) |
| reserved\_zero\_5bits | u(5) |
| **num\_circular\_images\_minus1** | u(8) |
| for( i = 0 ; i <= num\_circular\_images\_minus1; i++ ) { |  |
| **circular\_image\_center\_x**[ i ] | u(32) |
| **circular\_image\_center\_y**[ i ] | u(32) |
| **rect\_region\_top**[ i ] | u(32) |
| **rect\_region\_left**[ i ] | u(32) |
| **rect\_region\_width**[ i ] | u(32) |
| **rect\_region\_height**[ i ] | u(32) |
| **full\_radius**[ i ] | u(32) |
| **scene\_radius**[ i ] | u(32) |
| **camera\_center\_azimuth**[ i ] | i(32) |
| **camera\_center\_elevation**[ i ] | i(32) |
| **camera\_center\_tilt**[ i ] | i(32) |
| **camera\_center\_offset\_x**[ i ] | u(32) |
| **camera\_center\_offset\_y**[ i ] | u(32) |
| **camera\_center\_offset\_z**[ i ] | u(32) |
| **field\_of\_view**[ i ] | u(32) |
| **num\_polynomial\_coefs\_distortion**[ i ] | u(16) |
| for( j = 0 ;  j < num\_polynomial\_coefs\_distortion[ i ]; j++ ) |  |
| **polynomial\_coef\_k\_distortion**[ i ][ j ] | u(32) |
| } |  |
| } |  |

## Semantics

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

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

**view\_dimension\_idc** indicates alignment and viewing direction of fisheye lens.

– view\_dimension\_idc equal to 0 specifies that num\_circular\_images equal to 2, the values of camera\_center\_azimuth, camera\_center\_elevation, camera\_center\_tilt, camera\_center\_offset\_x, camera\_center\_offset\_y, and camera\_center\_offset\_z are such that the circular images have aligned optical axes and face opposite directions, and the sum of field\_of\_view values is greater than or equal to 360 \* 216.

– view\_dimension\_idc equal to 1 specifies that num\_circular\_images equal to 2, the values of camera\_center\_azimuth, camera\_center\_elevation, camera\_center\_tilt, camera\_center\_offset\_x, camera\_center\_offset\_y, and camera\_center\_offset\_z are such that the circular images have parallel optical axes that are orthogonal to the line intersecting the camera center points, and camera with i equal to 0 is the left view.

– view\_dimension\_idc equal to 2 specifies that num\_circular\_images equal to 2, the values of camera\_center\_azimuth, camera\_center\_elevation, camera\_center\_tilt, camera\_center\_offset\_x, camera\_center\_offset\_y, and camera\_center\_offset\_z are such that the circular images have parallel optical axes that are orthogonal to the line intersecting the camera center points, and camera with i equal to 0 is the right view.

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

– view\_dimension\_idc ranges from 3 to 6, inclusive, are reserved for future use.

**num\_circular\_images\_minus1** plus 1 specifies the number of circular images in the coded picture.

**circular\_image\_center\_x**[ i ] and **circular\_image\_center\_y**[ i ] specify the horizontal and vertical coordinate of the center of the i-th circular image in the coded picture, respectively, in units of 2−16 luma samples. The value of circular\_image\_center\_x[ i ] and circular\_image\_center\_y[ i ] shall be in the range of 0 to 65536 \* 216 – 1 (i.e., 4294967295), inclusive.

**rect\_region\_top**[ i ], **rect\_region\_left**[ i ], **rect\_region\_width**[ i ], and **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 fisheye circular image. These values are specified in units of luma samples.

**full\_radius**[ i ] specifies the radius of the i-th circular image that is defined as a length from the center of the i-th circular image represented by circular\_image\_center\_x[ i ] and circular\_image\_center\_y[ i ] to the outermost pixel boundary of the i-th circular image, in units of 2−16 luma samples, that corresponds to the maximum field of view of the i-th fisheye lens, specified by field\_of\_view[ i ]. The value of circular\_image\_radius[ i ] shall be in the range of 0 to 65536 \* 216 – 1 (i.e., 4294967295).

The actual pixel area of the i-th circular image is defined by the intersection of the rectangular region, defined by rect\_region\_top[ i ], rect\_region\_left[ i ], rect\_region\_width[ i ], and rect\_region\_height[ i ], and the circular region defined by image\_center\_x[ i ], image\_center\_y[ i ], and full\_radius[ i ].

**scene\_radius**[ i ] specifies the radius of a circular region within the i-th circular image in units of 2−16 luma samples, where the obstruction, such as the camera body, is not shown in the region defined by the circular\_image\_center\_x[ i ] and circular\_image\_center\_y[ i ] and scene\_radius[ i ]. The value of scene\_radius[ i ] shall be equal or smaller than full\_radius[ i ] and shall be in the range of 0 to 65536 \* 216 – 1 (i.e., 4294967295). The enclosed area is the suggested area for stitching as recommended by the content provider.

**camera\_center\_azimuth**[ i ] and **camera\_center\_elevation**[ i ] indicate the spherical coordinate that corresponds to the center of the i-th circular image in the cropped output picture, in units of 2−16 degrees. The value of camera\_center\_azimuth[ i ] shall be in the range of −180 \* 216 (i.e., −11796480) to 180 \* 216 − 1 (i.e., 11796479), inclusive, and the value of camera\_center\_elevation[ i ] shall be in the range of −90 \* 216 (i.e., −5898240) to 90 \* 216 (i.e., 5898240), inclusive.

**camera\_center\_tilt**[ i ] indicates the tilt angle of the i-th circular image of the cropped output picture, in units of 2−16 degrees. The value of camera\_center\_tilt[ i ] shall be in the range of −180 \* 216 (i.e., −11796480) to 180 \* 216 − 1 (i.e., 11796479), inclusive.

**camera\_center\_offset\_x**[ i ], **camera\_center\_offset\_y**[ i ] and **camera\_center\_offset\_z**[ i ] indicate the XYZ offset values, in units of 2−16 millimeters, of the focal center of the fisheye camera lens corresponding to the i-th circular image from the focal center origin of the overall fisheye camera configuration. The value of camera\_center\_offset\_x[ i ], camera\_center\_offset\_y[ i ], and camera\_center\_offset\_z[ i ] shall be in the range of 0 to 65536 \* 216 – 1 (i.e., 4294967295), respectively.

**field\_of\_view**[ i ] specifies the spherical domain coverage of the i-th circular image in the coded picture, in units of 2−16 degrees. The value of field\_of\_view[ i ] shall be in the range of 0 to 360 \* 216, inclusive.

**num\_polynomial\_coefs\_distortion**[ i ] specifies the number of polynomial coefficients corresponding to the i-th circular image.

**polynomial\_coef\_k\_distortion**[ i ][ j ] specifies the j-th polynomial coefficient value of the (i+1)-th curve function transformation of the radial distance between a luma sample in the i-th circular image and the center point of the circle, normalized by the full\_radius[ i ], to the angular value that is the angle between vectors corresponding to the luma sample and the center point both represented in the sphere coordinate whose origin corresponds to the focal point of the fisheye lens of the i-th circular image. The value of polynomial\_coef\_k\_distortion[ i ][ j ] shall be in the range of 0 to 256 \* 224 – 1 (i.e., 4294967295).

***D.X.X.X.X Conversion from a sample location of a circular image to sphere coordinates relative to the global coordinate axes***

For all circular images where i starts from 0 to num\_circular\_images\_minus1,

Input to this clause are:

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

– the centre location ( xc, yc ) of the i-th circular image given by circular\_image\_center\_x[ i ] and circular\_image\_center\_y[ i ], respectively, and full\_radius[ i ] ( rc ), all in units of 2−16 luma samples,

– field\_of\_view[ i ] ( θv ) in units of 2−16 degrees, and

– the rotation parameters ( αc, βc, γc ) where the values are given by camera\_center\_azimuth[ i ], camera\_center\_elevation[ i ], and camera\_center\_tilt[ i ], respectively, all in units of 2−16 degrees.

Output of this clause are:

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

The outputs are derived as follows:

ϕ' = ( Sqrt( ( x – xc \* 2-16 )2 + ( y – yc \* 2-16 )2 ) ÷ ( rc \* 2−16 ) ) \* ( θv \* 2−16 \* π ÷ 180 ) ÷ 2

θ' = Atan2( y − yc \* 2−16, x − xc \* 2−16 )

x1 = Cos( ϕ' )

y1 = Sin( ϕ' ) \* Cos( θ' )

z1 = Sin( ϕ' ) \* Sin( θ' )

α = ( αc \* 2−16 ) \* π ÷ 180

β = ( βc \* 2−16 ) \* π ÷ 180

γ = ( γc \* 2−16 ) \* π ÷ 180

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

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

1. **References**
2. 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.
3. JCTVC-AB0038, “Omnidirectional fisheye video SEI message”, H.-M. Oh, S. Oh, July 2017, Turin.
4. JCTVC-AC0034, “Omnidirectional fisheye video SEI message”, H.-M. Oh, S. Oh, Oct. 2017, Macao.
5. M41823, “MPEG#120 OMAF meeting agenda and minutes”, Y.-K. Wang, Y. Lim, Oct. 2017, Macao.

# Patent rights declaration(s)

**LG Electronics 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).**