|  |  |
| --- | --- |
| **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, 19–25 Oct. 2017 | Document: JCTVC-AC0034 |

|  |  |  |  |
| --- | --- | --- | --- |
| *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. | | |

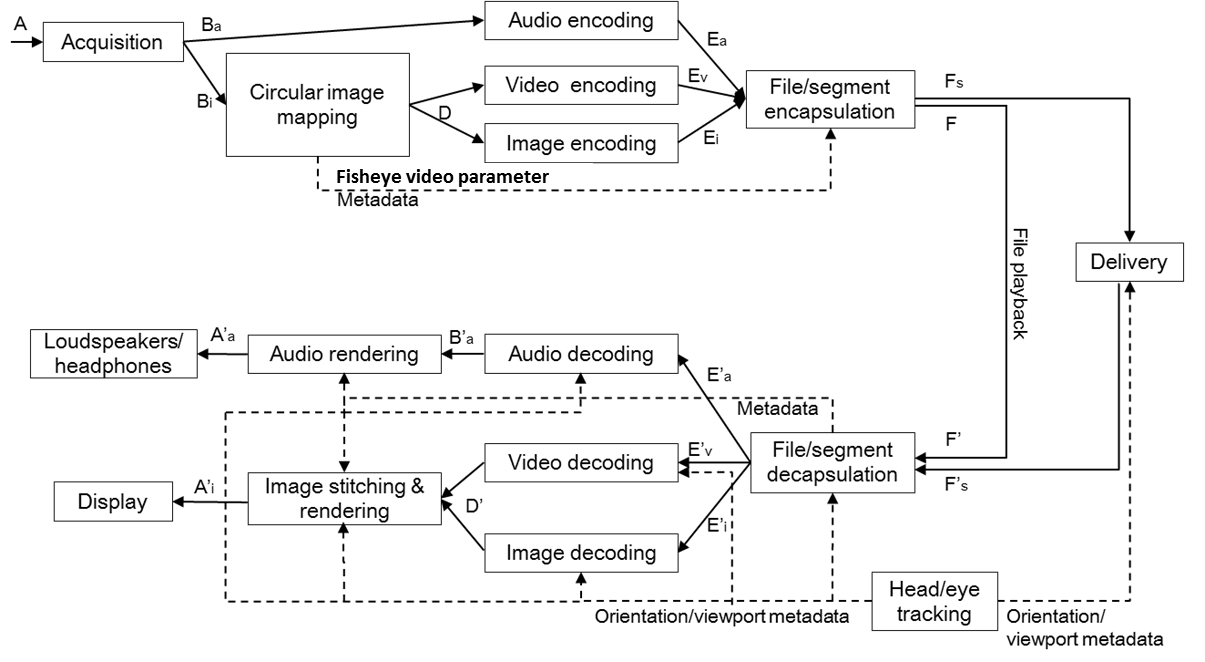
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# 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. With the proposed SEI message, the active pixel areas captured by the fisheye lens are indicated by the intersection of rectangular and circular regions. In aid of the field of view of the lens and the spherical coordinate matching information, the active pixels in each circular fisheye image could be mapped onto the 3D sphere by the post-decoder rendering process.

# 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 direct omnidirectional camera outputs in a rectangular format. When the projected picture is delivered, the picture is decoded and projected onto the screen of head-mounted display or any other display device according to the user viewport. Unlike this, when the fisheye circular picture is delivered, two additional steps are needed before playing on the display devices: stitching the circular images and then mapping it onto the sphere. In Figure 1, end-to-end content flow for fisheye video case is described.



**Figure** 1 – **Content flow process for omnidirectional media with fisheye video (modified from [1])**

To help the receivers to generate an omnidirectional video from the omnidirectional fisheye camera outputs, omnidirectional fisheye video SEI message based on the text elements in the OMAF DIS was proposed in the last JCT-VC meeting [2]. In the meeting, it was recommended to simplify the syntax elements and decided as further study. Based on the discussion, the essential elements that should be delivered to support post-decoder rendering process for fisheye video are proposed as follows.

* Indication of the omnidirectional fisheye video, which invokes the post-decoder rendering process.
* Description of actual pixel area in the decoded picture: intersection of a rectangular and a circular region.
* Spherical domain mapping information: a point on a sphere that matches to the center of the circular image and corresponding field of views.

In defining the region of active pixel area of each image, it is proposed to use the intersection of rectangular region, which represents the available pixel area of sensor, and the circular region, which represents the valid pixel area of lens. In Figure 2, three different cases of capturing an image from a sensor with respect to the focal distance. When focal distance is far relative to the size of the sensor, the full circular image could be captured on the sensor as shown in the first one. As the focal distance is getting closer, the cropping region becomes larger.

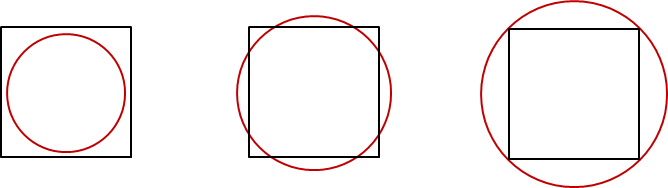


Figure 2 – Relationship between sensor pixel area and the circular image from the lens

In Figure 3, an example of actual picture is depicted which is captured by an omnidirectional camera which uses two fisheye lenses. The output picture contains two cropped circular images where the boundaries of the first image and the second image are met in the middle of the picture. In this case, the use of rectangular region with the circular region helps to prevent the receivers to include pixels from the other sensor output.



Figure 3 – Example of omnidirectional camera picture that contains two cropped circular images (unstitched 360 still picture [3])

# Proposal

In this contribution, we propose to define an SEI message to carry on information to indicate that the decoded picture is omnidirectional fisheye video and to inform receiver that the post-decoder rendering process is needed to construct the spherical video. According to the number of circular image produced from fisheye cameras, the detailed information with regard to the actual pixel area in the decoded picture and the projection on the spherical surface could be delivered.

## Syntax

|  |  |
| --- | --- |
| omnidirectional\_fisheye\_video ( payloadSize ) { | **Descriptor** |
| **num\_circular\_images\_minus1** | u(8) |
| for( i = 0 ; i <= num\_circular\_images\_minus1;  i++ ) { |  |
| **rect\_region\_top**[ i ] | u(16) |
| **rect\_region\_left**[ i ] | u(16) |
| **rect\_region\_width**[ i ] | u(16) |
| **rect\_region\_height**[ i ] | u(16) |
| **circular\_image\_center\_x**[ i ] | u(32) |
| **circular**\_**image\_center\_y**[ i ] | u(32) |
| **circular\_image\_radius**[ i ] | u(32) |
| **field\_of\_view**[ i ] | u(32) |
| **center\_azimuth**[ i ] | i(32) |
| **center\_elevation**[ i ] | i(32) |
| **center\_tilt**[ i ] | i(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 (usually two) 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.

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

**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.

**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.

**circular\_image\_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. The value of circular\_image\_radius[ i ] shall be in the range of 0 to 65536 \* 216 – 1 (i.e., 4294967295).

**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.

**center\_azimuth**[ i ] and **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 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 center\_elevation[ i ] shall be in the range of −90 \* 216 (i.e., −5898240) to 90 \* 216 (i.e., 5898240), inclusive.

**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 center\_tilt[ i ] shall be in the range of −180 \* 216 (i.e., −11796480) to 180 \* 216 − 1 (i.e., 11796479), inclusive.

For all circular images where i starts from 0 to num\_circular\_images\_minus1, the projection of a sample location (x, y) to spherical coordinate (ϕ, θ) is given as follow

– The active pixel area of the i-th fisheye image is determined 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 circular\_image\_center\_x[ i ], circular\_image\_center\_y[ i ], and circular\_image\_radius[ i ] in the units of luma sample, respectively.

– The projection for a sample location (x, y) in i-th active pixel area of fisheye camera is derived as follows where the output (ϕ, θ) represents the sample location in sphere coordinates in degrees relative to the global coordinate axes.

α = (+) ÷ (circular\_image\_radius[ i ] \* 2−16)) ÷ field\_of\_view[ i ] \* 2−16 \* π ÷ 180) ÷ 2

β = Atan2 ((y − circular\_image\_center\_y[ i ] \* 2−16), (x − circular\_image\_center\_x[ i ] \* 2−16))

Px = Sin α \* Cos β

Py = Cos α

Pz = Sin α \* Sin β

ϕ' = Atan2(Py, Px)

θ' = Atan2(Pz,)



ω′ = center\_tilt[ i ] \* 2−16 \* π ÷ 180

ϕ = Cos( ω′ ) \* ϕ' − Sin( ω′ ) \* θ' + center\_azimuth[ i ] \* 2−16

θ = Sin( ω′ ) \* ϕ' + Cos( ω′ ) \* θ'+ center\_elevation[ i ] \* 2−16

1. **References**
2. W16950, “Study of ISO/IEC DIS 23000-20 Omnidirectional Media Format”, B. Choi, Y.-K. Wang, M. M. Hannuksela, Y. Lim, A. Murtaza, July 2017, Turin.
3. JCTVC-AB0038, “Omnidirectional fisheye video SEI message”, H.-M. Oh, S. Oh, July 2017, Turin.
4. https://medium.com/@TIRIAS\_Research/360-degree-photos-and-video-will-move-social-media-d684958b1018

# 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).**