**Source: China Mobile**

**Title: [FS\_ Beyond2D] Scenario: Free-viewpoint TV**

**Document for: Agreement**

**Agenda Item: 9.9**

1 Introduction

This document contains a scenario related to the Study Item of Beyond 2D Video ( FS\_Beyond2D ) which is in document SP-240479.

2.Proposal

|  |
| --- |
| **1st Change** |

**1.Free-viewpoint TV**

**2.Motivation for the scenario**



Free-viewpoint TV (FTV) is a scenario that allows viewing of a 3D world by freely changing the viewpoint. It uses multiple high-precision cameras to shoot multi-angle images synchronously at the millisecond level to create a free viewpoint scene so that the user can freely move and watch from the selected view within the scene, including looking up and down, back and forth, left and right, i.e.6DoF (6-way freedom) experience. FTV has commercial value in sports events, TV programs, online education, advertising and marketing.

Multi-view video plus depth (MVD) is a source format of FTV which has limited viewpoints of video and associated depth maps. The depth maps and disparity maps associated with every view of the video that can be used to render arbitrary numbers of additional views via view synthesis to approximate the experience of free viewpoint. From the user’s perspective, they can choose the position and viewpoint arbitrarily and watch content of this view on autostereoscopic TVs[1] or traditional TVs with view control[2][3].

1. **Description of the scenario**



1. Capturing and processing: RGBD camera, color calibration, multi-stream synchronization, depth matching.
2. Encoding: texture map, depth map and camera parameters.
3. Packaging and delivery: HTTP adaptive streaming (e.g. DASH/HLS). Synthesis views or a view based on view control can be delivered to client.
4. Decoding: Synthesis views Rendering: rendering on client
5. General constraints on bandwidth: More views can tolerate better prediction error, but the bandwidth consumption will be greater. Therefore, the number of views should trade off prediction error and bandwidth.
6. **Supporting companies and 3GPP members**
7. **Source format properties**

For MVD source, the following information is considered to set input test sequences. It includes texture format properties and depth format properties of source as

|  |
| --- |
| Texture format properties |
| Spatial resolutions | .1920\*1080 |
| Chroma Format: | Y’CbCr |
| Chroma Subsampling: | 4:2:0, 4:2:2, 4:4:4 |
|  |  |
| Frame rates: | 25,30, 60 |
| Colour space formats: | BT.709,BT.2020, |
| Transfer Characteristics: |   |
| Bit depth | 10 |
| Viewpoints: | <tbd> |
| Camera parameters: | camera id, intrinsic parameters\*,extrinsic parameters |

|  |
| --- |
| Depth format properties |
| Spatial resolutions | 640\*480,1920\*1080 |
| Chroma Format: | <tbd> |
| Chroma Subsampling: | <tbd>  |
|  |  |
| Frame rates: | 25,30, 60 |
| Colour space formats: | <tbd> |
| Transfer Characteristics: | <tbd> |
| Bit depth | 8,16 |
| Viewpoints: | <tbd> |
| Camera parameters: | camera id, intrinsic parameters\*,extrinsic parameters |

The following information is considered to set output test sequences with synthesis views .

|  |
| --- |
| Synthesis multi-view format properties |
| Spatial resolutions | 3980\*2160, 7680\*4320 |
| Chroma Format: | Y’CbCr |
| Chroma Subsampling: | 4:2:0, 4:2:2, 4:4:4 |
| Aspect ratios | 16:9 |
| Frame rates: | 25,30, 60,90,120 |
| Colour space formats: | BT.709,BT.2020, |
| Transfer Characteristics: | <tbd> |
| Bit depth | 8,16 |
| Viewpoints: | <tbd> |

1. **Encoding and decoding constraints and settings**
2. **Performance Metrics and Requirements**
3. **Interoperability Considerations for the application**
4. **Test Sequences**
5. **Detailed test conditions**
6. **External Performance data**
7. **Additional Information**

|  |
| --- |
| **Change end** |

3 reference

[1] ThinkVision 27 3D <https://www.lenovo.com/us/en/p/accessories-and-software/monitors/professional/63f1uar3us?orgRef=https%253A%252F%252Fwww.bing.com%252F#tech_specs>

[2] Nguyen-Ha et al. Free-Viewpoint RGB-D Human Performance Capture and Rendering. In: Avidan, S., Brostow, G., Cissé, M., Farinella, G.M., Hassner, T. (eds) Computer Vision – ECCV 2022. ECCV 2022. Lecture Notes in Computer Science, vol 13676. Springer, Cham. <https://doi.org/10.1007/978-3-031-19787-1_27>

[3] P. Pérez et al., "Live Free-Viewpoint Video in Immersive Media Production Over 5G Networks," in IEEE Transactions on Broadcasting, vol. 68, no. 2, pp. 439-450, June 2022, doi: 10.1109/TBC.2022.3154612.