**3GPPSA4 131-bis-e MeetingS4-250596**

Online, 11 – 17 April 2025

**Source: China Mobile**

**Title: [FS\_Beyond2D] Clause 4.1 Introduction to Beyond 2D Video Formats**

**Agenda item: 9.6**

**Document for: Agreement**

**1. Introduction**

This document provides an introduction section to Clause 4, Beyond 2D Video Formats.

**2. Proposal**

It is proposed to agree the following changes to the 3GPP draft TR 26.956 V0.3.0

\* \* \* First Change \* \* \*

## 4.1 Introduction

This clause provides an overview of the Beyond 2D Video formats that have reached a certain amount of maturity as they can be generated from established and emerging capturing systems (including cameras for spatial video capturing) and can likely be rendered on existing display technologies (smartphones, VR HMDs, AR glasses, autostereoscopic and multiscopic displays). These formats include: stereoscopic 3D video, Multi-view plus Depth, dense dynamic point clouds and dynamic meshes. Emerging formats such as Neural Radiance Fields (NeRF), light fields, and 3D Gaussian Splatting (3DGS) are documented as formats under research. Table 4.1-1 summarizes the Beyond 2D Video formats documented in this study, highlighting their representation principles, advantages, challenges and compression technologies.

**Table 4.1-1 Summary of Beyond 2D Video Formats**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Name** | **Example** | **Definition** | **Codecs** | **Corresponding Section** |
| **Stereoscopic 3D and extensions** |  | A Stereoscopic View is defined as the perception of depth created by the brain's ability to fuse two slightly different images from each eye, based on the parallax difference between them. | Frame-packing and HEVC  MV-HEVC  ... | 4.3.2 |
| **Multi-view Plus Depth** |  | Multi-view video is a frame-based representation format whereby each frame of the video represents a still that can be viewed from any perspective within a viewing space that is informed by the provided camera positions. The representation optionally supports depth maps of same resolution. | MV-HEVC  MIV  ... | 4.3.4 |
| **Dense Dynamic Point Clouds** |  | A volumetric representation using 3D points with spatial coordinates and attributes (e.g., color, reflectance). Contains high-density point sets (>500K points/frame) enabling detailed, closed-surface rendering. | V-PCC  G-PCC  ... | 4.3.3 |
| **Dynamic Mesh** |  | A dynamic mesh is an object that represents a collection of vertices, edges and triangular faces (organized in polygons) defining the object's geometry that can be modified procedurally. | Draco  V-DMC  ... | 4.3.5 |
| **Light Fields** |  | A light field, or lightfield, is a vector function that describes the amount of light flowing in every direction through every point in a space | LVC  ... | 4.3.6.2 |
| **NeRF** |  | NeRF is the implicit representation of a 3D scene or object using a fully-connected (non-convolutional) deep network. | Under study | 4.3.6.1 |
| **3D Gaussian Splattings** |  | 3D Gaussian Splatting (3DGS), also referred as Gaussian Splatting Radiance Field, is an explicit radiance field based 3D representation that represents  3D scene or objects using a large number of discrete 3D anisotropic balls or particles, each defined by its spatial mean μ and covariance matric ∑. | Under study | 4.3.6.3 |

\* \* \* End of Changes \* \* \*