

附件 2：联盟标准制修订立项申请书

世界超高清视频产业联盟标准制、修订立项申请书

项目名称	支持六自由度交互的三维图像格式标准	制定 ☒ 修订 ☐	被修订标准号	
采用国际/内标准名称 (中文)	支持六自由度交互的三维图像格式标准			
申请单位	1. 华为技术有限公司 2. 中央广播电视总台 3. 咪咕文化科技有限公司 4. 行吟信息科技(上海)有限公司 5. 北京小鸟看看科技有限公司 6. 中国信息与通信研究院 7. 中国电子技术标准化研究院 8. 维沃移动通信有限公司 9. 凌云光技术股份有限公司 10. 寰宇信任(北京)技术有限公司 11. 杭州当虹科技股份有限公司 12. 北京中视广信科技有限公司	联系人	王志刚	
手机	18666221160	Email	wzg@huawei.com	
牵头单位	华为技术有限公司			
计划起止时间	2025 年 6 月-2025 年 12 月			
立项的目的、意义或必要性： 支持六自由度交互的三维图像是通过 3D 重建技术生成的物体/场景三维图像，作为照片的升级，这类三维图像让观察者以 6 自由度的视角看到场景细节，从而带来新的玩法与应用前景。 近年来以 3DGS 为代表的重建技术的发展使得普通消费者通过手机等轻量级终端低成本、快速、高质量的 3D 重建以及手机侧实时渲染浏览成为可能。然而当前产业界对于此类三维图像缺乏统一的交换格式，这样会导致在不同设备和应用上的 6 自由度三维图像文件不兼容，生态碎片化问题： (1) A 厂商的 6 自由度三维图像传输到 B 厂商的设备上时，可能无法识别，影响用户体验； (2) 各厂商封装格式不同，导致后续三方生态支持起来会比较复杂，标准不统一导致生态碎片化问题严重； 因此，本提案旨在建立一套通用的支持六自由度交互的三维图像格式标准，使得六自由度交互的三维图像文件在不同终端和不同应用之间可以交换分享，助力产业生态的进一步完善。				

适用范围或主要技术内容

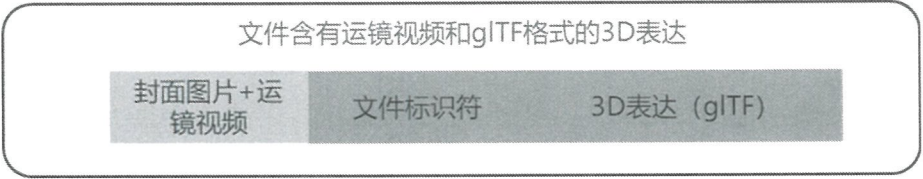
本文件规定了支持六自由度交互的三维图像格式标准，主要应用于以手机、PC、XR 眼镜等终端的六自由度交互的三维图像文件拍摄与重建、存储、编辑、分享使用场景，同时兼顾支持六自由度交互的三维图像文件在影视制作领域的二次创作需求。

从典型的用户使用终端拍摄到分享的全流程分析，该三维图像格式应该支持如下特性：

1. 支持预览，格式中存储预览图像
2. 支持存储 3DGS，mesh，点云等 3D 图元的原始表达以及压缩后的码流
3. 支持存储初始视角、最佳观测视角范围等 6DoF 浏览相关的参数
4. 支持对 3D 物体/场景添加标签、驱动、LOD、音频等扩展功能
5. 支持 HDR 6DoF 渲染
6. 与主流 APP 支持的媒体文件格式（如 HEIF、MP4 等）兼容，尽量复用当前流程以减少生态接入的开发成本
7. 支持导入 Blender 等主流 DCC 工具

为了满足以上需求，本提案提议利用 glTF 的现有生态和灵活的扩展能力，在其现有能力（对 mesh，点云等传统 3D 表达的支持，可存储视角等元数据）的基础上，新增扩展以实现特性[2~5,7]。通过将 glTF 封装进 MP4，以满足[1,6]。

支持六自由度交互的三维图像文件以MP4格式封装



在支持 3DGS 渲染的终端设备和应用上，用户可以通过读取 MP4 中封装的 glTF，以 6DOF 的方式浏览三维图像文件，在不支持 3DGS 渲染的设备和应用上以观看运镜视频的方式体验；

本标准包括支持如何在 MP4 中封装视频和 glTF、如何在 glTF 文件中扩展支持存储原始 3DGS、压缩后的 3DGS 码流以及六自由度交互需要的参数等内容。对于在三维图像格式中如何支持 HDR、空间音频、动画特效、标签等拓展特性以及是否将部分特性留作系列标准的后续内容，编写组将在编写过程中确定。

国内外情况简要说明

当前支持六自由度交互的三维图像文件没有现成可用的相关标准，与其有关联标准只能解决部分需求：



标准需求	PLY/SPZ/ (k)splate	glTF*	HEIF/JPE G	MP4	HEIF封装 glTF	MP4封装 glTF
生态兼容：兼容当前主流APP（社交媒体/即时通讯软件/电商APP等）支持的格式，使得3D动态图像格式可以在已有生态中无阻碍分享[1~3]	X	X	✓	✓	✓	✓
生态兼容：在不具备3D渲染能力的设备上可作为运镜视频观赏[1~3]	X	X	X(仅图片)	✓	X(仅图片)	✓
生态支持：格式中可存储或导出封面功能[1~3]	X	X	X	X	X	X
3D表达：支持3DGS	✓	X	X	X	X	X
3D表达和编辑：传统3D表达（如3D mesh）以及LOD、动画、音频、物理属性[1~3]	X	○	X	X	○	○
编辑：可对3D物体/场景添加标签[1,2]	X	X	X	X	X	X
6DoF交互浏览：含初始视角、最佳观测视角范围&运镜视角序列等6DoF浏览相关的参数[1,3]	X	○	X	X	○	○
6DoF交互浏览：应包含对音频（含空间音频）的支持[1~3]	X	○	X	○	○	○
6DoF交互浏览：支持HDR 6DoF渲染	X	○	X	X	○	○
传输：支持3DGS和3D mesh的编解码和传输[1~3]	X	○	X	X	○	○
版权：可通过元数据兼容针对3DGS和3D mesh的水印与加密的标准[1,2]	X	X	X	X	X	X

✓：满足需求 ○：部分满足 X：不满足需求

应用场景：[1]社交媒体 [2]电商 [3]即时通讯软件

*glTF 2.0 已成为ISO国际标准的glTF 2.0

第一类：支持存储 3DGS 的几种格式：

- (1) PLY: 学界多数重建工作选择以 PLY（Polygon File Format）格式作为 3DGS 的输出格式，多数工业界的重建软件也支持输出该格式。但 PLY 对 3DGS 的支持是通过自定义属性实现，同样是支持 3DGS 的 PLY，各方实现不完全一致，缺少标准化。且 PLY 缺少有效的压缩功能，大体积的文件增加了传输成本和用户等待时间。
- (2) SPZ: 由 Niantic 发布的 SPZ 格式通过量化和熵编码，实现对 3DGS 属性的轻量化压缩，有利于 3DGS 文件的存储和传输。缺点是压缩比偏低且不可调，难以控制编码的损失。
- (3) splate/ksplate: 可存储无损或是压缩后（基于量化的提供两档压缩，压缩比偏低）的 3DGS，主要用于基于 Three.js 的网页端 3DGS 渲染器中。与 Three.js 的结合使得其成为在网页端的 3DGS 应用的常用选择。

此外以上的几种格式存在共同的缺点：仅能用于存储 3D 表达，缺少对 6DoF 交体验的元数据的存储支持。同时以上几种格式也缺乏广泛的生态支持。

第二类：支持描述 3D 场景的格式：

- (1) glTF: 是一种开放且支持互操作的 3D 资产传输格式，关注运行效率，可部署在广泛的设备和平台上。glTF 将易于解析的 JSON 场景描述与一个或多个表示几何图形、动画和其他丰富数据的二进制资源相结合。这些二进制资源通常可直接加载到 GPU 缓冲区中，无需额外解析或处理，实现了高效传输和快速加载。glTF 支持存储点云/Mesh 等 3D 图元以及相机参数等各类 3D 体验的元数据，同时其具有良好的扩展机制。glTF2.0 已经是 ISO/IEC 国际标准，在 3D 领域有一定的生态支持。
- (2) USD: 同样可以描述 3D 场景，不同 glTF 关注传输和运行时效率，USD 更多关注制作域的需求。相比 glTF2.0 已经是 ISO/IEC 国际标准，USD 的标准化由 AOUSD（创始公司：PIXAR/NVIDIA/Adobe/Apple/Autodesk）掌控。

相对于 USD，关注传输和运行效率 glTF 更贴近本标准关注的应用场景。同时 AOUSD 和 Khronos 有 glTF 和 USD 的互转项目来保证互通。但当前两者都不原生支持 3DGS，需要扩展，且主流 APP 均不支持该类格式。

第三类：主流 APP 支持的媒体格式包括 HEIF/JPEG 等图片格式和以 MP4 为代表的视频格式。该类格式的可以被保存在手机的图库中并在各主流 APP 中被选中，上传，分享。但该类格式都不支持存储 3D 图元和相关元数据。



综上所述，现有格式无法满足应用诉求。为了满足可存储各类 3D 图元以及 6DoF 浏览所需的各类元数据的需求，同时降低生态接入的成本，应将第二类和第三类格式相结合。如上文所述，glTF 相比 USD 更适合本标准的应用场景；以视频格式封装相比以图片封装，在不支持 3DGS 渲染的设备上可以更好地向用户展现 6DoF 浏览的应有的效果，吸引用户在其他设备上进一步体验，同时 MP4 做为封装格式能力更强也便于体验的进一步演进。因此本标准选择将 glTF 封装进 MP4。	
申请立项单位意见	为技术有限公司 2025年6月23日 业务许可专用章 403071846096
同意发起立项申请。	
备注：	

1