**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC 1/SC 29/WG 2**

**MPEG TECHNICAL REQUIREMENTS**

**ISO/IEC JTC 1/SC 29/WG 2 m55222**

**Online – January 2021**

**Title: Annex C: Report of overview and Comparison of Scene Description Formats for Immersive Applications**

**Source: MPEG Requirements**

# Abstract

This contribution provides a copy of the publicly available Immersive Digital Experience Alliance (IDEA) whitepaper titled “An Overview of Scene Description Formats” released in September 2020 [1]. This whitepaper provides a high-level technical overview and comparison of five scene description formats currently employed in the contemporary computer-generated digital content and distribution workflow, as well as the incipient pipelines for immersive applications and displays. A comparative evaluation is accomplished by first defining then considering key parameters and capabilities required in a format representing a 3D scene.

# Introduction

Immersive content is authored using many digital content creation applications ultimately delivered or transmitted to end displays through game engines or the output of a render engine. To fully understand how a scene description fits into current content creation and distribution pipelines, as well as those that are emerging with immersive applications displays, we considered it necessary to provide a high-level technical overview and document the purpose, carriage capabilities, and functionality of the individual formats. This whitepaper considers the following parameters when evaluating the scene description formats: its purpose, the representation of virtual cameras, geometry, textures, materials, shading, volumes, animations, lighting, and audio, as well as the capability to compress, extend, and stream the format. The scene description formats considered and compared under these parameters are Filmbox (FBX), Alembic (ABC), GL Transmission Format (glTF), Universal Scene Description (USD), and Immersive Technologies Media Format (ITMF).

# Conclusions

A “better” scene description format is impractical to discern as each have different motivations, workflows, content types, and end use cases. The purpose of each scene description formats determines its overall design and scene hierarchical flow. They either follow a tree or directed acyclic graph data structure, both referred to as a “scene graph”. Mainly, these scene descriptions are motivated by interoperability between applications with few considering transmission to end user displays. FBX is widely adopted with its SDK enabling simple and consistent transport of 3D scene data. Alembic distills complex computations and rigging to an efficient geometric cache of flat scene data. USD enables a content creation workflow where artists can work on the same scene concurrently without loss of data. glTF as a publishing format allows for efficient transmission with minimal runtime processing in the last mile of content delivery. Finally, ITMF encloses all these scene descriptions as sub-formats in its package for interchange and distribution for immersive targeted 3D scenes.

Immersive applications and displays, many of these based on ray tracing techniques, add to the existing heterogenous pipelines and end client types alluding to the future necessity of a flexible, verbose format that addresses both interchange and distribution. By evaluating this prospective need, it recommends one that targets immersive media, is display-agnostic, and describes the greatest desirable representation that can be downscaled on an individual per end case basis.

# References

1. “An Overview of Scene Description Formats”, Public Release, available at [Immersive Digital Experiences Alliance](https://www.immersivealliance.org/download/download-idea-white-papers/), 30 September 2020.