

**ISO/IEC JTC 1/SC 29/WG 11 N19070**

**ISO/IEC JTC 1/SC 29/WG 11**

**Coding of moving pictures and audio**

**Convenorship: UNI (Italy)**

**Document type: Approved WG 11 document**

**Title: WD of Scene Description for MPEG Media**

**Status: Approved**

**Date of document: 2020-01-17**

**Source: Systems**

**Expected action: Approval by SC29**

**No. of pages: 5**

**Email of convenor: leonardo@chiariglione.org**

**Committee URL: mpeg.chiariglione.org**

**ISO 23090-14:2020(X)**

ISO TC JTC/SC 29/WG 11

Secretariat: JISC

**Information technology – Coded representation of immersive media - Part 14** **: Scene Description for MPEG Media**

WD stage

**Warning for WDs and CDs**

This document is not an ISO International Standard. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an International Standard.

Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

*To help you, this guide on writing standards was produced by the ISO/TMB and is available at* [*https://www.iso.org/iso/how-to-write-standards.pdf*](https://www.iso.org/iso/how-to-write-standards.pdf)

*A model manuscript of a draft International Standard (known as “The Rice Model”) is available at* [*https://www.iso.org/iso/model\_document-rice\_model.pdf*](https://www.iso.org/iso/model_document-rice_model.pdf)

© ISO 2020

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO’s member body in the country of the requester.

ISO copyright office

CP 401 • Ch. de Blandonnet 8

CH-1214 Vernier, Geneva

Phone: +41 22 749 01 11

Fax: +41 22 749 09 47

Email: copyright@iso.org

Website: www.iso.org

Published in Switzerland

Contents

[Foreword iv](#_Toc30035807)

[Introduction v](#_Toc30035808)

[1 Scope *(mandatory)* 1](#_Toc30035809)

[2 Normative references *(mandatory)* 1](#_Toc30035810)

[3 Terms and definitions 1](#_Toc30035811)

[4 Overview and Architecture 1](#_Toc30035812)

[4.1 Overview 1](#_Toc30035813)

[4.2 glTF 2.0 Overview and Extension Mechanisms 2](#_Toc30035814)

[4.3 Architecture 3](#_Toc30035815)

[4.4 Data Model 4](#_Toc30035816)

[5 Scene Description Extensions 4](#_Toc30035817)

[5.1 General Extensions 4](#_Toc30035818)

[5.2 Visual Extensions 4](#_Toc30035819)

[5.3 Audio Extensions 4](#_Toc30035820)

[6 Media Access API 4](#_Toc30035821)

[7 Processing Model 4](#_Toc30035822)

[Annex A (normative) Conformance and Validation 5](#_Toc30035823)

[A.1 General 5](#_Toc30035824)

[A.2 Conformance Point #1 5](#_Toc30035825)

[A.3 Conformance Point #2 5](#_Toc30035826)

[Bibliography 6](#_Toc30035827)

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see [www.iso.org/directives](https://www.iso.org/directives-and-policies.html)).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see [www.iso.org/patents](https://www.iso.org/iso-standards-and-patents.html)).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see [www.iso.org/iso/foreword.html](https://www.iso.org/foreword-supplementary-information.html).

This document was prepared by Technical Committee *[or Project Committee]* ISO/TC *[or ISO/PC]* ###, *[name of committee]*, Subcommittee SC ##, *[name of subcommittee]*.

This second/third/… edition cancels and replaces the first/second/… edition (ISO #####:####), which has been technically revised.

The main changes compared to the previous edition are as follows:

— xxx xxxxxxx xxx xxxx

A list of all parts in the ISO ##### series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user’s national standards body. A complete listing of these bodies can be found at [www.iso.org/members.html](https://www.iso.org/members.html).

Introduction

Type text.

*Identification of patent holders, if any.*

Information technology – Coded representation of immersive media - Part 14 : Scene Description for MPEG Media

# Scope

This part of the standard provides extensions to existing scene description formats in order to support MPEG media, in particular immersive media. MPEG media includes but is not limited to media encoded with MPEG codecs, media stored in MPEG containers, MPEG media and applications formats as well as media provided through MPEG delivery mechanisms. Extensions include scene description format syntax and semantics and the processing model when using these extensions in combination with a presentation engine. It also defines Media Access APIs for communication between the presentation engine and the Media Access Functions for these extensions. While the extensions defined in this part may be applicable to other scene description formats, a specific instantiation is provided for "The GL Transmission Format (glTF) 2.0" as defined by Khronos.

# Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

[1] Khronos Group, The GL Transmission Format (glTF) 2.0 Specification, <https://github.com/KhronosGroup/glTF/blob/master/specification/2.0/README.md>

# Terms and definitions

For the purposes of this document, the terms and definitions given in [1] and the following apply.

3.1

Presentation Engine

text of the definition

Note 1 to entry: Text of the note.

[SOURCE: …]

3.2

Media Access Function

text of the definition

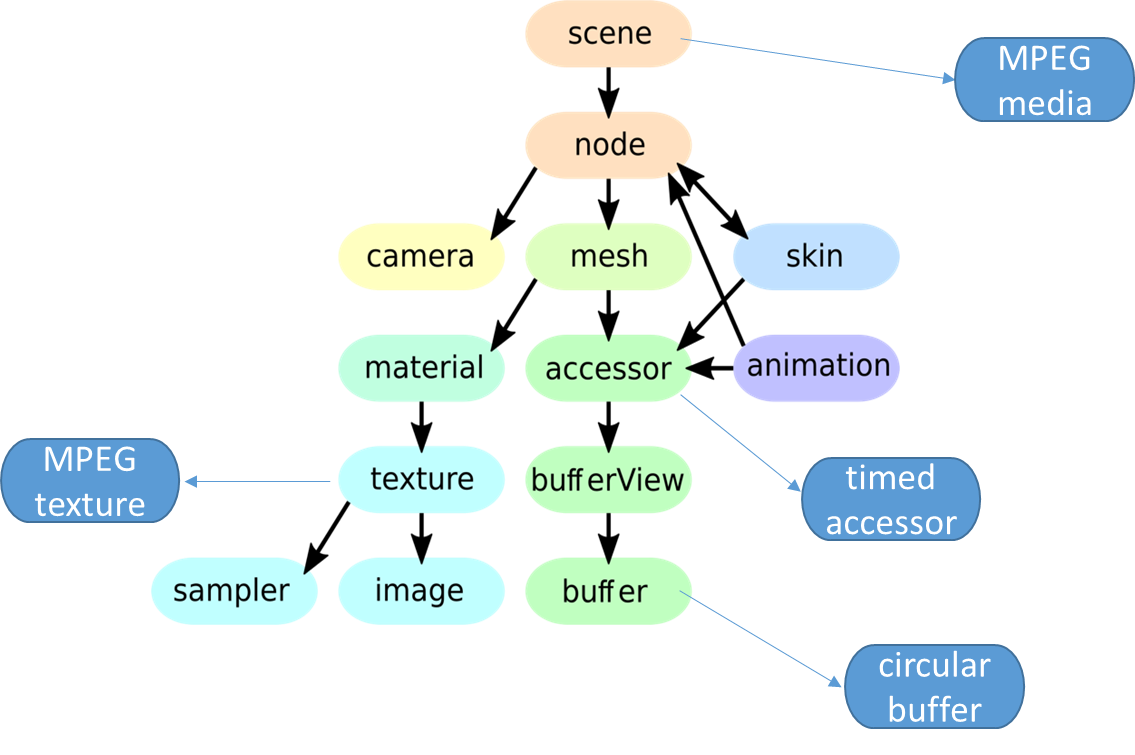
# Overview and Architecture

## Overview

Immersive media applications, for example those that aim to provide true AR and 6DoF experiences, require a scene description format that describes a rich 3D scene that enables physically-based rendering (PBR) of the audio-visual content.

Instead of specifying a new Scene Graph format for this purpose, this specification builds on the well-established glTF 2.0 [1] format that is standardized by the Khronos Group.

The following diagram depicts the glTF 2.0 format hierarchy and shows the extensions defined in this specification:



In addition to the extensions, which provide a tight integration of MPEG media with the Scene Description, the interface between the Presentation Engine and the Media Retrieval Engine is defined. Finally, a processing model as well as conformance and validation definitions of scene descriptions according to this specification are provided.

## glTF 2.0 Overview and Extension Mechanisms

glTF defines an extension mechanism [1] (section “Specifying Extensions”) that allows the base format to be extended with new capabilities. Any glTF object can have an optional extensions property that lists the extensions that are used by that object.

All extensions that are used in a glTF scene must be listed in the top-level extensionsUsed array object. Extensions that are required to correctly render the scene must also be listed in the extensionsRequired array.

As an example of possible extension, MPEG is currently defining a 6DoF audio encoder input format to address the MPEG-I requirements on 6DoF scene audio. glTF does not provide any support for audio scenes. To address this gap, a new node type and new material extension should be defined.

Similar to Javascript for HTML documents, an active processing may be supported in order to update a glTF scene description. This allows to update the description object model in an asynchronous manner (based on events such as interactivity or server events) as well as in a synchronous manner with a media source. In the latter case, a model as defined for Web Resource Track model for which updates are timed using an ISO BMFF track format aligned with ISO/IEC 29001-15 should be defined.

glTF 2.0 can be extended beyond the core specifications by basically 4 means:

* Vendor extensions:
* EXT Extensions:
* KHR Extensions
* KHX Extensions

This specification defines extensions to glTF under the vendor-specific extension framework with an MPEG namespace.

## Architecture

The scene description is consumed by a Presentation Engine to render a 3D scene to the viewer. The extensions defined in this specification allow for the creation of immersive experiences using MPEG media. The scene description extensions are designed with the goal of decoupling the Presentation Engine from the Media Retrieval Engine. Presentation Engine and Media Retrieval Engine communicate through the i-m interface, which allows the Presentation Engine to request media data required for the rendering of the scene. The Media Retrieval Engine will retrieve the requested media and make it available in a timely manner and in a format that can be immediately processed by the Presentation Engine. For instance, a requested media asset may be compressed and residing in the network, so the Media Retrieval Engine will retrieve and decode the asset and pass the resulting media data to the Presentation Engine for rendering. The media data is passed in form of buffers from the Media Retrieval Engine to the Presentation Engine. The requests for media data are passed through the Media Retrieval API from the Presentation Engine to the Media Retrieval Engine.

The following diagram depicts the reference architecture as described previously:



Figure 2: Scene Description Reference Architecture.

The interfaces and components in green are in scope for this specification.

The following assumptions apply:

* The format of the buffers is dictated by the scene description and is passed to the MAF through the Media Access API
* Decoder might need to perform necessary transformations to match the buffer format and layout declared in the scene description for that buffer
* The fetching of scene description and scene description updates may be triggered by the MAF.

NOTE: Upon definition of the scene update mechanism, the architecture might need to be adjusted.

## Data Model

# Scene Description Extensions

## General Extensions

## Visual Extensions

## Audio Extensions

# Media Access API

# Processing Model

1. (normative)  
     
   Conformance and Validation
   1. General
   2. Conformance Point #1
   3. Conformance Point #2

Bibliography

[1] ISO #####‑#, *General title — Part #: Title of part*

[2] ISO #####‑##:20##, *General title — Part ##: Title of part*