 ISO/IEC JTC 1/SC 29/WG 3 N01369

**ISO/IEC JTC 1/SC 29/WG 3**

**MPEG Systems   
Convenorship: KATS (Korea, Republic of)**

**Document type:** Output Document

**Title:** **Procedures for standard development, test scenarios and reference software for ISO/IEC 23090-14 (MPEG-I Scene Description)**

**Status:** Approved

**Date of document:** 2024-11-08

**Source:** ISO/IEC JTC 1/SC 29/WG 3

**Expected action:** ACT

**Action due date:** 2024-11-08

**No. of pages:** 17 (with cover page)

**Email of Convenor:** young.L@samsung.com

**Committee URL:** <https://isotc.iso.org/livelink/livelink/open/jtc1sc29wg3>

**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC 1/SC 29/WG 3**

**CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC 1/SC 29/WG 3 N** **1369**

**Kemer, TK – November 2024**

|  |  |
| --- | --- |
| **Source** | **WG03 (MPEG Systems)** |
| **Title** | **Procedures for standard development, test scenarios and reference software for ISO/IEC 23090-14 (MPEG-I Scene Description)** |
| **Editor** | **Thomas Stockhammer** |
| **MPEG number** | **24432** |

Contents

[1 Scope 4](#_Toc185248540)

[2 Time Plans and Projects 4](#_Toc185248541)

[3 Extending Khronos glTF2.0 6](#_Toc185248565)

[3.1 General 6](#_Toc185248566)

[3.2 Template for MPEG Extensions submitted to Khronos 6](#_Toc185248567)

[3.3 Status Extension Submission for first Edition 7](#_Toc185248568)

[3.4 Status Extension Submission for second Edition 8](#_Toc185248569)

[3.5 Process and Workflow 9](#_Toc185248570)

[3.6 Status 11](#_Toc185248571)

[4 Communication with Khronos 11](#_Toc185248588)

[4.1 Overview 11](#_Toc185248589)

[4.2 Communication prior to MPEG#146 12](#_Toc185248590)

[MPEG#133 12](#_Toc185248591)

[MPEG#135 12](#_Toc185248592)

[MPEG#136 12](#_Toc185248593)

[MPEG#137 12](#_Toc185248594)

[MPEG#138 12](#_Toc185248595)

[MPEG#139 12](#_Toc185248596)

[MPEG#140 13](#_Toc185248597)

[MPEG#141 13](#_Toc185248598)

[MPEG#142 13](#_Toc185248599)

[MPEG#143 14](#_Toc185248600)

[MPEG#144 14](#_Toc185248601)

[MPEG#145 14](#_Toc185248602)

[4.3 Communication from MPEG#146 14](#_Toc185248603)

[4.4 Communication from MPEG#147 14](#_Toc185248604)

[4.5 Communication from MPEG#148 15](#_Toc185248605)

[5 Requirements, Scenarios and Test Assets 15](#_Toc185248606)

[5.1 Requirements 15](#_Toc185248607)

[5.2 Scenarios 15](#_Toc185248608)

[5.3 Template for Test Scenario 15](#_Toc185248609)

[5.4 Continuous Call for Test Data 16](#_Toc185248610)

[5.5 Timeline 16](#_Toc185248611)

[5.6 Available Test Assets 16](#_Toc185248612)

[6 Roadmap 16](#_Toc185248613)

[6.1 Introduction 16](#_Toc185248614)

[6.2 Agreed Timeline 16](#_Toc185248615)

[6.3 Relevant Documents 17](#_Toc185248616)

[6.4 Topic 1: New and features enhancements to Ed 2. 17](#_Toc185248617)

[6.5 Topic 2: Support for scene understanding 17](#_Toc185248618)

[Overview of XR Spatial Computing of real environment 19](#_Toc185248619)

[Meta 19](#_Toc185248620)

[Space Setup 19](#_Toc185248621)

[SceneScript 20](#_Toc185248622)

[Google / ARCore 21](#_Toc185248623)

[Scene Semantics API 21](#_Toc185248624)

[Apple / ARKit 23](#_Toc185248625)

[RoomPlan API 23](#_Toc185248626)

[6.6 Topic 3: Immersive audio 25](#_Toc185248627)

[6.7 Topic 4: V-DMC support 25](#_Toc185248628)

[6.8 Topic 5 Support for Haptics Phase 2 26](#_Toc185248629)

[6.9 Topic 6: Multi-users support and avatar integration. 26](#_Toc185248630)

[7 Contributions for Extensions 27](#_Toc185248631)

[7.1 General 27](#_Toc185248632)

[7.2 Extension Principles 28](#_Toc185248633)

[8 Reference Software 28](#_Toc185248634)

[9 Gitlab Management 29](#_Toc185248635)

[10 Coordinators for Efforts until MPEG#149 29](#_Toc185248636)

# Scope

This document provides information and agreed processes in order to support the development of ISO/IEC 23090-14, "MPEG-I Scene Description" as well as ISO/IEC 23090-24, "Conformance and Reference Software for MPEG-I Scene Description".

# Time Plans and Projects

* ISO/IEC FDIS 23090-14:2023 Information technology — Coded representation of immersive media — Part 14: Scene Description for MPEG Media 🡺 published
  + <https://www.iso.org/standard/80900.html>
  + <https://sd.iso.org/projects/project/80900/overview>
  + published
* ISO/IEC 23090-24:2024 Information technology — Coded representation of immersive media — Part 24: Conformance and Reference Software for Scene Description for MPEG Media
  + Editors: Gurdeep Singh Bhullar
  + <https://www.iso.org/standard/83696.html>
  + <https://sd.iso.org/projects/project/83696/overview>
  + published
* ISO/IEC DIS 23090-14:2023/AMD 1:2023 Information technology — Coded representation of immersive media — Part 14: Scene description — Amendment 1: Support for immersive media codecs in scene description
  + <https://www.iso.org/standard/84769.html>
  + <https://sd.iso.org/projects/project/84769/overview>
  + published
* ISO/IEC DIS 23090-14/AMD 2 Information technology — Coded representation of immersive media — Part 14: Scene description — Amendment 2: Support for Haptics, Augmented Reality, Avatars, Interactivity, MPEG-I Audio and Lighting
  + https://www.iso.org/standard/86439.html
  + <https://sd.iso.org/projects/project/86439/overview>
  + Editors: Imed Bouazizi, Emmanuel Thomas, Patrice Hirtzlin
  + Project cancelled

A screenshot of a computer

Description automatically generated

* ISO/IEC DIS 23090-14:2024 Information technology — Coded representation of immersive media — Part 14: Scene description
  + Editors: Bouazizi Imed Dr, Hirtzlin Patrice M., Stockhammer Thomas Mr Dr.
  + https://www.iso.org/standard/90191.html
  + <https://sd.iso.org/projects/project/90191/overview>
  + Final text received or FDIS registered for formal approval (Mayumi, Youngkwon)

A screenshot of a computer

Description automatically generated

* ISO/IEC DIS 23090-14/AWI Amd 1: Information technology — Coded representation of immersive media — Part 14: Scene description — Amendment 1: Support of MPEG-I audio, scene understanding and other extensions
  + Editors: Bouazizi Imed Dr, Lelievre Sylvain M.
  + https://www.iso.org/standard/90213.html
  + <https://sd.iso.org/projects/project/90213/overview>
  + Project registered

A screenshot of a computer

Description automatically generated

* ISO/IEC 23090-24 Information technology — Coded representation of immersive media — Part 24: Conformance and reference software for scene description — Amendment 1: Conformance and reference software for scene description on haptics, augmented reality, avatars, interactivity and lighting
  + Editors: Imed Bouazizi, Gurdeep Singh Bhullar
  + https://www.iso.org/standard/87584.html
  + <https://sd.iso.org/projects/project/87584/overview>
  + Timeline extended

**A screenshot of a computer

Description automatically generated**

# Extending Khronos glTF2.0

## General

Based on the agreement during MPEG#128, MPEG-I Scene Description is developed as an extension to Khronos' glTF2.0 specification. This specification can be accessed here: <https://github.com/KhronosGroup/glTF/blob/master/specification/2.0/README.md>

According to the specification, glTF defines an extension mechanism that allows the base format to be extended with new capabilities. Any glTF object can have an optional extensions property. For details see <https://github.com/KhronosGroup/glTF/blob/master/specification/2.0/README.md#specifying-extensions>. For more information on glTF extensions, consult the [extensions registry specification](https://github.com/KhronosGroup/glTF/blob/master/extensions/README.md).

glTF supports different ways on extending the specification as documented here: <https://github.com/KhronosGroup/glTF/blob/master/extensions/README.md#promoting-extensions>

The following principles are agreed:

* MPEG develops extensions to Khronos glTF2.0 under the *Vendor Extensions framework*. Contributing companies should be aware of this. If contributions do not provide a statement that says otherwise, it is expected that the proponents agree to this.
* MPEG has been assigned an extension with the prefix MPEG <https://github.com/KhronosGroup/glTF/blob/master/extensions/Prefixes.md>. Contact person is the MPEG convenor, the JTC1 SC29 WG3 MPEG Systems chair as well as the chair of the MPEG-I Scene Description BOG. An update request has been submitted here: https://github.com/KhronosGroup/glTF/issues/2247
* If MPEG contributors are generally interested that their proposal may be considered as a KHR extension without any binding commitment, input contributions may state so. However, such a statement or the absence of such a statement does not impact the processing of a contribution in the context of the MPEG-I scene description work.

## Template for MPEG Extensions submitted to Khronos

It is proposed that all MPEG agreed extensions after DIS and FDIS has been issued, are added to the Khronos repository as follows

* Contributors
  + ISO/IEC SC29 WG3 (MPEG Systems) - Scene Description Breakout Group
  + Contacts: Thomas Stockhammer (MPEG-I Scene Description BoG Chair, tsto@qti.qualcomm.com)
* Status
  + Draft at DIS
  + Stable at FDIS
  + Based on [ISO/IEC FDIS 23090-14](https://www.iso.org/standard/80900.html)
* Dependencies
  + Written against the glTF 2.0 spec
* Overview:
  + Introduction should be provided on the extension
  + Pointer to ISO/IEC 23090-14 where the extension is defined
* glTF Schema Updates
  + Pointer to MPEG schema updates
* JSON Schema
  + Link to schema
* Known Implementation
  + Pointer to reference software: [ISO/IEC 23090-24](https://www.iso.org/standard/83696.html)
* Resources:
  + [ISO/IEC FDIS 23090-14](https://www.iso.org/standard/80900.html), Information technology — Coded representation of immersive media — Part 14: Scene Description
  + [ISO/IEC WD 23090-24](https://www.iso.org/standard/83696.html), Information technology — Coded representation of immersive media — Part 24: Conformance and Reference Software for Scene Description for MPEG Media
  + Comments may be submitted here: <https://github.com/MPEGGroup/Scene-Description>
* Best Practices:
  + Implementation Guidelines, Fallback mechanisms, etc.
* License:
  + Copyright (c), ISO/IEC
  + The use of the "MPEG scene description extensions" is subject to the license as accessible here: <https://standards.iso.org/> and is subject to the IPR policy as accessible here: <https://www.iso.org/iso-standards-and-patents.html>.

The extensions are also collected in the internal github here: here <http://mpegx.int-evry.fr/software/MPEG/Systems/SceneDescription/Specification/23090-14/-/tree/master/Extensions>.

## Status Extension Submission for first Edition

The pull request was finally merged on Feb 28, 2023

Khronos adds MPEG-I Scene Description Extensions to glTF2.0  
  
As chairman of the MPEG-I Scene Description group, it is my great pleasure to announce a milestone that we achieved in course of the collaboration with [The Khronos Group](https://www.linkedin.com/company/the-khronos-group/), in particular the 3D Formats working group: Khronos adopts the MPEG-I Scene Description extensions as defined in ISO/IEC 23090-14 as extensions to glTF2.0.  
  
This work happened within a long-lasting collaboration and dedication of many individuals, in particular to mention [Imed Bouazizi](https://www.linkedin.com/in/ACoAAABzII8BOGSlKHQJx0qsSleydHO4Jle_u9g), [Lukasz Kondrad](https://www.linkedin.com/in/ACoAAAA_W7MBW6HtQKZOIdZu4ULEFmu3sUi9KV0) [Yago Sanchez de la Fuente](https://www.linkedin.com/in/ACoAABbg1eYBz6ghlVhx02wl-M8UjdGaJ1rb2T8) [Ozgur Oyman](https://www.linkedin.com/in/ACoAAAAlWk8BoqVRVOMs6K9VmNP-L2WBMX7_7YQ) [Mary-Luc Champel](https://www.linkedin.com/in/ACoAAAAwDbEBviagkk33BWszQ4QoaM7aAb-7ihQ) Gurdeep Singh [Gaëlle Martin-Cocher](https://www.linkedin.com/in/ACoAAAAUBEUBKbCilTsq_kuS_2_6wxYyEkp6uWE) [Emmanuel Thomas](https://www.linkedin.com/in/ACoAAAJvsVcBFQvM_uLkeZX2_oN1i2hZlOIWky4) [Neil Trevett](https://www.linkedin.com/in/ACoAAAAdd-gBwt27NDRKFIyeasKdoIOrxKb8SXM) [Youngkwon Lim](https://www.linkedin.com/in/ACoAABeikS4BduV21BklETEPce2ccT18_ydnwrk) [Alexey Medvedev](https://www.linkedin.com/in/ACoAAAHPxh0BdrQofs-Yme26VI5QOGpQW9mGVy4) [Alexey Knyazev](https://www.linkedin.com/in/ACoAADWpIJoBljbJMD89dBLMPXIGS6zt6Rh745U) [Leonardo Chiariglione](https://www.linkedin.com/in/ACoAAAABm4wBfQbOVw0iZ0JuU4-i-m4E7-tiP0w). The extensions are documented here: https://github.com/KhronosGroup/glTF/blob/main/extensions/README.md   
  
These efforts are the foundation work in the [Metaverse Standards Forum](https://www.linkedin.com/company/metaverse-standards-forum/) and [3GPP](https://www.linkedin.com/company/3gpp/), and are only the starting point. More extensions to be expected. For details on the extensions and MPEG-I scene description, refer to our white paper [here](https://lnkd.in/eazc69Ba).



In addition, the extensions are added to the main extension page: <https://github.com/haudiobe/glTF/blob/main/extensions/README.md>

## Status Extension Submission for second Edition

A first set of extensions was created as output from MPEG#141 in

|  |  |  |  |
| --- | --- | --- | --- |
| [MDS22339](https://dms.mpeg.expert/doc_end_user/current_document.php?id=86339&id_meeting=193) | Draft registration of Khronos extensions 2nd edition | |  | | --- | | [MDS22339\_WG03\_N00815](https://dms.mpeg.expert/doc_end_user/documents/141_OnLine/wg11/MDS22339_WG03_N00815.zip) | |

First amendment (2  branches, 2 extensions)

* [MPEG\_primitive\_V3C](https://github.com/haudiobe/glTF/tree/MPEG_primitive_V3C/extensions/2.0/Vendor/MPEG_primitive_V3C)
* [MPEG\_sampler\_YCbCr](https://github.com/haudiobe/glTF/tree/MPEG_sampler_YCbCr/extensions/2.0/Vendor/MPEG_sampler_YCbCr)

Second amendment (5  branches, 7 extensions)

* [MPEG\_node\_avatar](https://github.com/haudiobe/glTF/tree/MPEG_avatar/extensions/2.0/Vendor/MPEG_node_avatar)
* <https://github.com/haudiobe/glTF/tree/MPEG_haptic/extensions/2.0/Vendor>
  + [MPEG\_haptic](https://github.com/haudiobe/glTF/tree/MPEG_haptic/extensions/2.0/Vendor/MPEG_haptic)
  + [MPEG\_material\_haptic](https://github.com/haudiobe/glTF/tree/MPEG_haptic/extensions/2.0/Vendor/MPEG_material_haptic)
* [MPEG\_lights\_texture\_based](https://github.com/haudiobe/glTF/tree/MPEG_lights_texture_based/extensions/2.0/Vendor/MPEG_lights_texture_based)
* <https://github.com/haudiobe/glTF/tree/MPEG_interactivity/extensions/2.0/Vendor>
  + [MPEG\_scene\_interactivity](https://github.com/haudiobe/glTF/tree/MPEG_interactivity/extensions/2.0/Vendor/MPEG_scene_interactivity)
  + [MPEG\_node\_interactivity](https://github.com/haudiobe/glTF/tree/MPEG_interactivity/extensions/2.0/Vendor/MPEG_node_interactivity)
* <https://github.com/haudiobe/glTF/tree/MPEG_anchor/extensions/2.0/Vendor>
  + [MPEG\_scene\_anchor](https://github.com/haudiobe/glTF/tree/MPEG_anchor/extensions/2.0/Vendor/MPEG_scene_anchor)
  + [MPEG\_node\_anchor](https://github.com/haudiobe/glTF/tree/MPEG_anchor/extensions/2.0/Vendor/MPEG_node_anchor)

However, we identified, that a workflow through a non-private GitHub repository and an internal approval process is preferable in acting towards Khronos.

## Process and Workflow

For the workflow, the following aspects should be considered:

1. Every feature in MPEG-I SD creates its own pull request
2. A feature may consist of multiple extensions
3. The extensions should be submitted as part of addition of the technology to the standard to MPEG systems
4. A public repository in MPEG is used to host the mirror, but also some extensions that are not yet approved by Khronos. This GitHub repository can also be used by the public to provide comments
5. Care should be taken on keeping consistency with what is added to the standard and also to preliminary drafts sent for ballot
6. We also maintain a repository internally that needs to be taken care of
7. A timely visibility of the extensions to Khronos and general public is important.
8. It should not contradict ISO rules and policies.

A high-level workflow is shown in Figure 1.

A diagram of a diagram

Description automatically generated

Figure 1 High-level workflow

The following detailed workflow implementation was proposed

* Initial set up (only happens once in the course of developing the MPEG-I SD standard):
  + A fork of the Khronos glTF repository on GitHub is created under the MPEGGroup account on GitHub. This should include all 1st edition extensions.
  + The glTF repository fork under MPEGGroup on GitHub is cloned as a new repository on the MPEG GitLab repository under the Scene Description project.
* From that point onwards, the MPEG/extensions GitLab repository is the repository where all the updates are collected from the SD BoG decisions.
* When work on a new feature commences in MPEG-I SD, a new branch in the MPEG/extensions GitLab repository is created for the new feature and updates are made to that internal branch as modifications are agreed by the group.
* Once the document to which the new feature belongs (e.g., an amendment or a new edition) reaches CD stage and a ballot is to be initiated, the following must be done:
  + the MPEG/extensions GitLab branch is tagged with the edition number and the ISO stage
  + the MPEG/glTF GitLab branch is tagged with the edition number and the ISO stage
* When the document reaches DIS stage and a ballot is initiated, the following must be done:
  + the MPEG/extensions and MPEG/glTF GitLab branches are tagged with the edition number and the ISO stage. The following tag is proposed
    - iso\_number|ned|iso\_document
    - Examples
      * 23090-14|1ed|CD
      * 23090-14|2ed|DAmd 2
  + the MPEG/extensions and MPEG/glTF GitLab branches are pushed to the corresponding MPEGGroup repository on GitHub
    - executed manually by somebody from a local repo with both remote endpoint GitLab and GitHub
    - we need a responsible person. A script may be created
  + a pull request from the MPEGGroup/glTF is created against the Khronos/glTF GitHub repository to start soliciting feedback and comments from the DIS ballot
    - a draft at DIS stage and later changed to a final pull request at FDIS
  + inform Khronos of the existence of these draft extensions in an LS
* Any feedback or comments on the pull request created on the Khronos GitHub repository that the group agrees is useful and should be captured by a national body (NB) comment on the ballot.
* Other feedback may also be received from the MPEGGroup/extensions which should also be addressed via MPEG input contributions and/or NB comments.
* When the document reaches FDIS stage and a ballot is initiated, the following must be done:
  + the MPEG/extensions and MPEG/glTF GitLab branches are tagged with the edition number and the ISO stage
  + The following tag is proposed
    - iso\_number|ned|iso\_document
    - Examples
      * 23090-14|1ed|CD
      * 23090-14|2ed|DAmd 2
  + the MPEG/extensions and MPEG/glTF GitLab branches are pushed to the corresponding MPEGGroup repository on GitHub
    - executed manually by somebody from a local repo with both remote endpoint GitLab and GitHub
    - we need a responsible person. A script may be created
* When the pull request on the Khronos GiHub repository is accepted and merged, the master (main) branches on both the MPEGGroup GitHub repository and the internal MPEG GitLab repositories should by synched with the Khronos GitHub repository.

Alternative workflow (update to above workflow):

* When DIS is issued

1. Create fork on MPEG GitHub repository of Khronos glTF repository, and take the extensions from the MPEG GitLab repository and add them "manually"
2. Create a pull request to Khronos as “draft”
3. All updates to the extensions are done on the fork on MPEG GitHub repository
4. When we release a new version of the standard (e.g., improvement of DIS, FDIS, etc.), we create a clone of the public MPEG GitHub repository and add it to the internal MPEG GitLab in order to maintain spec consistency

## Status

In previous meetings, we describe scripts that were used to automate the generation of the README files for the new MPEG extensions to glTF 2.0. Unfortunately, these scripts did not support the generation of the companion JSON schemas.

We had to generate the JSON schemas manually as it turned out to be quite hard to generate them automatically from the description. The effort to automate this process will still continue in the future.

The following extensions are now considered complete with their JSON schemas and we want to invite reviewers prior to submitting to Khronos:

* MPEG\_anchor: <https://github.com/MPEGGroup/Scene-Description/tree/main/MPEG_anchor>
* MPEG\_interactivity: <https://github.com/MPEGGroup/Scene-Description/tree/main/MPEG_interactivity>
* MPEG\_lighting: <https://github.com/MPEGGroup/Scene-Description/tree/main/MPEG_lighting>
* MPEG\_node\_avatar: <https://github.com/MPEGGroup/Scene-Description/tree/main/MPEG_node_avatar>

Note that the extensions in the annexes are not created yet. We consider these to be related to MPEG codecs and not necessarily of interest to Khronos. However, we still plan to generate them for hosting on the public GitHub.

We invite the experts to contribute to this process by reviewing the documentation and schemas to prepare for the submission to Khronos.

# Communication with Khronos

## Overview

Khronos has active work in the context of glTF2.0, see the KHR extensions under development here: <https://github.com/KhronosGroup/glTF/blob/master/extensions/README.md>. It is also identified that there is an overlap between MPEG members and glTF participants. Khronos and graphics experts meet in Khronos meetings, but also at developer and research conferences such as GDC and Siggraph. For meetings, please refer to <https://www.khronos.org/events/>.

Khronos Member Meetings occur 3 times per year and offer the opportunity for Khronos members to come together in a face-to-face environment to discuss technical work, industry feedback, network with colleagues and have some fun.

## Communication prior to MPEG#146

### MPEG#133

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| [MDS20159](https://dms.mpeg.expert/doc_end_user/current_document.php?id=78184&id_meeting=185) | WG 03 | 00180 | All | Liaison to Khronos on Scene Description for MPEG Media | WG 03 MPEG Systems | [MDS20159\_WG03\_N00180](https://dms.mpeg.expert/doc_end_user/documents/133_OnLine/wg11/MDS20159_WG03_N00180.zip) |

### MPEG#135

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| [MDS20563](https://dms.mpeg.expert/doc_end_user/current_document.php?id=79965&id_meeting=187) | WG 03 | 00309 | WG 03 All | Liaison statement to Khronos on MPEG-I Scene Description | WG 03 MPEG Systems | [MDS20563\_WG03\_N00309](https://dms.mpeg.expert/doc_end_user/documents/135_OnLine/wg11/MDS20563_WG03_N00309.zip) |

### MPEG#136

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| [MDS21056](https://dms.mpeg.expert/doc_end_user/current_document.php?id=81135&id_meeting=188) | WG 03 | 00434 | WG 03 All | Liaison statement to Khronos on MPEG-I Scene Description | WG 03 MPEG Systems | [MDS21056\_WG03\_N00434](https://dms.mpeg.expert/doc_end_user/documents/136_OnLine/wg11/MDS21056_WG03_N00434.zip) |

### MPEG#137

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| [MDS21327](https://dms.mpeg.expert/doc_end_user/current_document.php?id=82177&id_meeting=189) | WG 03 | 00511 | WG 03 All | Liaison statement to Khronos on MPEG-I Scene Description | WG 03 MPEG Systems | [MDS21327\_WG03\_N00511](https://dms.mpeg.expert/doc_end_user/documents/137_OnLine/wg11/MDS21327_WG03_N00511.zip) |

### MPEG#138

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| [MDS21435](https://dms.mpeg.expert/doc_end_user/current_document.php?id=82962&id_meeting=190) | WG 03 | 00542 | WG 03 MPEG-I | Registration of Khronos extensions | WG 03 MPEG Systems | [MDS21435\_WG03\_N00542](https://dms.mpeg.expert/doc_end_user/documents/138_OnLine/wg11/MDS21435_WG03_N00542.zip) |
| [MDS21603](https://dms.mpeg.expert/doc_end_user/current_document.php?id=83120&id_meeting=190) | WG 03 | 00588 | WG 03 All | Liaison statement to Khronos on mesh attributes in glTF 2.0 | WG 03 MPEG Systems | [MDS21603\_WG03\_N00588](https://dms.mpeg.expert/doc_end_user/documents/138_OnLine/wg11/MDS21603_WG03_N00588.zip) |

### MPEG#139

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| [MDS21744](https://dms.mpeg.expert/doc_end_user/current_document.php?id=83961&id_meeting=191) | WG 03 | 00615 | WG 03 MPEG-I | Registration of Khronos extensions | WG 03 MPEG Systems | [MDS21744\_WG03\_N00615](https://dms.mpeg.expert/doc_end_user/documents/139_OnLine/wg11/MDS21744_WG03_N00615.zip) |
| [MDS21885](https://dms.mpeg.expert/doc_end_user/current_document.php?id=84100&id_meeting=191) | WG 03 | 00671 | WG 03 All | Liaison statement to Khronos on MPEG-I Scene description | WG 03 MPEG Systems | [MDS21885\_WG03\_N00671](https://dms.mpeg.expert/doc_end_user/documents/139_OnLine/wg11/MDS21885_WG03_N00671.zip) |

### MPEG#140

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| [MDS21968](https://dms.mpeg.expert/doc_end_user/current_document.php?id=85069&id_meeting=192) | WG 03 | 00689 | 2022-10-28 15:29:21 | 2023-01-15 18:33:04 | WG 03 MPEG-I | Final registration of Khronos extensions for 1st edition | WG 03 MPEG Systems | [MDS21968\_WG03\_N00689](https://dms.mpeg.expert/doc_end_user/documents/140_Mainz/wg11/MDS21968_WG03_N00689.zip) |
| [MDS22198](https://dms.mpeg.expert/doc_end_user/current_document.php?id=85289&id_meeting=192) | WG 03 | 00751 | 2022-10-28 15:29:21 | 2023-01-15 18:37:40 | WG 03 MPEG-I | Draft registration of Khronos extensions 2nd edition | WG 03 MPEG Systems | [MDS22198\_WG03\_N00751](https://dms.mpeg.expert/doc_end_user/documents/140_Mainz/wg11/MDS22198_WG03_N00751.zip) |
| [MDS22200](https://dms.mpeg.expert/doc_end_user/current_document.php?id=85291&id_meeting=192) | WG 03 | 00753 | 2022-10-28 15:29:21 | 2023-01-15 18:38:56 | WG 03 All | Liaison to Khronos on MPEG-I Scene Description | WG 03 MPEG Systems | [MDS22200\_WG03\_N00753](https://dms.mpeg.expert/doc_end_user/documents/140_Mainz/wg11/MDS22200_WG03_N00753.zip) |

### MPEG#141

|  |
| --- |
|  |
| [MDS Number](https://dms.mpeg.expert/doc_end_user/current_meeting.php?id_meeting=193&type_order=0&sql_type=document_number) | [Group](https://dms.mpeg.expert/doc_end_user/current_meeting.php?id_meeting=193&type_order=0&sql_type=document.id_sub_group) | [G-Number](https://dms.mpeg.expert/doc_end_user/current_meeting.php?id_meeting=193&type_order=0&sql_type=document_gnumber) | [Created](https://dms.mpeg.expert/doc_end_user/current_meeting.php?id_meeting=193&type_order=0&sql_type=document_date_time) | [Uploaded](https://dms.mpeg.expert/doc_end_user/current_meeting.php?id_meeting=193&type_order=0&sql_type=upload_document_date_time) | [Standard](https://dms.mpeg.expert/doc_end_user/current_meeting.php?id_meeting=193&type_order=0&sql_type=document.id_group) | [Title](https://dms.mpeg.expert/doc_end_user/current_meeting.php?id_meeting=193&type_order=0&sql_type=title) | [Source](https://dms.mpeg.expert/doc_end_user/current_meeting.php?id_meeting=193&type_order=0&sql_type=authors) | |  |  | | --- | --- | | [Download](javascript:%20submitform()) |  | |
| [MDS22293](https://dms.mpeg.expert/doc_end_user/current_document.php?id=86293&id_meeting=193) | WG 03 | 00769 | 2023-01-21 14:27:20 | 2023-03-20 18:01:49 | WG 03 MPEG-I | Revised registration of Khronos extensions for 1st edition | WG 03 MPEG Systems | |  |  | | --- | --- | | [MDS22293\_WG03\_N00769](https://dms.mpeg.expert/doc_end_user/documents/141_OnLine/wg11/MDS22293_WG03_N00769.zip) |  | |
| [MDS22339](https://dms.mpeg.expert/doc_end_user/current_document.php?id=86339&id_meeting=193) | WG 03 | 00815 | 2023-01-21 14:27:20 | 2023-04-24 10:03:30 | WG 03 MPEG-I | Draft registration of Khronos extensions 2nd edition | WG 03 MPEG Systems | |  | | --- | | [MDS22339\_WG03\_N00815](https://dms.mpeg.expert/doc_end_user/documents/141_OnLine/wg11/MDS22339_WG03_N00815.zip) | |

### MPEG#142

|  |
| --- |
|  |
| [MDS22619](https://dms.mpeg.expert/doc_end_user/current_document.php?id=87765&id_meeting=194) | WG 03 | 00877 | 2023-04-29 09:06:31 |  | WG 03 MPEG-I | Draft registration of Khronos extensions 2nd edition | WG 03 MPEG Systems | |  |  | | --- | --- | | MDS22619\_WG03\_N00877 | A white paper with a black line  Description automatically generated | |
| [MDS22647](https://dms.mpeg.expert/doc_end_user/current_document.php?id=87793&id_meeting=194) | WG 03 | 00904 | 2023-04-29 09:06:31 | 2023-07-16 10:41:32 | WG 03 MPEG-I | Proposed Khronos blog post on MPEG-I Scene Description | WG 03 MPEG Systems | |  |  | | --- | --- | | [MDS22647\_WG03\_N00904](https://dms.mpeg.expert/doc_end_user/documents/142_Antalya/wg11/MDS22647_WG03_N00904.zip) |  | |
| [MDS22679](https://dms.mpeg.expert/doc_end_user/current_document.php?id=87825&id_meeting=194) | WG 03 | 00918 | 2023-04-29 09:06:31 | 2023-05-11 17:15:03 | WG 03 All | Liaison statement to Khronos on MPEG-I Scene Description Status update | WG 03 MPEG Systems | |  | | --- | | [MDS22679\_WG03\_N00918](https://dms.mpeg.expert/doc_end_user/documents/142_Antalya/wg11/MDS22679_WG03_N00918.zip) | |

### MPEG#143

No communication, but MPEG was invited by Khronos to a meetup that was finally scheduled on July 13, 2023.

Details of the event are here: <https://www.khronos.org/events/gltf-meetup-July2023>

* Public Slides in pdf: [glTF 2.0 Extensions in MPEG and 3GPP - Real time exchange formats for 3D Experiences](https://www.khronos.org/developers/linkto/gltf-2.0-extensions-in-mpeg-and-3gpp-real-time-exchange-formats-for-3d-experiences)
* Video recording: [glTF 2.0 Extensions in MPEG and 3GPP - Real time exchange formats for 3D Experiences](https://www.khronos.org/developers/linkto/gltf-2.0-extensions-in-mpeg-and-3gpp-real-time-exchange-formats-for-3d-experiences-vid)

In addition, Khronos invited to use the presentation and the transcript of the webinar to transfer this into a blog. This effort will happen over the next few weeks and we will share the draft with MPEG colleagues for comments.

### MPEG#144

|  |
| --- |
|  |
| [MDS23188](https://dms.mpeg.expert/doc_end_user/current_document.php?id=90514&id_meeting=196) | WG 03 | 01035 | 2023-10-20 22:21:23 |  | WG 03 MPEG-I | Draft registration of Khronos extensions 2nd edition | WG 03 MPEG Systems | |  |  | | --- | --- | | MDS23188\_WG03\_N01035 | A white paper with a black line  Description automatically generated | |
| [MDS23221](https://dms.mpeg.expert/doc_end_user/current_document.php?id=90547&id_meeting=196) | WG 03 | 01068 | 2023-10-20 22:21:23 |  | WG 03 MPEG-I | Proposed Khronos blog post on MPEG-I Scene Description | WG 03 MPEG Systems | |  | | --- | | MDS23221\_WG03\_N01068 | |

### MPEG#145

No communication from MPEG#145

## Communication from MPEG#146

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| [MDS23811](https://dms.mpeg.expert/doc_end_user/current_document.php?id=93489&id_meeting=198) | WG 03 | 01201 | 2024-04-27 15:44:28 |  | WG 03 MPEG-I | Draft registration of Khronos extensions 2nd edition | WG 03 MPEG Systems |
| [MDS24029](https://dms.mpeg.expert/doc_end_user/current_document.php?id=93695&id_meeting=198) | WG 03 | 01259 | 2024-04-27 15:44:28 |  | WG 03 All | Liaison letter to Khronos on glTF validator software | WG 03 MPEG Systems |

## Communication from MPEG#147

No communication was sent.

## Communication from MPEG#148

No communication was sent.

# Requirements, Scenarios and Test Assets

## Requirements

The work of the MPEG-I scene description is based on the requirements defined in N18965, later revised to N19511. The coverage of the requirements and the progress is documented in

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| [MDS23203](https://dms.mpeg.expert/doc_end_user/current_document.php?id=90529&id_meeting=196) | WG 03 | 01050 | 2023-10-20 22:21:23 |  | WG 03 MPEG-I | Requirements Coverage of MPEG-I Scene Description | WG 03 MPEG Systems | |  | | --- | | MDS23203\_WG03\_N01050 | |

Updated Requirements are available here

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| [MDS23869](https://dms.mpeg.expert/doc_end_user/current_document.php?id=93547&id_meeting=198) | WG 02 | 00359 | 2024-04-27 15:44:28 |  | WG 02 MPEG-I | Draft of MPEG-I Phase 3 requirements | WG 02 MPEG Technical requirements | |  | | --- | | MDS23869\_WG02\_N00359 | |

## Scenarios

Providing Extension to MPEG-I Scene Description is based on well-defined and agreed scenarios. WG3\_N0761 also covers the mapping of requirements to scenarios.

Scenarios include:

* Description of the scenario
* A set of test assets that are needed for the scenario

Agreed scenarios and test assets can be accessed:

* <https://gitlab.com/mpeg-i/scene-description/scenarios/>

Agreed Test Assets can be accessed here.

* <http://mpegfs.int-evry.fr/mpegcontent/ws-mpegcontent/MPEG-I/Part14-SceneDescriptions>

Note: access and contribution to this requires an account. To request an account, please contact the test asset coordinators (see clause 9)

For adding new scenarios, please provide an input contribution to MPEG with the following information

* Description of the scenario
* A set of test assets that are needed for the scenario

A template for the scenario is provided in clause 5.3.

## Template for Test Scenario

The following table should be used to propose test scenarios for scene description:

|  |  |
| --- | --- |
| Item | Description |
| Title | <give it a catchy title, e.g. as those listed in clause 2> |
| Description | * What is the basic use case? * How does it relate to MPEG-I Requirements and Use Cases? |
| Required test assets | * 3D scene, real-time assets for media (2D/3D) * Anything else * References to test assets |
| Current Support | * How can glTF Scene Description be used today * What are gaps/inefficiencies of glTF2.0 to address this scenario? |
| Criteria | * What are relevant criteria for the user experience/QoE? * What are relevant criteria for passing the test scenario? |

## Continuous Call for Test Data

Among others, we solicit the following material to be used as content for the creation and validation of MPEG-Scene Descriptions:

* 2D content that can server as overlays, video textures
* 2D and 3D content that is captured from a local camera, e.g. representing a conference room or flat surfaces for overlay
* 3D game content, e.g. provided in Unity, that can be used for the online gaming scenario
* 3D cinematographic content that includes complete scenes
* VR content and 3D mesh and point cloud content that can be used for VR scenes
* etc…

We welcome contributions of content that can be made available to the MPEG community for the sake of the MPEG-I Scene Description activity.

## Timeline

The data sets should be submitted as input contributions to the 142nd MPEG meeting (April 2023), but early submission into AHG is welcome.

## Available Test Assets

The following table lists the available assets and provides a brief description:

http://mpegx.int-evry.fr/software/MPEG/Systems/SceneDescription/test-assets

# Roadmap

## Introduction

As ISO/IEC 23090-14 AMD2/Edition 2 is on a path to be finalized, this contribution discusses the potential topics to be addressed in subsequent amendments, and their timelines. The requirement coverage document [m65644](https://dms.mpeg.expert/doc_end_user/documents/144_Hannover/wg11/m65644-v2-m65644-SD-MPEG-I_requirement_coverage-r1.zip), the MPEG-I Requirement document (m68115) the latest TuC (n1127) and the TuC status document ([m66557](https://dms.mpeg.expert/doc_end_user/documents/145_OnLine/wg11/m66557-v1-m66557-%5bSD%5dProposedUpdatestoTuC.zip)) were also considered, in particular for the requirements identified for phase 3.

## Agreed Timeline

The following amendments and part are proposed to be created:

* Amd1 to ISO/IEC 23090-14 second edition.
  + on new and feature enhancement, scene understanding, immersive audio and multi-users support.
  + Request a subdivision for the amendment at this meeting.
  + Expected timeline: WD in July, Cdam in Nov 2024, Dam 05 2025.
* Amd2 to ISO/IEC 23090-14 second edition.
  + on haptic phase 2, VDMC support, and integration of avatar (s).
  + Amd to be created in jan. 2025.
* A new part of 23090 for MPEG-Avatar
  + Add a mandate to the ad-hoc group to develop the process to address the MPEG Media Avatar requirements defined in the MPEG-I requirements document.
  + Created a new Part.
  + The Annex H of 23090-14 will be integrated in this new part while remaining in the second edition of 23090-14 for the time being.

The details of the topics for the proposed amd1 and amd2 are described below.

## Relevant Documents

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| [MDS23869](https://dms.mpeg.expert/doc_end_user/current_document.php?id=93547&id_meeting=198) | WG 02 | 00359 | 2024-04-27 15:44:28 |  | WG 02 MPEG-I | Draft of MPEG-I Phase 3 requirements | WG 02 MPEG Technical requirements |
| [MDS23969](https://dms.mpeg.expert/doc_end_user/current_document.php?id=93637&id_meeting=198) | WG 03 | 01250 | 2024-04-27 15:44:28 |  | WG 03 MPEG-I | Draft request for IS/EC 23090-14 2nd edition AMD 1 Support of MPEG-I immersive audio, scene understanding and other extensions | WG 03 MPEG Systems |

## Topic 1: New and features enhancements to Ed 2.

Codec support in the TuC:

**= Support for multiple atlases for MIV applications (MPEG142)**

Source: https://mpeg.expert/software/MPEG/Systems/SceneDescription/MPEG-Contributions/-/issues/484[m62515]

**= Supporting Multiple Viewers in the Media Access Function**

Source: http://mpegx.int-evry.fr/software/MPEG/Systems/SceneDescription/MPEG-Contributions/-/issues/242[m58510]

Improvement to the lighting feature.

MPEG-I requirements phase 3:

General

## Topic 2: Support for scene understanding

Scene understanding can be defined as the human capability to interpret scenes based on what is seen and, on the ability to infer what is not seen. In Computer vision, two major technologies are used to enable a scene understanding in 3D/XR scene format: Segmentation & Annotation, and Spatial mapping. These can be realized by a variety of means, including AI tools.

In glTF or MPEG SD, there is currently no support, whether in the scene graph or in other MPEG extensions, for **enabling the use of** segmentation/annotation mechanisms & results nor for spatial mapping information.

Examples of industry practice:

* <https://learn.microsoft.com/en-us/windows/mixed-reality/design/scene-understanding>
* <https://developers.facebook.com/blog/post/2023/04/25/presence-platform-overview/>
* <https://developer.apple.com/documentation/realitykit/realitykit-scene-understanding>

Related activities in 3GPP:

3GPP is developing in release 19, studies on Localized Mobile Metaverse Services, which will likely lead to normative work in 6G releases for the support of more advanced XR capabilities.

[SP-220353](https://www.3gpp.org/ftp/Information/WI_Sheet/SP-220353.zip) from SA1, notably considers the support of:

* Interactive XR media shared among multiple users in a single location, …
* Acquisition, use and exposure of local (physical and digital) information to enable metaverse services, including acquiring and exposing local spatial/environmental information and user/UE(s) information (including viewing angle, position and direction)”

At the same time, in SA4, some contributions for Split Rendering over IMS, relate to Spatial computing and scene understanding: [[**S4-240583**](https://www.3gpp.org/ftp/TSG_SA/WG4_CODEC/TSGS4_127-bis-e/Docs/S4-240583.zip)](http://S4-240583) provides a new scenario, while [**S4-240688**](https://www.3gpp.org/ftp/TSG_SA/WG4_CODEC/TSGS4_127-bis-e/Docs/S4-240688.zip)identifies network capability to manage a scene understanding on behalf of UEs.

Collected problem statements and industry needs in the TuC:

**= On the support of real environment data**

Source: http://mpegx.int-evry.fr/software/MPEG/Systems/SceneDescription/MPEG-Contributions/-/issues/444[m61811]

**= Semantic representation**

Source: http://mpegx.int-evry.fr/software/MPEG/Systems/SceneDescription/MPEG-Contributions/-/issues/576[m64402]

MPEG-I requirements:

|  |  |  |
| --- | --- | --- |
| Spatial Computing Server requirements for eXtended Reality (XR) | | |
| Requirements on SCS | | |
| 134 | The SCS shall provide XR Spatial Description in a standard representational format (e.g. scene description), upon request of XR devices (UEs) on different platforms (desktop and mobile). | May need to Phase 3 |
| 135 | The SCS shall accept the XR spatial data (XR Features) captured by XR UEs over the network or direct connection. | Phase 3 |
| 136 | The SCS shall deliver the XR Spatial Description to XR UEs over the network or direct connection. | Phase 3 |
| 137 | The SCS shall convert other representation formats to XR Spatial Description. | Phase 3 |
| XR Spatial Description | | |
| 139 | The Description shall contain XR features for indoor and/or outdoor localization/tracking purpose. | Partially completed in AMD2  To be continued in Phase3. |
| 143 | The Description shall allow navigation of the XR client from one space to another when multiple indoor XR spaces are used. | partially completed.  Some potential work in Phase 3. |
| XR Spatial Description | | |
| Requirements on XR Client | | | |
| 151 | The XR Client shall capture environment visual data (image, video, or other visual cues), and/or non-visual features such as environment audio data (audio sources, proximity, or other audio cues), environment location data (from IMU and GPS sensors), and environment object data (geometry, surface and material characteristics, proximity, or other object cues), that is, the XR features; and generate a collection of XR features (for the privacy concern) at multiple locations within the environment. | SD phase 3 support for gaze capture, and environmental data (sensor) (e.g. how to interact with sensor data) |
| 152 | The XR Client should capture its spatial (local or global) location(s) through the sensors, if available. | Mostly out of scope of SD |
| 153 | The XR Client should generate a new XR spatial description and upload it to the SCS for further processing (SCS-002) over the network or direct connection. | API and SD spatial description uplink proposed for phase 3 |
| 154 | The XR Client should update an existing XR spatial description received from SCS and upload it to the SCS, if feasible | API and SD spatial description uplink Proposed for phase 3 |
| 155 | The XR Client should load new XR Descriptions when navigating from one physical place to another (with the help of 144).  *144: The Description shall include XR Anchor objects with presentation properties as the placement for XR Media object* | To be addressed in phase 3 with interactivity/ event-based udpate |

During MPEG#147, as use case in this context was agreed:

# Overview of XR Spatial Computing of real environment

## Meta

### Space Setup

Assisted Space Setup is one of the new features ([1]) in a major update (v64 update) that has been recently released for Meta Quest headsets. It runs on Meta Quest 3.

Before that update, the mapping of room was done manually. In Table 1, the supported Meta labels are mapped to the AR Foundation labels in Unity.

|  |
| --- |
| **Meta Label Unity AR Foundation Label** |
| DESK Table  COUCH Seat  FLOOR Floor  CEILING Ceiling  WALL\_FACE Wall  DOOR\_FRAME Door  WINDOW\_FRAME Window  SCREEN Other  LAMP Other  PLANT Other  STORAGE Other  BED Other  OTHER Other |

Table 1: mapping between manual based Meta Labels and Unity labels

Space Setup automatically identifies and marks furniture in an indoor room. For now, it works for windows, doors, tables, sofas, storage, screens and beds, with additional furniture types supported over time.

Segmentation:

Objects are represented with bounding boxes around each identified object. The segmentation can be adjusted manually.

classification:

Meta provides the following labels:

|  |
| --- |
| **Space Setup: Labels** |
| Floor (Not displayed)  Ceiling (Not displayed)  Wall (Not displayed)  Door  Table  Sofa  Storage  Screen  Bed |

Table 2: Space Setup/Labels

Processing:

The processing is done on the fly.

XR Spatial computing with Space Setup:

A screenshot of a computer

Description automatically generated

Figure 1:Space Setup ([2])

### SceneScript

SceneScript has been introduced by Reality Labs Research ([3]). SceneScript is a method for representing and inferring scene geometry, it’s an output of [Project Aria](https://www.projectaria.com/). This project is a research device from Meta, which is worn like a regular pair of glasses, and enables researchers to study the future of always-on egocentric perception.

Currently, SceneScript runs on the fly on simulation platform but not on Quest 3.

SceneScript is a feature that uses AI to automatically identify a room's architecture and objects such as walls, doors, windows, tables, and more. It’s an indoor mapping.

Segmentation:

Objects are represented within the SceneScript as a set of simple geometric primitives (for instance, a table is a set of cylinders and a box).

classification:

Somes labels are provided by SceneScript. This set of labels can be enriched.

|  |
| --- |
| **SceneScript: Labels (from Aria Synthetic Environments)** |
| table, sofa, shelf, chair, bed, floor\_mat, exercise\_weight, cutlery, container, clock, cart, vase, tent, flower\_pot, pillow, mount, lamp, ladder, fan, cabinet, jar, picture\_frame, mirror, electronic\_device, dresser, clothes\_rack, battery\_charger, air\_conditioner, window. |

Table 3:Scenescript / labels

XR Spatial computing with SceneScript:

A room with chairs and tables

Description automatically generated

Figure 2: SceneScript([3])

## Google / ARCore

### Scene Semantics API

Built on a machine learning model, the Scene Semantics API ([4])provides real-time outdoor semantic information, which complements existing geometric information in ARCore.

Given an image of an outdoor scene, the API returns a label for each pixel across a set of useful semantic classes. In addition to pixel labels, the Scene Semantics API also offers confidence values for each pixel label and an easy-to-use way to query the prevalence of a given label in an outdoor scene.

Segmentation:

Objects are segmented on the current image (2D segmentation). Each pixel has a label and a confidence value.

classification:

ARCore provides the following classification with 12 labels (with an additional “unlabeled” label):

|  |  |
| --- | --- |
| **Semantic label quality tiers** |  |
| Main scene components | * sky * building * tree * road * vehicle |
| Major scene details | * sidewalk * terrain * structure * water |
| Minor scene details | * object * person |

Table 4: Scene Semantics / Labels

Processing:

The processing is done on the fly.

XR Spatial computing with Scene Semantics:

A screenshot of a video game

Description automatically generated

Figure 3: Scene Semantics ([4])

## Apple / ARKit

### RoomPlan API

RoomPlan ([5]) is a new Swift API that utilizes the camera and LiDAR Scanner on iPhone and iPad to create a 3D floor plan of a room, including key characteristics such as dimensions and types of furniture.

The 3D room layout estimator uses two neural networks, one that detects walls and openings, and another that detects doors and windows. The estimator detects walls and openings as lines and lifts them into 3D using estimated wall height. It detects doors and windows on 2D wall planes and later projects them into 3D space, given the wall information and camera position.

The 3D object-detection pipeline recognizes 16 object categories directly in 3D, covering major room-defining furniture types, such as sofa, table, and refrigerator.

Segmentation:

Objects are represented with bounding boxes.

classification:

ARKit provides the following labels:

|  |
| --- |
| **RoomPlan: Labels** |
| Storage  Sofa  Table  Chair  Bed  Refrigerator  Oven  Stove  Dishwasher  Washer or dryer  Fireplace  Sink  Bathtub  Toilet  Stairs  TV |

Table 5: RoomPlan / Labels

Processing:

The processing is done on the fly.

XR Spatial computing with RoomPlan:

A screen shot of a room

Description automatically generated

Figure 4: RoomPlan ([5])

## Topic 3: Immersive audio

As immersive Audio was not addressed in 23090-14 Amd2/Second edition, it should be addressed in the Amd1 of the second edition.

MPEG-I Audio in Scene Description in the TuC:

Contains a few proposals and notably:

**= Immersive audio extension**

Source: https://mpeg.expert/software/MPEG/Systems/SceneDescription/MPEG-Contributions/-/issues/538[m63549]

<https://git.mpeg.expert/MPEG/Systems/SceneDescription/MPEG-Contributions/-/issues/659>

And m67011, 66705.

MPEG-I requirements:

Requirements 5, 32, 33, 48 to 74, 78, 80 (in the MPEG-I requirements document).

## Topic 4: V-DMC support

V-DMC reaches CD Stage in April 2024.

MPEG-I Requirements including:

* 11 - The scene description shall support audio, video and other media formats standardized by MPEG.
* 87 - The scene description shall support information to enable a renderer to output raster data (image, and video), volumetric data (point clouds, meshes, arrays of voxels, and reflectance fields) and audio.

## Topic 5 Support for Haptics Phase 2

Haptics Phase 2 is expected to reach the CD stage by the end of the year.

* SD should provide extensions for supporting spatial and interactive Haptics
* Extensions are proposed to be developed and validated in the Haptics AhG and then proposed to SD.
* Regular synchronization to be set up.

MPEG-I Requirements:

|  |  |  |
| --- | --- | --- |
| 80 | The scene description shall support parametric models for use in rendering environmental acoustic behaviour (e.g. reverberation, occlusion and directivity). | Partially completed.  Immersive audio and Haptic phase 2 in SD phase 3 |
| I-l Interface: Local capture Interface | | |
| 84 | It shall be possible to provide feedback through available actuators | Should partially be covered by haptic phase 2. Need support in phase 3 of SD. |
| I-i Interface: User inputs Interface | | |
| 86 | it shall be possible to define custom interactivity procedures based on input from the user or from the user’s devices and sensors | Partially completed with Interactivity extension and OpenXR.  SD Phase 3 in particular for haptics and avatar |
| Haptic and interaction model | | |
| 130 | [Haptics Phase 2B] The specification shall support coding and presentation of interactivity models related to avatar-scene or avatar-avatar interactions. | Partially addressed in SD AMD2. To be completed in SD phase 3 |

## Topic 6: Multi-users support and avatar integration.

The following enhancements are proposed to be developed, taking in account alignment with the 3GPP SA4 architecture:

* Extensions in SD needed to support the below MPEG-I requirements for multi-users & social VR, and the integration of the MPEG animated avatar.
* Support for Animation stream integration in the MAF
* Support for Interaction between avatars and, between avatar(s) and the SD scene.

The work on these features will be align with the 3GPP architecture and developed in 3GPP SA4.

MPEG-I requirements:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| I-i Interface: User inputs Interface | | | | |
| 86 | | | it shall be possible to define custom interactivity procedures based on input from the user or from the user’s devices and sensors | Partially completed with Interactivity extension and OpenXR.  SD Phase 3 in particular for haptics and avatar |
| Social VR | | | | |
| 104 | | The specification shall support metadata that provides the direction of view of the local user in its immersive environment | | Partially completed with avatar extension.  Clarify the requirement with pose metadata (completed in MAF) and gaze metadata (not completed – phase 3) |
| 105 | | The specification shall support metadata that indicates which remote user is being looked at by the local user | | Partially completed with avatar extension.  Clarify the requirement with pose metadata (completed) and gaze metadata (not completed – phase 3) |
| 117.1 | | The specification shall enable making the omnidirectional environment consistent: | | Phase 3 | |
| 117.1.1 | | - for all users that are embedded in their omnidirectional environment | | Phase 3 | |
| 117.1.2 | | -  between the users (they can look at one another, and can see when another person is looking at them) | | Phase 3 for gaze support and additional metadata | |
| 117.1.3 | | -  for multiple users looking at / pointing at a common element in the omnidirectional environment (e.g., a video screen) | | Phase 3 for gaze support and additional metadata | |
| 117.2 | | The specification shall support synchronization of user viewpoints and orientations (i.e. where each user is looking at), as well as content playback status, between users. Note: these individual environments need not necessarily be the same, as long as they are internally consistent for all participants individually Note: “consistent” means the right visual perspective for all objects as well as audio/visual alignment | | Phase 3 for gaze support | |
| 117.3 | | The specification shall enable bringing multiple users together in the same omnidirectional environment even when they are captured using their own individual 3DoF coordinate system | | Phase 3 | |
| Multi-User Interactivity | | | | | |
| 93 | The specification shall enable realistic composition of user-embodiment within 6DoF content media.  Note: Such composition may, e.g., include delivering proper lighting information and some form of geometry information of the scene so user-embodiment is rendered with realistic lighting and shadows. | | | Partial support with Avatar extension,  Additional support likely needed in phase 3 | |
| 94 | The specification shall support rendering of other users in 6DoF content media, including possible speech or audio from other users. | | | Partial support with Avatar extension  Additional support likely needed in phase 3 | |
| 95 | The specification shall enable multi-user immersive applications in which several users are experiencing the same immersive experience together. | | | Not completed yet, To be addressed in phase 3 | |
| 95.1 | It shall be possible to detect & render interactions between users within the immersive environment. | | | Not completed yet, To be addressed in phase 3 | |

# Contributions for Extensions

## General

For every extension documented in ISO/IEC 23090-14 under the framework in clause 3 the following information is expected to be provided:

* The schema for the extension as part of the standard as well as a json document
* The semantics for the extension
* The processing model on the "Presentation Engine"
* The conformance description, i.e. conformance requirements for the Presentation Engine that supports the extension
* *A promise for example content that uses the extension that is finally available within 1 meeting after the technology was added. If not fulfilled, the feature is expected to be removed and this will be documented as a note in the draft standard.*
* *A promise of a reference implementation in one of the agreed reference software libraries as documented in clause 7, that is finally available within 2 meetings after the technology was added. If not fulfilled, the feature is expected to be removed and this will be documented as a note in the draft standard.*

Hence, contributions addressing extensions to glTF under the framework in clause 3 should include the following:

* The scenarios that this extension is addressing.
* All information from above

As long as not all the above information is available, a documented extension is not moved into the WD/CD, but is maintained in the Technology under Consideration (TuC) document. The status of the completed information and the missing one is documented in the TUC.

The following text processes is recommended, but needs final verification:

*To fulfil the requirement on the reference software, it is sufficient to demonstrate that the reference software is able to properly process the test scenario. The test scenario content shall at least have a scene description file in glTF textual format that makes use of the proposed extension. The test scene description glTF document should use one of the available assets. The proposal must indicate any dependencies on other extensions.*

*The following is an example of this procedure:*

* *A test scenario is defined around support for video textures*
* *The proposal is to make use of the MPEG\_video\_texture extension*
* *A sample content is proposed based on the "conferenceroom" glTF file, which is part of the assets. The glTF file is extended to include the MPEG\_video\_texture extension. The bbb.mp4 asset is used to describe the video texture, which is attached to a rectangular mesh in the "conferenceroom" scene.*
* *The reference software is run with the modified scene description document and the expected behavior is demonstrated, showing the video texture.*

## Extension Principles

The following extension principles apply

* If the extension adds a new top-level array (by extending the root glTF object), its elements should inherit all properties of glTFChildOfRootProperty.schema.json.
* Other objects introduced by the extension should inherit all properties of glTFProperty.schema.json.
* By glTF 2.0 conventions, schemas should allow additional properties.
* Names MUST begin with an MPEG prefix, followed by an underscore.
* Names MUST use lowercase snake-case following the prefix, e.g. MPEG\_materials\_sand.
* Names SHOULD be structured as MPEG\_<scope>\_<feature>, where scope is an existing glTF concept (e.g. mesh, texture, image) and feature describes the functionality being added within that scope. This structure is recommended, but not required.
* Scope SHOULD be singular (e.g. mesh, texture), except where this would be inconsistent with an existing Khronos extension (e.g. materials, lights).

# Reference Software

The reference software for the scene description is documented in ISO/IEC 23090-24 as available in :



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| [MDS24437](https://dms.mpeg.expert/doc_end_user/current_document.php?id=96499&id_meeting=200) | WG 03 | Working Draft of ISO/IEC 23090-24 AMD 1 Conformance and reference software for scene description on haptics, augmented reality, avatars, interactivity, MPEG-I audio and lighting | | WG 03 MPEG Systems |
| [MDS24438](https://dms.mpeg.expert/doc_end_user/current_document.php?id=96500&id_meeting=200) | WG 03 | Procedures for test scenarios and reference software development for MPEG-I Scene Description | | WG 03 MPEG Systems |

# Gitlab Management

For details on test assets, conformance and reference software, as well as test scenarios, refer to WG3 N0782. A summary of the assets is provided here:

|  |  |  |
| --- | --- | --- |
| **Asset** | **Hosting** | **Location name** |
| Repository | Gitlab.com | https://gitlab.com/mpeg-i/scene-description |
| MPEG Trimesh (mpegtrimesh) Reference software | Gitlab.com | https://gitlab.com/mpeg-i/scene-description/mpegtrimesh |
| Conformance software | Gitlab.com | https://gitlab.com/mpeg-i/scene-description/conformance |
| Scenarios | Gitlab.com | https://gitlab.com/mpeg-i/scene-description/scenarios |
| Test vectors | Gitlab.com with LFS for binary files | <https://gitlab.com/mpeg-i/scene-description/test-vectors> |
| Test assets | MPEG content | <http://mpegfs.int-evry.fr/mpegcontent/ws-mpegcontent/MPEG-I/Part14-SceneDescriptions> |

For access to the project, please register an account on GitLab.com at <https://gitlab.com/users/sign_in> and collect the following information:

* GitLab.com username
* GitLab.com email address

Please then send an email containing this information to the gitlab managers as listed in clause 9.

For uploading content to the Test Assets, please bring an input contribution to the MPEG meeting.

# Coordinators for Efforts until MPEG#149

* BOG Chair:
  + Thomas Stockhammer (tsto@qti.qualcomm.com)
* AHG Chairs:
  + Thomas Stockhammer (tsto@qti.qualcomm.com)
  + Mary-Luc Champel ([champelmaryluc@xiaomi.com](mailto:champelmaryluc@xiaomi.com))
  + Gaëlle Martin-Cocher
* Editor of ISO/IEC 23090-14:2022
  + Imed Bouazizi (bouazizi@qti.qualcomm.com)
  + Lukasz Kondrad ([lukasz.kondrad@nokia.com](mailto:lukasz.kondrad@nokia.com))
  + Yago Sanchez ([yago.sanchez@hhi.fraunhofer.de](mailto:yago.sanchez@hhi.fraunhofer.de))
  + Thomas Stockhammer (tsto@qti.qualcomm.com)
* Editor of ISO/IEC 23090-14:2022/Amd.1
  + Imed Bouazizi (bouazizi@qti.qualcomm.com)
  + Gurdeep Bhullar ([Gurdeep.Bhullar@InterDigital.com](mailto:Gurdeep.Bhullar@InterDigital.com))
  + Thomas Stockhammer (tsto@qti.qualcomm.com)
* Editor of ISO/IEC 23090-14:2024 (second edition)
  + Imed Bouazizi (bouazizi@qti.qualcomm.com)
  + Emmanuel Thomas (thomase@xiaomi.com)
  + Patrice Hirtzlin (Patrice.Hirtzlin@InterDigital.com)
  + Thomas Stockhammer (tsto@qti.qualcomm.com)
* Editor of ISO/IEC 23090-14:2024/Amd.1
  + Imed Bouazizi ([bouazizi@qti.qualcomm.com](mailto:bouazizi@qti.qualcomm.com))
  + Sylvain LeLievres
  + Thomas Stockhammer (tsto@qti.qualcomm.com)
* Editor of Technology under Considerations Document
  + Lukasz Kondrad ([lukasz.kondrad@nokia.com](mailto:lukasz.kondrad@nokia.com))
  + Imed Bouazizi (bouazizi@qti.qualcomm.com)
* Test Asset and Scenario Coordinator
  + Emmanuel Thomas (thomase@xiaomi.com)
  + Imed Bouazizi ([bouazizi@qti.qualcomm.com](mailto:bouazizi@qti.qualcomm.com))
* Gitlab Management
  + Emmanuel Thomas (thomase@xiaomi.com)
  + Imed Bouazizi ([bouazizi@qti.qualcomm.com](mailto:bouazizi@qti.qualcomm.com))
* Editor of ISO/IEC 23090-24
  + Ahmed Hamza (Ahmed.Hamza@InterDigital.com)
  + Gurdeep Bhullar ([Gurdeep.Bhullar@InterDigital.com](mailto:Gurdeep.Bhullar@InterDigital.com))
  + Imed Bouazizi (bouazizi@qti.qualcomm.com)
  + Emmanuel Thomas (thomase@xiaomi.com)
  + Thomas Stockhammer (tsto@qti.qualcomm.com)