International Standard

ISO/IEC 23090-28

Information technology — Coded representation of immersive media —

Part 28: Interchangeable scene-based media representations

Technologies de l'information — Représentation codée de média immersifs —

Partie 28: Représentations de médias basés sur des scènes interchangeables

First edition
2026-02

Reference number
ISO/IEC 23090-28:2026(en)

© ISO/IEC 2026



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2026

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
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword ... **iv**

Introduction ... **v**

1 Scope ... **1**

2 Normative references ... **1**

3 Terms and definitions ... **1**
3.1 General ... 1
3.2 Abbreviated terms ... 3

4 Architectures for scene-based media ... **3**
4.1 General ... 3
4.2 Architecture for scene graph representations ... 4
4.3 Scenes nested within scenes ... 5
4.4 Architecture for annotated scene graph using IMS ... 5
4.5 Architecture for translation between scene graph representations using IMS ... 5

5 Independent mapping space (IMS) ... **7**
5.1 General ... 7
5.2 Architecture of IMS ... 7
5.3 IMS systems ... 7
5.3.1 General ... 7
5.3.2 IMS logical organization component ... 9
5.3.3 IMS physical organization component ... 18
5.3.4 IMS processing directions component ... 23

Annex A (normative) **Annotation using IMS and ITMF** ... **32**

Annex B (normative) **Mapping of IMS to glTF 2.0** ... **37**

Annex C (normative) **Annotation process for glTF** ... **39**

Bibliography ... **45**

© ISO/IEC 2026 – All rights reserved

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

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 document 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 or www.iec.ch/members_experts/refdocs).

ISO and IEC draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and https://patents.iec.ch. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

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. In the IEC, see www.iec.ch/understanding-standards.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

A list of all parts in the ISO/IEC 23090 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.

© ISO/IEC 2026 – All rights reserved

Introduction

Emerging display technologies capable of creating fully formed holograms, e.g. 3D aerial images comprised of waveforms emitted from a display panel surface, require input source media that are sufficiently dense in information such that each of the objects in a holographic visual scene is described in terms of its complete and explicit geometry, and surface properties to characterize how the surface responds to the presence of light. Such visual information can furthermore facilitate display technologies to produce visual scenes of objects that cannot be distinguished from the same objects when viewed in the natural world; i.e., the visual information of the media is sufficiently dense such that the display is enabled to produce a photorealistic result.

Commonly used raster media formats such as those specified in ITU-R BT.601, ITU-R BT.709, ITU-R BT.2020, ITU-R BT.2100 were originally developed to efficiently provide information sufficient for 2D displays to produce 2D visual presentations, albeit with varying degrees of spatial resolutions, frame rates, and pixel depth. These raster formats, however, neither provide sufficient information in a practical nor efficient manner, to enable a holographic display to produce holograms. As an alternative to raster media, 3D graphics formats used in digital content creation tools and production workflows offer a solution to provide sufficiently dense information to enable photorealistic results where raster formats do not. The use of 3D graphics technologies hence serves as a fundamental technology in enabling immersive displays to create their optimal visual presentations.

To date, immersive displays provide a media input interface that is capable of ingesting 3D graphics formats for presentation by the display. Such interfaces can be enabled and shipped as part of the onboard processing features in the display by use of game engine software development kits. However, 3D scene media that are targeted for processing by the display's choice of renderer must first be translated into a format that is consistent with the internal scene representations of the renderer. To facilitate the translation and distribution of 3D scene media into the renderer interface provided by the display, an independent mapping space (IMS) is specified as a part of the ISO/IEC 23090 series.

At the core of the IMS is a vocabulary of terms and definitions that can be used to annotate a scene graph or scene asset with metadata, i.e. to facilitate the translation of media from one format to another. Such translation from one format to another format is a commonly encountered challenge in the distribution of scene-based media, as a number of formats have emerged to gain popularity in various application areas, e.g., gaming, social media, e-commerce, and content production. The corresponding vocabulary of terms and definitions describes nodes, media formats, media attributes, and render processing instructions of commonly-used scene graph formats and media representations. Elements from the IMS can be used to annotate a scene graph or media format so that its translation to another format is more straightforward and consistently performed. Thus, the IMS facilitates the distribution and interoperability of a variety of scene-based media into render applications that are limited in the types of formats that they are designed to support.

Another application of this document is to guide the translation of media assets for use across a diverse set of applications comprising the Metaverse, e.g., gaming, social media, retail applications. In the absence of a common format that specifies media for use across the variety of Metaverse applications, the IMS can facilitate the translation from one asset representation to another.

© ISO/IEC 2026 – All rights reserved

Information technology — Coded representation of immersive media —

Part 28: Interchangeable scene-based media representations

1 Scope

This document specifies an architecture and data model for interchange of scene-based media for use in systems and applications that employ renderers for the presentation of immersive media, including audio, visual, tactile and other representations of media types.

The architecture and data model specified in this document applies to use cases where there is a need for interchange of scene-based media for presentation by systems with 3D render-based technologies. One specific application of this document is immersive display technologies that utilize a real-time renderer of 3D media to create their visual presentations, as opposed to a video decoder and renderer of 2D raster-based media. Another application of this document is for the distribution of media assets for use across a large variety of applications comprising the Metaverse.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

3.1 General

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

3.1.1
scene graph
general data structure commonly used by vector-based graphics editing applications and modern computer games, that arranges the logical and often (but not necessarily) spatial representation of a graphical scene; a collection of nodes and vertices in a graph structure

3.1.2
immersive technology media format
ITMF
suite of specifications developed by the Immersive Digital Experiences Alliance

3.1.3
node
fundamental element of the scene graph comprised of information related to the logical or spatial or temporal representation of visual or audio information

© ISO/IEC 2026 – All rights reserved

3.1.4
node graph
collection of nodes that can be positioned or modified as a single unit and linked to the remainder of the scene graph by linker nodes

3.1.5
node type
general category of node that performs a certain function, for example, a camera node type can be a panoramic lens camera node, an open shading language (OSL) node, or other types of cameras

3.1.6
attribute
metadata associated with a node used to describe a particular characteristic or feature of a node either in a canonical or more complex form (e.g. in terms of another node)

3.1.7
container
serialized format to store and exchange information to represent all natural, all synthetic, or a mixture of synthetic and natural scenes including a scene graph and all of the media resources that are required for rendering the scene

3.1.8
serialization
process of translating data structures or object state into a format that can be stored (for example, in a file or memory buffer) or transmitted (for example, across a network connection link) and reconstructed later (possibly in a different computer environment)

Note 1 to entry: When the resulting series of bits is reread according to the serialization format, it can be used to create a semantically identical clone of the original object.

3.1.9
renderer
software-based application or process, based on a selective mixture of disciplines related to: acoustic physics, light physics, visual perception, audio perception, mathematics, and software development, that, given an input scene graph and asset container, emits a visual and/or audio signal suitable for presentation on a targeted device or conforming to the desired properties as specified by attributes of a render target node in the scene graph

Note 1 to entry: For visual-based media assets, a renderer may emit a visual signal suitable for a targeted display, or for storage as an intermediate asset (e.g. repackaged into another container i.e. used in a series of rendering processes in a graphics pipeline); for audio-based media assets, a renderer may emit an audio signal for presentation in a multi-channel loudspeaker and/or binauralized headphones, or for repackaging into another (output) container.

3.1.10
scene-based media
audio, visual, haptic, and other primary types of media and media-related information organized logically and spatially by a use of a scene graph

3.1.11
RGB
additive colour model in which red, green, and blue are the primary colours

Note 1 to entry: The abbreviation RGB is derived from the first letter of each of the primary colours.

3.1.12
asset
one of potentially multiple objects comprising the media for a scene in scene-based media; each object fully formed and packaged such that it can be rendered separately from other objects in the scene

Note 1 to entry: Individual assets may be shared, e.g. reused across multiple scenes.

© ISO/IEC 2026 – All rights reserved

3.1.13
pin
input parameter to a node, or an output from a node

3.1.14
this document
referring to this International Standard

3.1.15
independent mapping space label
IMS label
sequence of alphanumeric characters using dot notation comprised of an IMS prefix followed by a node label, which may optionally be followed by labels derived from the ITMF Data Encoding Specification[4]

3.1.16
independent mapping space prefix
IMS prefix
sequence of alphanumeric characters using dot notation comprised of "ims" followed by an alphanumeric label to denote the IMS component, subsequently followed by an alphanumeric label to denote the IMS system for a particular node

3.2 Abbreviated terms

2D	two dimensional
3D	three dimensional
AOV	arbitrary output variables
IDEA	Immersive Digital Experiences Alliance
ITMF	Immersive Technology Media Format
IMS	independent mapping space
OSL	Open Shading Language
RGB	red, green, blue
USD	Universal Scene Description
glTF	graphics language Transmission Format

4 Architectures for scene-based media

4.1 General

This clause illustrates architectures related to the organization of scene-based media and how such media can be annotated by a process that stores IMS metadata into the media. Furthermore, this clause provides a hypothetical architecture for a process that translates one scene-based media format to another.

A more in-depth description of these architectures is provided in ISO/IEC TR 23090-27.[2]

NOTE This document references only some of the architectures in ISO/IEC TR 23090-27[2] for the purposes of providing additional context.

© ISO/IEC 2026 – All rights reserved