INTERNATIONAL STANDARD

ISO/IEC 23090-7

First edition
2022-11

Information technology — Coded representation of immersive media —

Part 7: Immersive media metadata

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

Partie 7: Métadonnées de media immersifs

Reference number
ISO/IEC 23090-7:2022(E)

© ISO/IEC 2022



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2022

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

© ISO/IEC 2022 – All rights reserved

Contents

Page

Foreword ..... **v**

Introduction ..... **vi**

1 Scope **1**

2 Normative references **1**

3 Terms, definitions and symbols **1**
3.1 Terms and definitions 1
3.2 Symbols 4

4 Overview **5**
4.1 General 5
4.2 Variables 5
4.3 Processes 5
4.4 Syntax structures 5

5 Common metadata **6**
5.1 Reference coordinate system 6
5.2 Coordinate system rotation 6
5.3 Common metadata data structures 8
5.3.1 Rotation structure 8
5.3.2 Content coverage structure 8
5.3.3 Viewpoint information structures 8
5.3.4 Sphere region structure 9
5.3.5 Spherical region-wise quality ranking - Syntax 11
5.3.6 2D region-wise quality ranking structure- Syntax 12
5.4 Common metadata semantics 12
5.4.1 Rotation structure - Semantics 12
5.4.2 Content coverage structure - Semantics 12
5.4.3 Viewpoint information structures - Semantics 13
5.4.4 Sphere region structure - Semantics 14
5.4.5 Spherical region-wise quality ranking - Semantics 14
5.4.6 2D region-wise quality ranking structure - Semantics 15

6 Video and image metadata **16**
6.1 Projection formats 16
6.1.1 List of projection formats 16
6.1.2 Equirectangular projection process 17
6.1.3 Cubemap projection process 17
6.2 Region-wise packing formats 20
6.2.1 List of packing formats 20
6.2.2 Rectangular region-wise packing process 20
6.3 Sample location mapping process 21
6.3.1 Relation of decoded pictures to global coordinate axes 21
6.3.2 Mapping of luma sample locations within a decoded picture to sphere coordinates relative to the global coordinate axes 23
6.3.3 Conversion from a sample location in a projected picture to sphere coordinates relative to the global coordinate axes 24
6.3.4 Conversion from a sample location of an active area in a fisheye decoded picture to sphere coordinates relative to the global coordinate axes 25
6.4 Fisheye omnidirectional video 27
6.5 Video and image metadata data structures 27
6.5.1 Projection format structure - Syntax 27
6.5.2 Region-wise packing structure 27
6.5.3 Fisheye omnidirectional video structure 30
6.6 Video and image metadata semantics 32
6.6.1 Projection format structure - Semantics 32

© ISO/IEC 2022 – All rights reserved

6.6.2 Region-wise packing structure 32
6.6.3 Fisheye omnidirectional video structure 36

Bibliography **44**

© ISO/IEC 2022 – 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).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC 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) or the IEC list of patent declarations received (see https://patents.iec.ch).

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 and IEC websites.

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 and www.iec.ch/national-committees.

© ISO/IEC 2022 – All rights reserved

Introduction

This document is organized as follows.

— Clauses 5 describes common metadata applicable to immersive media. This includes reference co-ordinate system related metadata and other common metadata syntax and semantics.

— Clauses 6 describes metadata that applies to video and images. This includes projection formats and packing region-wise formats metadata which applies to video and images.

The goal of this document is to allow reuse of the commonly defined metadata to be referenced by other standards.

The International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) draw attention to the fact that it is claimed that compliance with this document may involve the use of a patent.

ISO and IEC take no position concerning the evidence, validity and scope of this patent right.

The holder of this patent right has assured ISO and IEC that he/she is willing to negotiate licences under reasonable and non-discriminatory terms and conditions with applicants throughout the world. In this respect, the statement of the holder of this patent right is registered with ISO and IEC. Information may be obtained from the patent database available at www.iso.org/patents or https://patents.iec.ch.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights other than those in the patent database. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

© ISO/IEC 2022 – All rights reserved

Information technology — Coded representation of immersive media —

Part 7: Immersive media metadata

1 Scope

This document specifies common immersive media metadata focusing on immersive videos (including 360° videos) and images.

2 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.

ISO/IEC 14496-12, *Information technology — Coding of audio-visual objects — Part 12: ISO base media file format*

ISO/IEC 23008-12, *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 12: Image file format*

3 Terms, definitions and symbols

3.1 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 14496-12 and ISO/IEC 23008-12 and the following 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
azimuth
first of the two *sphere coordinates* (3.1.22) describing the location of a point on the sphere

3.1.2
azimuth circle
circle on the sphere connecting all points with the same *azimuth* (3.1.1) value

Note 1 to entry: An azimuth circle is always a *great circle* (3.1.12).

3.1.3
circular image
image captured with a *fisheye lens* (3.1.9)

© ISO/IEC 2022 – All rights reserved

3.1.4
common reference coordinate system
3D Cartesian coordinate system with the centre being (X, Y, Z) equal to (0, 0, 0), used as the reference coordinate system for all viewpoints within a *viewpoint group* (3.1.27)

3.1.5
content coverage
one or more *sphere regions* (3.1.23) that are covered by the content represented by the track or by an image item

3.1.6
elevation
second of the two *sphere coordinates* (3.1.22) describing the location of a point on the sphere

3.1.7
elevation circle
circle on the sphere connecting all points with the same *elevation* (3.1.6) value

Note 1 to entry: When the elevation is zero, an elevation circle is also a *great circle* (3.1.12). This coincides with the equator on Earth.

3.1.8
field of view
extent of the observable world in captured/recorded content or in a physical display device

3.1.9
fisheye lens
wide-angle camera lens that usually captures an approximately hemispherical *field of view* (3.1.8) and projects it as a *circular image* (3.1.3)

3.1.10
fisheye video
video captured by *fisheye lenses* (3.1.9)

3.1.11
global coordinate axes
coordinate axes that are associated with audio, video, and images representing the same acquisition position and intended to be rendered together

3.1.12
great circle
intersection of the sphere and a plane that passes through the centre point of the sphere

Note 1 to entry: A great circle is also known as an orthodrome or Riemannian circle.

Note 2 to entry: The centre of the sphere and the centre of a great circle are co-located.

3.1.13
guard band
area in a *packed picture* (3.1.16) that is not rendered but may be used to improve the rendered part of the packed picture to avoid or mitigate visual artifacts such as seams

Note 1 to entry: Guard bands are associated with *packed regions* (3.1.17) as described in 6.5.2.

3.1.14
local coordinate axes
coordinate axes obtained after applying rotation to the *global coordinate axes* (3.1.11)

© ISO/IEC 2022 – All rights reserved

3.1.15
omnidirectional video
video and its associated audio that enable rendering according to the user's *viewing orientation* (3.1.26), if consumed with a head-mounted device, or according to user's desired *viewport* (3.1.28), otherwise, as if the user was in the spot where and when the media was captured

3.1.16
packed picture
picture that is represented as a coded picture in the coded video bitstream

3.1.17
packed region
region in a *packed picture* (3.1.16) that is mapped to a *projected region* (3.1.19) as specified by the *region-wise packing* (3.1.21) signalling

3.1.18
projected picture
picture that has a representation format specified by an *omnidirectional video* (3.1.15) *projection* (3.1.20) format

3.1.19
projected region
region in a *projected picture* (3.1.18) that is mapped to a *packed region* (3.1.17) as specified by the *region-wise packing* (3.1.21) signalling

3.1.20
projection
inverse of the process by which the samples of a *projected picture* (3.1.18) are mapped to a set of positions identified by a set of *azimuth* (3.1.1) and *elevation* (3.1.6) coordinates on a unit sphere

3.1.21
region-wise packing
inverse of the process of transformation, resizing, and relocating of *packed regions* (3.1.17) of a *packed picture* (3.1.16) to remap to *projected regions* (3.1.19) of a *projected picture* (3.1.18)

3.1.22
sphere coordinates
azimuth (ϕ) (3.1.1) and *elevation* (θ) (3.1.6) that identify a location of a point on the unit sphere

3.1.23
sphere region
region on a sphere, specified either by four *great circles* (3.1.12) or by two *azimuth circles* (3.1.2) and two *elevation circles* (3.1.7), or such a region on the rotated sphere after applying certain amount of yaw, pitch, and roll rotations

3.1.24
SDL
syntactic description language
language that allows the description of a bitstream's syntax

Note 1 to entry: Syntactic description language is defined in ISO/IEC 14496-1:2010, Clause 8.

3.1.25
tilt angle
angle indicating the amount of tilt of a *sphere region* (3.1.23), measured as the amount of rotation of the sphere region along the axis originating from the sphere origin passing through the centre point of the sphere region, where the angle value increases clockwise when looking from the origin towards the positive end of the axis

© ISO/IEC 2022 – All rights reserved

3.1.26
viewing orientation
triple of *azimuth* (3.1.1), *elevation* (3.1.6), and *tilt angle* (3.1.25) characterizing the orientation that a user is consuming the audio-visual content

Note 1 to entry: In case of image or video, viewing orientation characterizes the orientation of the *viewport* (3.1.28).

3.1.27
viewpoint group
group of viewpoints that share the same *common reference coordinate system* (3.1.4)

3.1.28
viewport
region of omnidirectional image or video suitable for display and viewing by the user

3.2 Symbols

+	Addition.
−	Subtraction (as a two-argument operator) or negation (as a unary prefix operator).
*	Multiplication, including matrix multiplication.
x^y	Exponentiation. Specifies *x* to the power of *y*. In other contexts, such notation is used for superscripting not intended for interpretation as exponentiation.
/	Integer division with truncation of the result toward zero. For example, 7 / 4 and −7 / −4 are truncated to 1 and −7 / 4 and 7 / −4 are truncated to −1.
÷	Used to denote division in mathematical equations where no truncation or rounding is intended.
$\frac{x}{y}$	Used to denote division in mathematical equations where no truncation or rounding is intended.
$\sum_{i=x}^{y} f(i)$	The summation of f(*i*) with *i* taking all integer values from *x* up to and including *y*.
x % *y*	Modulus. Remainder of *x* divided by *y*, defined only for integers *x* and *y* with *x* >= 0 and y > 0.
Asin(*x*)	The trigonometric inverse sine function, operating on an argument *x* that is in the range of −1.0 to 1.0, inclusive, with an output value in the range of −π÷2 to π÷2, inclusive, in units of radians.
Atan(*x*)	The trigonometric invers tangent function, operating on an argument *x* that is any real number, with an output value in the range of −π÷2 to π÷2, inclusive, in units of radians.

© ISO/IEC 2022 – All rights reserved