International Standard

ISO/IEC 23090-4

First edition
2025-11

Information technology — Coded representation of immersive media —

Part 4: MPEG-I immersive audio

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

Partie 4: Audio immersif MPEG-I

Reference number
ISO/IEC 23090-4:2025(en)

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

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 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.

Introduction

0.1 General

MPEG-I Immersive audio reproduction with six degrees of freedom (6DoF) movement of the listener in an audio scene enables the experience of virtual acoustics in a Virtual Reality (VR) or Augmented Reality (AR) simulation. Audio effects and phenomena known from real-world acoustics like, for example, localization, distance attenuation, reflections, reverberation, occlusion, diffraction and the Doppler effect are modelled by a renderer that is controlled through metadata transmitted in a bitstream with additional input of interactive listener position data.

Along with other parts of MPEG-I (i.e. ISO/IEC 23090-12, “Immersive Video”, ISO/IEC 23090-5, “Visual Volumetric Video-based Coding (V3C) and Video-based Point Cloud Compression” and ISO/IEC 23090-2, “Systems Support”), the ISO/IEC 23090 series of standards supports a complete audio-visual VR or AR presentation in which the user can navigate and interact with the simulated environment using 6DoF, that being spatial navigation (x, y, z) and user head orientation (yaw, pitch, roll).

While VR presentations impart the feeling that the user is actually present in the virtual world, AR enables the enrichment of the real world by virtual elements that are perceived seamlessly as being part of the real world. The user can interact with the virtual scene or virtual elements and, in response, cause sounds that are perceived as realistic and matching the users’ experience in the real world.

This document provides means for rendering a real-time interactive audio presentation while permitting the user to have 6DoF movement. It defines metadata to support this rendering and a bitstream syntax that enables efficient storage and streaming of the MPEG-I Immersive Audio content.

0.2 Typesetting of variables

For improved text readability, algorithmic variables are distinguished from descriptive text by their dedicated typesetting. Throughout the text, variables that are input directly from the bitstream are typeset in boldface letters. In the bitstream syntax the use of a boldface variable represents reading bits from the bitstream and converting to an appropriate type (based on the indicated mnemonics and number of bits), and assigning the resulting value to the boldfaced variable. Variables that are derived by computation are typeset in pseudo-code font or italics.