International Standard

ISO/IEC 23090-31

First edition
2025-01

Information technology — Coded representation of immersive media —

Part 31:
Haptics coding

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

Partie 31: Codage haptique

Reference number
ISO/IEC 23090-31:2025(en)

Contents

Page

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

Haptics provide an additional layer of entertainment and sensory immersion to the user. Therefore, the user experience and enjoyment of media content, from a ISOBMFF files, broadcast channel, streaming games or mobile advertisements can be significantly enhanced by the judicious addition of haptics to the audio/video content. To that end, haptics has been proposed as a first-order media type, akin to audio and video, in ISOBMFF. Further, haptics has also been proposed as an addition to ISO/IEC 23009-1 (MPEG-DASH) to signal the presence of haptics in MP4 segments to DASH streaming clients. Lastly, the ISO/IEC 23090 (MPEG-I) use cases have been augmented with haptics

Haptics digital encoding is the storing of tactile data in a digital format. As with audio and video, digital encoding is of fundamental importance to allow digital haptic devices to function. Haptics encoding gained relevance with the increased market importance of wideband haptics in consumer peripherals such as smartphones with haptic engines and game consoles with haptic enabled controllers. The prior generation of haptics peripherals was based on less expressive haptic actuators usually based on state machine control processes.

In the field of haptics, the signal encoding usually takes one of two approaches:

— Quantized: This representation is generally made from measured data. The samples from the original phenomenon are stored at a specific acquisition frequency inside the file to represent this signal. One example of a quantized haptic signal is proposed through WAV files, originally developed for audio. WAV file formalism allows the capture of real-world data and the representation of complex wideband haptic feedback. This type of haptics encoding has the disadvantage of being difficult to modify once encoded due to the inability to access the primitives used to create the signal.

— Descriptive: This representation is used to encode haptic signals as a combination of functions to be synthesized. Examples of such vectorized formats include AHAP and IVS. These formats have the advantage of being created with a composition of primitives. They are easily modifiable at runtime by an application and by dedicated editing tools. Currently, these solutions support only vibrotactile perception, but can easily be extended for other forms of haptics such as kinaesthetic, temperature and textures. They also tend to be memory inefficient with increasing signal complexity and cannot encode non-periodic phenomena

This document describes the coded representation allowing to encode both descriptive and quantized data in a human readable JSON format (.hjif) used as an exchange format. This format can be compressed and packetized into a binary file format for distribution and streaming purposes (.hmpg).

Information technology — Coded representation of immersive media —

Part 31:
Haptics coding

1 Scope

This document specifies technology that supports the efficient transmission and rendering of haptic signals for the playback of immersive experiences in a wide variety of scenarios. The document describes in detail a robust coded representation of haptic media covering the two most popular haptic perceptions leveraged by devices today: vibrotactile and kinaesthetic. Support for other haptic modalities has also been integrated.

The coded representation allows to encode both descriptive and quantized data in a human readable JSON format used for exchange purposes, and a compressed bitstream version, optimised for memory usage and distribution purposes. This approach also allows to meet the expectations for compatibility with both descriptive and quantized formats, as required by the market, as well as interoperability between devices for 3D immersive experiences, mobile applications and other distribution purposes.

Information provided in this document related to the decoder is normative, while information related to the encoder and renderer is informative.

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 8601-1, *Date and time — Representations for information interchange — Part 1: Basic rules*

ISO 8601-2, *Date and time — Representations for information interchange — Part 2: Extensions*

ISO/IEC 21778:2017, *Information technology — The JSON data interchange syntax*

IETF RFC 3986, *Uniform Resource Identifier (URI): Generic Syntax*, available at: https://www.rfc-editor.org/info/rfc3986

IETF RFC 8259, *The JavaScript Object Notation (JSON) Data Interchange Format*, available at: https://www.rfc-editor.org/info/rfc8259

IETF RFC 4648, *The Base16, Base32, and Base64 Data Encodings*, available at: https://www.rfc-editor.org/info/rfc4648

3 Terms, definitions and abbreviated terms

3.1 Terms and definitions

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
actuator
component of a device for rendering haptic sensations

3.1.2
avatar
body (or part of body) representation

3.1.3
band
component in a channel for containing effects for a specific range of frequencies

3.1.4
channel
component in a perception containing one or more bands rendered on a device at a specific body location

3.1.5
device
physical system having one or more actuators configured to render a haptic sensation corresponding with a given signal

3.1.6
effect
component of a band for defining a signal, consisting of a haptic waveform or one or more haptic keyframes

3.1.7
experience
top level haptic component containing perceptions and metadata

3.1.8
haptics
tactile sensations

3.1.9
keyframe
component of an effect mapping a position in time or space to an effect parameter such as amplitude or frequency

3.1.10
metadata
global information about an experience, perception, channel or band

3.1.11
MIHS format
self-contained stream for transporting MPEG-I haptic data

3.1.12
MIHS initialization unit
MIHS unit containing metadata necessary to reset and initialize a haptic decoder

3.1.13
MIHS packet
MIHS data packet which includes metadata or binary effect data

3.1.14
MIHS unit
MIHS data unit covering a duration of time

3.1.15
modality
type of haptics, such as vibration, force, pressure, position, velocity or temperature

3.1.16
perception
haptic perception containing one or more channels of a specific modality

3.1.17
signal
representation of the haptics associated with a specific modality to be rendered on a device

3.1.18
transient
short momentary effect defined by an amplitude and, directly or indirectly, a frequency

3.1.19
vector
direction in space which can be used for haptic signals spatialization

3.2 Abbreviated terms

AC	Arithmetic coding/coder
AHAP	Apple Haptic and Audio Pattern - JSON-like file format that specifies a haptic pattern
ATSC	Advanced Television Systems Committee
CDF9/7	Cohen–Daubechies–Feauveau 9/7
CRC	Cyclic redundancy check
DASH	Dynamic Adaptive Streaming over HTTP (specified in ISO/IEC 23009)
FFT	Fast Fourier Transform
HJIF	Haptics JSON Interchange Format
ID	Identifier
ISOBMFF	ISO Base media file format (specified in ISO/IEC 14496-12)
IVS	XML-based file format for representing haptic effects
JSON	JavaScript Object Notation
LOD	Level of Detail
MIHS	MPEG-I Haptic Stream
MPEG	Moving Picture Experts Group
MPEG-I	MPEG Immersive media (specified in this and other parts of ISO/IEC 23090)
OHM	Object Haptic Metadata - Text file format for haptics metadata
PCM	Pulse-code modulation
SPIHT	Set Partitioning In Hierarchical Trees

3.3 Mnemonics

The following data types are used for data specifications:

bslbf	Bit string with left bit first, where “left” is the order in which bit strings are written. Bit strings are written as a string of 1s and 0s within single quote marks, for example ‘10000001’.
vlcs8	Variable length character string. Contains string data stored as a character array encoded in UTF-8.
uimsbf	Unsigned integer with most significant bit first.
imsbf	Integer with most significant bit first.
vlclbf	Variable length code with left bit first, where “left” refers to the order in which the variable length codes are written.
duimsbf	Decimal stored as unsigned integer with most significant bit first in a given range. Default range is [-1,1]
islif	Integer stream with left integer first, where “left” is the order in which integer strings are written. Only for buffering.
vlislif	Variable length integer stream with left integer first, where “left” refers to the order in which the integers are written.

4 Overview and architecture

4.1 Overview

This document describes a coded representation of haptics based on a data model that enables the encoding of both descriptive and quantized haptic data. Two complementary formats based on this shared data model (Clause 5) are detailed in the specifications: an interchange format (.hjif) detailed in Clause 6, and a packetized compressed binary format for streaming that can also be stored in a binary file (.hmpg) detailed in Clause 7. In addition, a normative decoder and an informative encoder and synthesizer are described in detail.

The .hjif format is a human-readable format based on JSON and is not optimized for memory usage. It can easily be parsed and manually edited, which makes it an ideal interchange format, especially when designing or creating content. For distribution purposes, the .hjif data can be compressed and packetized into a more memory efficient binary .hmpg bitstream. This compression is lossy, with different parameters impacting the encoding depth of amplitudes and frequencies composing the bitstream. The compressed and packetized data can be directly distributed as a MPEG-I haptic stream (MIHS).

4.2 Codec architecture

The generic codec architecture can process both waveform PCM signals (WAV) and descriptive haptic files such as AHAP, IVS or HJIF, the proposed MPEG format. Metadata information is provided to the codec through OHM input files (described in Annex B). An overview of the codec architecture is depicted in Figure 1, and a more detailed description of the input files and encoder architecture is provided in Clause 8.

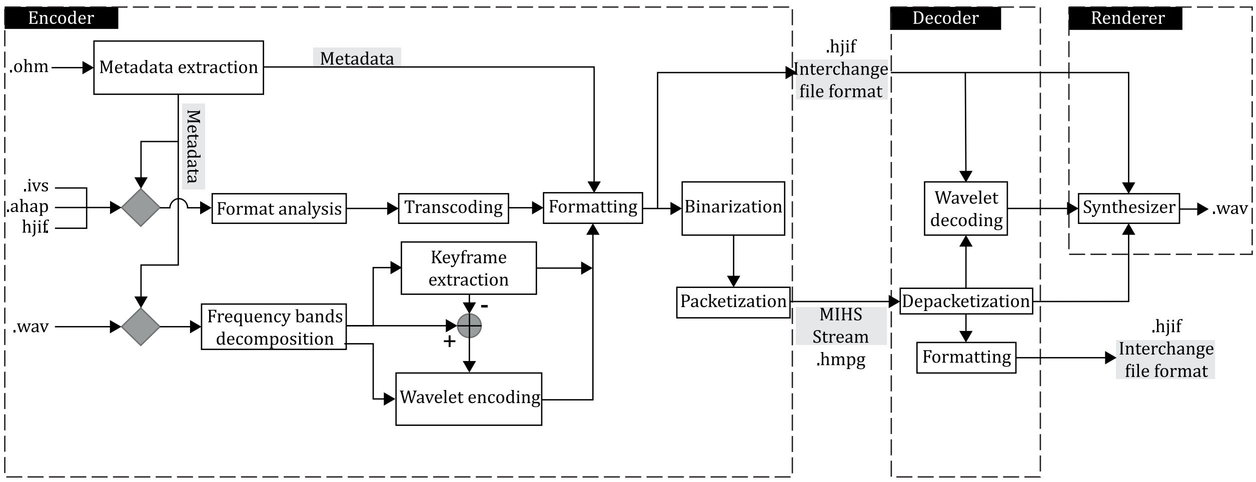


Figure 1 — Overview of the codec architecture

The encoder processes the two types of input files differently. For descriptive content, the input is analysed semantically to transcode (if necessary) the data into the proposed coded representation. For PCM content, the signal analysis process is split into two sub-processes. After performing a frequency band decomposition on the signal, the different bands (one or more) can be encoded as a set of keyframes. In Figure 1, an efficient example is given where the low frequencies are encoded using a keyframe extraction process. The low frequency band(s) is(are) then reconstructed, and the error between this signal and the original low frequency signal is computed. This residual signal is then added to the original high frequency band(s), before encoding using wavelet transforms. This multi-bands hybrid encoding scheme is detailed in Clause 8.

The encoder can output two types of formats: an interchange file format (.hjif) encoded in JSON, and a binary encoded streaming format (MIHS). that can also be stored as a binary file (.hmpg). The two formats have complementary purposes, and a lossy one-to-one conversion can be operated between them. The encoder is informative and is detailed in subclause 8.2.

The decoder takes as input a binary .hmpg file (the MIHS bitstream) or a .hjif file and either outputs a .hjif file or directly synthesizes the haptic data. Haptic data contained in a .hjif file can be rendered directly on haptic devices or using an intermediate synthesizer generating PCM data. The decoder is normative and is detailed in subclause 8.3.

The synthesizer allows to render haptic data from a .hjif input file or MIHS stream into a PCM output file. The synthesizer is informative and is detailed in subclause 8.3.4.

5 Data model

5.1 Data structure overview

This Clause focuses on the data model of the coded representation of haptics. It specifies the information required by a synthesizer to render the haptic data. The following subclauses provide detailed definitions for every property of each element of the data model depicted in Figure 2. The data structure introduced in this Clause is shared by the interchange format and the streaming format defined respectively in Clauses 6 and 7.

The data structure of the two formats follows the hierarchical organization illustrated in Figure 2.

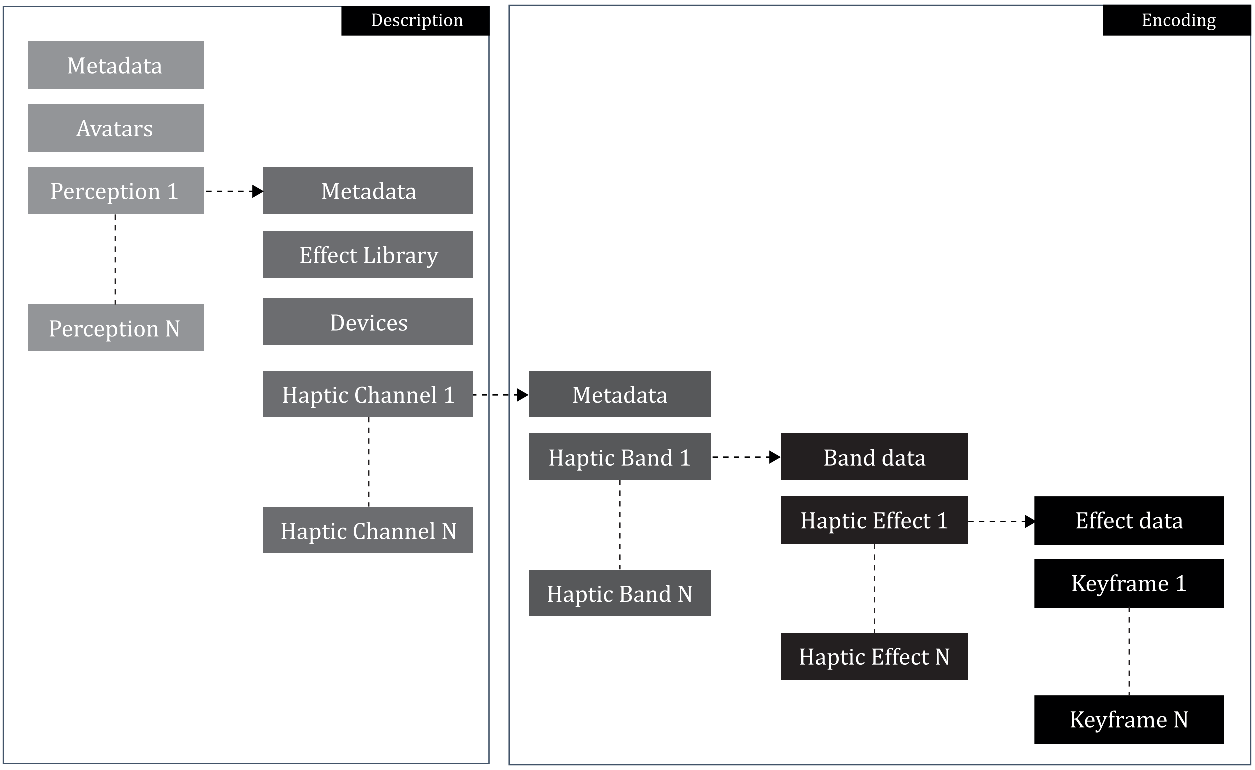
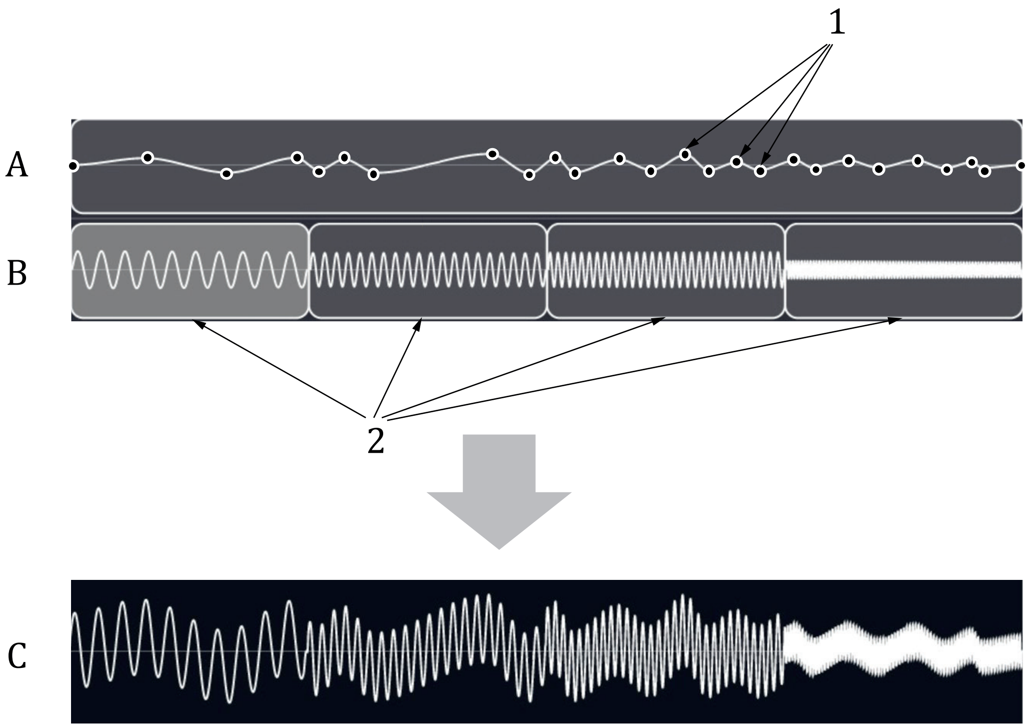


Figure 2 — General data model

The highest level of the structure describes the entire haptic experience defined in the file or stream. It contains some high-level metadata information. It also provides a list of avatars (i.e., body representation) referenced later in the file to specify the desired location of haptic stimuli on the body. Finally, the haptics are described through a list of perceptions. These perceptions correspond to haptic signals associated with specific perception modalities (e.g., vibration, force, position, velocity, temperature).

In addition to specific metadata, a perception contains a list of channels. The data in each channel are decomposed into one or more frequency bands. Each band defines part of the signal in a frequency range. A band is described by a list of haptic effects, each containing a list of keyframes. The haptic signal in a channel can then be reconstructed by combining the data in the different bands as illustrated in Figure 3. For instance, by adding the high and low frequency bands in this figure, the original signal can be reconstructed.

The perception and channel information describe the content (Description part of Figure 2), while the bands, effects and keyframes contain the data of the encoded signals (Encoding part of Figure 2).



Key

1 keyframes

2 effects

A band 1 (*curve* band)

B band 2 (vectorial *wave* band)

C channel 1

Figure 3 — Haptic channel (bottom) and its decomposition in two frequency bands (top)

The format proposes four types of bands: transient bands, curve bands, vectorial wave bands and wavelet wave bands. Each band is composed of a series of effects of the same type as the band, each defined by a list of keyframes. The data contained in the effects and keyframes are interpreted differently for different types of bands.

5.2 Haptic experience

The haptic experience defines the root of the hierarchical data model. It provides information on the date of the file, the version of the format and the encoding profile and level. It also describes the haptic experience, lists the different avatars (i.e., body representations) used throughout the experience, and defines all the haptic perceptions. Table 1 details the properties of a haptic experience.

Table 1 — Haptic experience properties

Property	Description
version	Year of the edition and amendment of ISO/IEC 23090-31 that this file conforms to, in the following format: XXXX or XXXX-Y, where XXXX is the year of publication and Y is the amendment number, if any.
profile	Name of the profile used to generate the encoded stream as defined in the normative Annex D.
level	Number of the level used to generate the encoded stream as defined in the normative Annex D.
date	Creation date of this haptic experience.
description	A user description of this haptic experience.

Table 1 *(continued)*

Property	Description
timescale	Number of ticks per second.
avatars	List of avatars defining body representations used in the haptic experience. See subclause 5.3.
perceptions	List of perceptions describing a haptic signal. See subclause 5.4.

5.3 Haptic avatar

Haptic avatars are used as body representations. Avatars can reference a custom 3D mesh from a companion file. Each perception may be associated with an avatar, which allows spatialization of effects at the channel level. The same avatar can be used by multiple perceptions.

A 3D mesh can provide resolution and accuracy with variable vertex density depending on the application. For instance, the density can be representative of the spatial acuity of a specific perception modality (Figure 4). The format of custom 3D meshes is out of the scope of this document.

Figure 4 — Example of haptic avatar body representations

Table 2 defines the list of properties of a haptic avatar: