FINAL DRAFT
International Standard

ISO/IEC
FDIS
23008-12

Information technology — High efficiency coding and media delivery in heterogeneous environments —

Part 12: Image File Format

Technologies de l'information — Codage à haute efficacité et livraison des medias dans des environnements hétérogènes —

Partie 12: Format de fichier d'image

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

Secretariat: **JISC**

Voting begins on:
2025-02-28

Voting terminates on:
2025-04-25

Reference number
ISO/IEC FDIS 23008-12:2025(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

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 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*, SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

This third edition cancels and replaces the second edition (ISO/IEC 23008-12:2022), which has been technically revised.

The main changes are as follows:

— clarification on the signalling of colour information in image items;

— support for the signalling of camera intrinsic and extrinsic matrices;

— support for progressive decoding, rendering and refinement;

— support for region annotations for image sequence or video track;

— support for renderable text items.

A list of all parts in the ISO/IEC 23008 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

The Image File Format is designed to enable the interchange of images and image sequences, as well as their associated metadata. It forms part of a family of specifications that are box-structured, and is built using tools defined in the ISO base media file format. This document specifies both structural brands that can be used with any codec and brands specific to High Efficiency Video Coding (HEVC). The file format specified in this document is referred to as the High Efficiency Image File Format (HEIF). It is suggested that HEIF be pronounced "heaff" (like heath with an ff ending). When the requirements of the HEVC-specific brands are applied, the file format can be referred to as the HEVC Image File Format.

This document is organized as follows:

Clause 5 specifies general requirements on files and file readers conforming to the Image File Format.

Clause 6 specifies the file structures for the storage of a single image and an image collection. Additionally, general requirements that shall be supported in all files using the Image File Format for the storage of a single image or an image collection are specified.

Clause 7 specifies the file structures for the storage of image sequences. Additionally, general requirements that shall be supported in all files using the Image File Format for the storage of image sequences are specified.

Clause 8 specifies the metadata structures for a single image, an image collection, and image sequences.

Clause 9 specifies enhancements to the ISO base media file format.

Clause 10 specifies structural brands for a single image and an image collection, as well as image sequences. Requirements on both files and file readers are specified.

Clause 11 specifies tools to associate annotations, e.g. metadata or images with one or more regions of an image or an image sequence.

Annex A specifies the format for storing Exif, XMP, and MPEG-7 metadata in files conforming to the Image File Format.

Annex B specifies the format for encapsulating HEVC-coded images, image collections, and image sequences according to the Image File Format. Annex B also specifies HEVC-specific brands for a single image and an image collection as well as image sequences. Requirements on both files and file readers are specified.

Annex C and Annex D specify the MIME type registration for a single image or an image collection, and image sequences, respectively, for the structural and HEVC-specific brands.

Annex E specifies the format for encapsulating AVC-coded images, image collections, and image sequences according to the Image File Format.

Annex F and Annex G specify the MIME type registration for a single image or an image collection, and image sequences, respectively, for the AVC-specific brands.

Annex H specifies the format for encapsulating JPEG-coded images, image collections, and image sequences according to the Image File Format.

Annex I contains guidelines on defining new image formats and brands.

Annex J contains informative examples of single image and image collection file structures conforming to the Image File Format.

Annex K provides examples of content encoding, file structures and player operations for progressive rendering, progressive decoding and progressive refinement with the Image File Format .

Annex L specifies the format for encapsulating VVC-coded images, image collections, and image sequences according to the Image File Format. Annex L also specifies VVC-specific brands for a single image and an image collection as well as image sequences. Requirements on both files and file readers are specified.

Annex M specifies the format for encapsulating EVC-coded images, image collections, and image sequences according to the Image File Format. Annex M also specifies EVC-specific brands for a single image and an image collection as well as image sequences. Requirements on both files and file readers are specified.

Annex N contains considerations on privacy and security relating to the use of the Image File Format.

Information technology — High efficiency coding and media delivery in heterogeneous environments —

Part 12: Image File Format

1 Scope

This document specifies the Image File Format, an interoperable storage format for a single image, a collection of images, and sequences of images.

The format defined in this document is built on tools defined in ISO/IEC 14496-12 and enables the interchange, editing, and display of images, as well as the carriage of metadata associated with those images. The Image File Format defines structures used to contain metadata, how to link that metadata to the images, and defines how metadata of certain forms is carried.

This document also specifies brands for the storage of images and image sequences conforming to High Efficiency Video Coding (HEVC), Advanced Video Coding (AVC), JPEG, Versatile Video Coding (VVC) and Essential Video Coding (EVC).

NOTE The storage of HEVC, AVC, VVC and EVC video sequences is out of scope and is provided in ISO/IEC 14496-15.

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 10918-1, *Information technology — Digital compression and coding of continuous-tone still images: Requirements and guidelines*

ISO/IEC 14496-10, *Information technology — Coding of audio-visual objects — Part 10: Advanced Video Coding*

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

ISO/IEC 14496-15, *Information technology — Coding of audio-visual objects — Part 15: Carriage of network abstraction layer (NAL) unit structured video in the ISO base media file format*

ISO/IEC 23008-2, *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 2: High efficiency video coding*

ISO/IEC 23090-3, *Information technology — Coded representation of immersive media — Part 3: Versatile video coding*

ISO/IEC 23090-7, *Information technology — Coded representation of immersive media — Part 7: Immersive media metadata*

ISO/IEC 23094-1, *Information technology — General video coding — Part 1: Essential video coding*

IETF RFC 3937, *A Uniform Resource Name (URN) Namespace for the International Press Telecommunications Council (IPTC)*

3 Terms, definitions, abbreviated terms and symbols

3.1 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 14496-12, ISO/IEC 14496-15 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
alternate group
group of *entities* (3.1.11) that are alternatives to each other and out of which only one should be selected for processing

3.1.2
associated image item
image item (3.1.19) that is associated with the *item property* (3.1.29) through the `ItemPropertiesBox`

3.1.3
auxiliary image
image (3.1.17) that may not be intended to be displayed but provides supplemental information, such as transparency data, complementing a respective *master image* (3.1.30)

3.1.4
coded image
coded representation of an *image* (3.1.17)

3.1.5
coded image item
item (3.1.28) whose data is a *coded image* (3.1.4)

3.1.6
crop-rotate-mirror derived image item
derived image item (3.1.8) of type `'iden'` that is not associated with any other types of *essential item properties* (3.1.12) than `'colr'`, `'irot'`, `'clap'`, and `'imir'`

3.1.7
derived image
representation of an *image* (3.1.17) as an *operation* (3.1.34) on other images

3.1.8
derived image item
item (3.1.28) whose data is a *derived image* (3.1.7)

3.1.9
derived region item
item (3.1.28) whose data defines a *region* (3.1.43) within an *image item* (3.1.19), with which the item is associated via item reference, as an *operation* (3.1.34) on other *region items* (3.1.45)

3.1.10
descriptive item property
item property (3.1.29) that describes rather than transforms the associated item

3.1.11
entity
item or track

3.1.12
essential item property
item property (3.1.29) that readers are required to process

3.1.13
font item
item (3.1.28) whose data is font data

Note 1 to entry: Font items associated with text items provide font for rendering the text data.

3.1.14
HEVC image item
image item (3.1.19) of type `'hvc1'` or `'lhv1'`

3.1.15
hidden image
image (3.1.17) that is not intended to be displayed

3.1.16
hidden sample
sample that is not intended to be displayed

3.1.17
image
one or more arrays of pixels of different colour components described by an *image item* (3.1.19) or a sample

3.1.18
image collection
set of *images* (3.1.17) stored as *items* (3.1.28) of a single file according to this document

3.1.19
image item
coded image item (3.1.5) or *derived image item* (3.1.8)

3.1.20
Image File Format
file format specified by this document

3.1.21
image property
item property (3.1.29) for an *image item* (3.1.19)

3.1.22
image sequence
sequence of *coded images* (3.1.4) which may be associated with advisory timing and in which images may use *inter prediction* (3.1.25)

3.1.23
image sequence track
track that contains an *image sequence* (3.1.22)

3.1.24
input image
image (3.1.17) that is used as an input for the *operation* (3.1.34) of the *derived image item* (3.1.8)

3.1.25
inter prediction
prediction derived in a manner that is dependent on data elements (e.g. sample values or motion vectors) of *images* (3.1.17) other than the current image

3.1.26
intra coding
coding of an *image* (3.1.17) that may use *intra prediction* (3.1.27) and does not use *inter prediction* (3.1.25)

3.1.27
intra prediction
prediction derived from only data elements (e.g. sample values) of the same decoded image

3.1.28
item
data that does not require timed processing, as opposed to sample data, and is described by the boxes contained in a `MetaBox`

3.1.29
item property
descriptive or transformative information about an *item* (3.1.28) as stored in the item properties array

3.1.30
master image
image that is stored as an *item* (3.1.28) and is not an *auxiliary image* (3.1.3) or a thumbnail image

3.1.31
master image sequence
sequence of images that is stored as an *image sequence track* (3.1.23) and is not an *auxiliary image* (3.1.3) sequence or a thumbnail *image sequence* (3.1.22)

3.1.32
metadata item
item (3.1.28) containing metadata that may for example describe an *image item* (3.1.19)

Note 1 to entry: ISO/IEC 14496-12 uses the terms item and metadata item interchangeably to refer to an item of any type. This document overrides the metadata item definition of the ISO base media file format.

3.1.33
non-essential item property
item property (3.1.29) that readers are allowed to ignore

3.1.34
operation
for a *derived image item* (3.1.8), manipulation, identified by the item type, that produces a *reconstructed image* (3.1.41) from a set of *input images* (3.1.24). For a *derived region item* (3.1.9), manipulation, identified by the item type, that produces the shape, position and size of regions of an *image* (3.1.17) from a set of regions from input *region items* (3.1.45).

3.1.35
output image
image (3.1.17) that results when the reconstructed image of the *image item* (3.1.19) is transformed according to the *transformative item properties* (3.1.50) of the image item

3.1.36
pre-derived coded image
coded image (3.1.4) that has been derived from one or more other images

3.1.37
predictively coded image item
image item (3.1.19) that has a decoding dependency to one or more other *coded image items* (3.1.5)

3.1.38
progressive decoding
decoding a bitstream with a single decoder instance in successive steps where each step improves the perceived image quality over that of the previous step

3.1.39
progressive refinement
progressive rendering (3.1.40) of an *image item* (3.1.19) or sample in a file while downloading the file

3.1.40
progressive rendering
displaying an *image item* (3.1.19) or a sample in successive steps where each step improves the perceived image quality over that of the previous step and is superimposed over the *image* (3.1.17) of the previous step in the same displaying window

Note 1 to entry: A progressive rendering step can improve the perceived image quality over the complete image as a whole, or region by region resulting in a region-wise progressive rendering.

3.1.41
reconstructed image
image (3.1.17) that results when the *coded image item* (3.1.5) is decoded or when the *operation* (3.1.34) of the *derived image item* (3.1.8), if any, is applied

3.1.42
reference image
image (3.1.17) that may be used as a reference for *inter prediction* (3.1.25) of another image

3.1.43
region
area represented by a shape, position and size encompassing a part of an *image* (3.1.17)

3.1.44
region annotation
metadata or data representing an annotation associated with a *region* (3.1.43)

3.1.45
region item
item (3.1.28) whose data defines a *region* (3.1.43) within an *image item* (3.1.19) with which the item is associated via item reference

3.1.46
region track
track whose samples define a *region* (3.1.43) within samples of another track with which the track is associated via track reference

3.1.47
source image item
image item (3.1.19) referred to by the `'dimg'` item reference from the *derived image item* (3.1.8) or from another derived image item that is a source image item for the derived image item

Note 1 to entry: In other words, an image item is a source image item for a derived image item when it is required for deriving the output image of the derived image item.

Note 2 to entry: The definition of the source image item is recursive: an image item is a source image item for a particular derived image item, when the output image of the image item is used as an input image for any derived image item in the `'dimg'`-item-reference-linked chain of derived image items ending at that particular derived image item, inclusive.

3.1.48
text item
item (3.1.28) whose data contains human-readable text content that would require a text layout and rendering engine if displayed

3.1.49
time-parallel sample
sample in the reference track that has the same or, when a sample with the same decoding time is not available, the closest preceding decoding time relative to that of the particular sample in the particular track

3.1.50
transformative item property
item property (3.1.29) that transforms the reconstructed representation of the item content

Note 1 to entry: A transformative item property may for example specify rotation by 90, 180, or 270 degrees of a reconstructed image of an image item.

3.1.51
unique ID
identifier for either an item, an entity group or a track that fulfils the requirements of the `'unif'` brand

Note 1 to entry: Requirements on the `'unif'` brand are specified in ISO/IEC 14496-12.

3.1.52
visual context
visual rendering surface such as a screen buffer, which may already contain visual material, and onto which an *image* (3.1.17) can be rendered

3.2 Abbreviated terms

ASCII	American Standard Code for Information Interchange
AVC	Advanced Video Coding (Rec. ITU-T H.264 | ISO/IEC 14496-10)
DCF	Design rule for Camera File system (JEITA CP-3461)
EVC	Essential Video Coding (ISO/IEC 23094-1)
Exif	Exchangeable Image File Format (JEITA CP-3451)
HDR	high dynamic range
HEIF	High Efficiency Image File Format (this document: ISO/IEC 23008-12)
HEVC	High Efficiency Video Coding (Rec. ITU-T H.265 | ISO/IEC 23008-2)
MD5	Message Digest algorithm 5
MIME	Multi-purpose Internet Mail Extensions
NAL	network abstraction layer
PPS	picture parameter set
SEI	supplemental enhancement information
SPS	sequence parameter set
TIFF	Tagged Image File Format
URI	Uniform Resource Identifier
URN	Uniform Resource Name
UTF-8	Universal Character Set Transformation Format — 8-bit
VCL	video coding layer
VPS	video parameter set
VVC	Versatile Video Coding (Rec. ITU-T H.266 | ISO/IEC 23090-3)