FINAL DRAFT

INTERNATIONAL STANDARD

ISO/IEC FDIS 18181-1

ISO/IEC JTC 1/SC 29

Secretariat: **JISC**

Voting begins on: **2021-10-14**

Voting terminates on: **2021-12-09**

Information technology — JPEG XL Image Coding System —

Part 1: Core coding system

Technologies de l'information - Système de codage d'images JPEG

iTeh ST^{XL}NDARD PREVIEW Partie 1: Système de codage de noyau (standards.iteh.ai)

ISO/IEC FDIS 18181-1 https://standards.iteh.ai/catalog/standards/sist/2141a071-847c-436f-bac1f26c71e8dfcd/iso-iec-fdis-18181-1

RECIPIENTS OF THIS DRAFT ARE INVITED TO SUBMIT, WITH THEIR COMMENTS, NOTIFICATION OF ANY RELEVANT PATENT RIGHTS OF WHICH THEY ARE AWARE AND TO PROVIDE SUPPORTING DOCUMENTATION.

IN ADDITION TO THEIR EVALUATION AS BEING ACCEPTABLE FOR INDUSTRIAL, TECHNO-LOGICAL, COMMERCIAL AND USER PURPOSES, DRAFT INTERNATIONAL STANDARDS MAY ON OCCASION HAVE TO BE CONSIDERED IN THE LIGHT OF THEIR POTENTIAL TO BECOME STAN-DARDS TO WHICH REFERENCE MAY BE MADE IN NATIONAL REGULATIONS.



Reference number ISO/IEC FDIS 18181-1:2021(E)

iTeh STANDARD PREVIEW (standards.iteh.ai)

ISO/IEC FDIS 18181-1 https://standards.iteh.ai/catalog/standards/sist/2141a071-847c-436f-bac1f26c71e8dfcd/iso-iec-fdis-18181-1



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2021

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

Forew	rord	v
Introd	luction	vi
1	Scope	
2	Normative references	
3	Terms and definitions3.1Data storage3.2Inputs3.3Processes3.4Image organization3.5DCT	1 2 2 3 4 5
4	Abbreviated terms	
5	Conventions5.1Mathematical symbols5.2Functions5.3Operators5.4Pseudocode	
6	Functional concepts 6.1 Image organization 6.2 Group splitting STANDARD PREVIEW 6.3 Codestream and bitstream 6.4 Multiple frames 6.5 Mirroring	8 8 9 10 10
7 8	Encoder requirements ISO/IEC FDIS 18181-1 https://standards.iteh.ai/catalog/standards/sist/2141a071-847c-436f-bac1- Decoder requirements PriceThe8tfcd/ison-jerg-fdis-18181-1	
9	Codestream. 9.1 Syntax	10 10 11 11 11 11 11 11 11 11
10	Decoding process	
Annex	x A (normative) Headers	
Annex	x B (normative) ICC profile	
Annex	c (normative) Frames	
Annex	x D (normative) Entropy decoding	
Annex	x E (normative) Weighted Predictor	
Annex	x F (normative) Adaptive quantization	70
Annex	x G (normative) Chroma from luma	71

ISO/IEC FDIS 18181-1:2021(E)

Annex H (normative) Extensions	.72
Annex I (normative) Integral transforms	.73
Annex J (normative) Restoration filters	.84
Annex K (normative) Image features	.87
Annex L (normative) Colour transforms	.92
Annex M (informative) Encoder overview	.98
Bibliography	L 01

iTeh STANDARD PREVIEW (standards.iteh.ai)

ISO/IEC FDIS 18181-1 https://standards.iteh.ai/catalog/standards/sist/2141a071-847c-436f-bac1f26c71e8dfcd/iso-iec-fdis-18181-1

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.iso.org/directiv

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 www.iso.org/patents) or the IEC list of patent declarations received (see www.iso.org/patents) or the IEC list of patent declarations received (see www.iso.org/patents) or the IEC list of patent declarations received (see www.iso.org/patents) or the IEC list of patent declarations received (see www.iso.org/patents) or the IEC list of patent declarations received (see https://www.iso.org/patents) or the IEC list of patent declarations received (see https://www.iso.org/patents) or the IEC list of patent declarations received (see https://www.iso.org/patents) or the IEC list of patent declarations received (see https://www.iso.org/patents) or the IEC list of patent declarations received (see https://www.iso.org/patents) or the IEC list of patent declarations received (see https://www.iso.org/patents) or the IEC list of patent declarations received (see https://www.iso.org/patents) or the IEC list of patent declarations received (see https://www.iso.org/patents) or the IEC list of patents) or th

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 18181 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 <u>www.iso.org/members.html</u> and <u>www.iec.ch/national-committees</u>.

Introduction

The International Organization for Standardization (ISO) and 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.

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.

iTeh STANDARD PREVIEW (standards.iteh.ai)

ISO/IEC FDIS 18181-1 https://standards.iteh.ai/catalog/standards/sist/2141a071-847c-436f-bac1f26c71e8dfcd/iso-iec-fdis-18181-1

Information technology — JPEG XL Image Coding System —

Part 1: Core coding system

1 Scope

This document defines a set of compression methods for coding one or more images of bi-level, continuous-tone greyscale, or continuous-tone colour, or multichannel digital samples.

This document:

- specifies decoding processes for converting compressed image data to reconstructed image data;
- specifies a codestream syntax containing information for interpreting the compressed image data;
- provides guidance on encoding processes for converting source image data to compressed image data.

2 Normative references STANDARD PREVIEW

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 15076-1:2010, Image technology colour management Architecture, profile format and data structure — Part 1: Based on ICC.1:2010

ISO/IEC 60559, Information technology — Microprocessor Systems — Floating-Point arithmetic

IEC 61966-2-1, Multimedia systems and equipment — Colour measurement and management — Part 2-1: Colour management — Default RGB colour space — sRGB

Rec. ITU-R BT.2100-2, Image parameter values for high dynamic range television for use in production and international programme exchange

Rec. ITU-R BT.709-6, Parameter values for the HDTV standards for production and international programme exchange

SMPTE ST 428-1, D-Cinema Distribution Master — Image Characteristics

3 Terms and definitions

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

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

- ISO Online browsing platform: available at <u>https://www.iso.org/obp</u>
- IEC Electropedia: available at <u>https://www.electropedia.org/</u>

3.1 Data storage

3.1.1

byte

8 consecutive bits encoding a value between 0 and 255

3.1.2

big endian

value representation with bytes in most to least-significant order

3.1.3

bitstream

sequence of bytes from which bits are read starting from the least-significant bit of the first byte

3.1.4

codestream

bitstream representing compressed image data

3.1.5

bundle

structured data consisting of one or more fields

3.1.6

field

numerical value or bundle, or an array of either

iTeh STANDARD PREVIEW

3.1.7 histogram

array of unsigned integers representing a probability distribution, used for entropy coding

3.1.8

ISO/IEC FDIS 18181-1 https://standards.iteh.ai/catalog/standards/sist/2141a071-847c-436f-bac1set unordered collection of elements f26c71e8dfcd/iso-iec-fdis-18181-1

3.2 Inputs

3.2.1

pixel

vector of dimension corresponding to the number of channels, consisting of samples

3.2.2

sample

integer or real value, of which there is one per channel per pixel

3.2.3

grevscale

image representation in which each pixel is defined by a single sample representing intensity (either luminance or luma depending on the ICC profile)

3.2.4

continuous-tone image

image having samples consisting of more than one bit

3.2.5

opsin

photosensitive pigments in the human retina, having dynamics approximated by the XYB colour space

3.2.6

burst

sequences of images typically captured with identical settings

3.2.7

animation

series of pictures and timing delays to display as a video medium

3.2.8

composite

series of images that are superimposed

3.2.9

frame

single image (possibly part of a burst or animation or composite)

3.2.10

preview

lower-fidelity rendition of one of the frames (e.g. lower resolution), or a frame that represents the entire content of all frames

3.3 Processes

3.3.1

decoding process

process which takes as its input a codestream and outputs a continuous-tone image

3.3.2

decoder embodiment of a decoding process ANDARD PREVIEW

3.3.3

(standards.iteh.ai)

encoding process

process which takes as its input continuous-tone image(s) and outputs compressed image data in the form of a codestreamos://standards.iteh.ai/catalog/standards/sist/2141a071-847c-436f-bac1-

f26c71e8dfcd/iso-iec-fdis-18181-1

3.3.4

encoder embodiment of an encoding process

3.3.5

lossless

descriptive term for encoding and decoding processes in which the output of a decoding procedure is identical to the input to the encoding procedure

3.3.6

lossy

descriptive term for encoding and decoding processes which are not lossless

3.3.7

upsampling

procedure by which the (nominal) spatial resolution of a channel is increased

3.3.8

downsampling

procedure by which the spatial resolution of a channel is reduced

3.3.9

entropy encoding

lossless procedure designed to convert a sequence of input symbols into a sequence of bits such that the average number of bits per symbol approaches the entropy of the input symbols

3.3.10

entropy encoder

embodiment of an entropy encoding procedure

3.3.11

entropy decoding

lossless procedure which recovers the sequence of symbols from the sequence of bits produced by the entropy encoder

3.3.12

entropy decoder

embodiment of an entropy decoding procedure

3.3.13

Gabor-like transform

convolution with default or signalled 3x3 kernel for deblocking

3.3.14

tick

unit of time such that animation frame durations are integer multiples of the tick duration

3.4 Image organization

3.4.1

grid

2-dimensional array; a [x, y] means addressing an element of grid a at row y and column x. Where so specified, addressing elements with coordinates outside of bounding rectangle (x < 0, or y < 0, or x >= width, or y >= height) is allowed

3.4.2

iTeh STANDARD PREVIEW

sample grid

for all calstandards itch ai) of acordin

common coordinate system for all samples of an image, with top-left coordinates (0, 0), the first coordinate increasing towards the right, and the second increasing towards the bottom

<u>ISO/IEC FDIS 18181-1</u>

3.4.3 channel https://standards.iteh.ai/catalog/standards/sist/2141a071-847c-436f-bac1f26c71e8dfcd/iso-iec-fdis-18181-1

component

rectangular array of samples having the same designation, regularly aligned along a sample grid

3.4.4

rectangle

rectangular area within a channel or grid

3.4.5

width

width in samples of a sample grid or a rectangle

3.4.6

height

height in samples of a sample grid or a rectangle

3.4.7

raster order

access pattern from left to right in the top row, then in the row below and so on

3.4.8

naturally aligned

positioning of a power-of-two sized rectangle such that its top and left coordinates are divisible by its width and height, respectively

3.4.9 block

naturally aligned square rectangle covering up to 8 × 8 input pixels

3.4.10

group

naturally aligned square rectangle covering up to $2^n \times 2^n$ (with n between 7 and 10, inclusive) input pixels

3.4.11

table of contents

data structure that enables seeking to a group or the next frame within a codestream

3.4.12

section

part of the codestream with an offset and length that are stored in a frame's table of contents

3.5 DCT

3.5.1

coefficient input value to the inverse DCT

3.5.2

quantization

method of reducing the precision of individual coefficients

3.5.3

variable-size rectangle of input pixels NDARD PREVIEW

3.5.4

(standards.iteh.ai)

dct_block

an array with 64 elements corresponding to DGT coefficients of a (8 × 8) block

3.5.5 https://standards.iteh.ai/catalog/standards/sist/2141a071-847c-436f-bac1f26c71e8dfcd/iso-iec-fdis-18181-1

var-DCT

lossy encoding of a frame that applies DCT to varblocks

3.5.6

LF coefficient

lowest frequency DCT coefficient, containing the average value of a block or the lowest-frequency coefficient within the 8×8 rectangle of a varblock of size greater than 8×8

3.5.7

HF coefficients

all DCT coefficients apart from the LF coefficients, i.e. the high frequency coefficients

3.5.8

pass

data enabling decoding of successively higher resolutions

3.5.9

LF group

 $2^{n} \times 2^{n}$ LF values from a naturally aligned rectangle covering up to $2^{n+3} \times 2^{n+3}$ input pixels

3.5.10

quantization weight

factor that a quantized coefficient is multiplied by prior to application of the inverse DCT in the decoding process

3.5.11

channel decorrelation

method of reducing total encoded entropy by removing correlations between channels

3.5.12

channel correlation factor

factor by which a channel should be multiplied by before adding it to another channel to undo the channel decorrelation process

4 Abbreviated terms

DCT: discrete cosine transform (DCT-II as specified in <u>I.2</u>)

IDCT: inverse discrete cosine transform (DCT-III as specified in <u>I.2</u>)

LF: N / $8 \times M$ / 8 square of lowest frequency coefficients of N $\times M$ DCT coefficients

RGB: additive colour model with red, green, blue channels

LMS: absolute colour space representing the response of cone cells in the human eye

XYB: absolute colour space based on gamma-corrected LMS, in which X is derived from the difference between L and M, Y is an average of L and M (behaves similarly to luminance), and B is derived from the S ("blue") channel

5 Conventions

5.1 Mathematical symbols Teh STANDARD PREVIEW

[a,	b],	(c,	d),	[e,	f)	closed or open or half open intervals containing all integers or real numbers x (depending on context) such that
						$a \le x \le b$, $\underline{eSO/IEC \le FBIS e_8 + 8 \times 1} \le f$.
{a,	b, (с}				ordered sequence of elements. 1811-1
п						the smallest positive zero of the sine function

5.2 Functions

sqrt(x)	square root, such that $(sqrt(x))^2 = x$ and $sqrt(x) >= 0$. Undefined for $x < 0$.
cbrt(x)	cube root, such that $(cbrt(x))^3 = x$.
cos(r)	cosine of the angle r (in radians)
erf(x)	Gauss error function: erf(x) = $\frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-t^2} dt$
log(x)	natural logarithm of x. Undefined for $x \le 0$.
log2(x)	base-two logarithm of x. Undefined for x <= 0.
floor(x)	the largest integer that is less than or equal to x
ceil(x)	the smallest integer that is greater than or equal to x
abs(x)	absolute value of x: equal to -x if x < 0, otherwise x
sign(x)	sign of x, 0 if x is 0, +1 if x is positive, -1 if x is negative
UnpackSigned(u)	equivalent to u / 2 if u is even, and -(u + 1) / 2 if u is odd

ISO/IEC FDIS 18181-1:2021(E)

clamp(x, lo, hi)	<pre>equivalent to min({max({lo, x}), hi})</pre>
InterpretAsF16(u)	the real number resulting from interpreting the unsigned 16-bit integer ${\tt u}$ as a binary16 floating-point number representation (cf. ISO/IEC 60559)
InterpretAsF32(u)	the real number resulting from interpreting the unsigned 32-bit integer ${\tt u}$ as a binary32 floating-point number representation (cf. ISO/IEC 60559)
len(a)	length (number of elements) of array a
sum(a)	sum of all elements of the array/tuple/sequence a
max(a)	maximal element of the array/tuple/sequence a
min(a)	smallest element of the array/tuple/sequence a

5.3 Operators

This document uses the operators defined by the C++ programming language ^[2], with the following differences:

×	multiplication
×=	a ×= b is equivalent to a = a × b
/	division of real numbers without truncation or rounding. Division by zero is undefined.
ХÄ	exponentiation, x to the power of yrds.iteh.ai)
<<	left shift: x << s is defined as x/re2°FDIS 18181-1
>>	https://standards.iteh.ai/catalog/standards/sist/2141a071-847c-436f-bac1- right shift: $x >> s$ is defined as floor ($x = -\frac{1}{100} e^{-1}$) (18181-1
Umod	a Umod d is the unique integer r in [0, d) for which a == r + $q \times d$ for a suitable integer q
Idiv	a Idiv b is equivalent to a / b, rounded towards zero to an integer value

The order of precedence for these operators is listed below in descending order. If several operators appear in the same line, they have equal precedence. When several operators of equal precedence appear at the same level in an expression, evaluation proceeds according to the associativity of the operator (either from right to left or from left to right).

Operators	Type of operation	Associativity
++x,x	prefix increment/decrement	right to left
ХĂ	exponentiation	right to left
!, ~	logical/bitwise NOT	right to left
×, /, Idiv, Umod	multiplication, division, integer division, remainder	left to right
+, -	addition and subtraction	left to right
<<, >>	left shift and right shift	left to right
< , >, <=, >=	relational	left to right
=	assignment	right to left
+=, -=, ×=	compound assignment	right to left

5.4 Pseudocode

This document describes functionality using pseudocode formatted as follows:

// Informative comment
var = u(8); // Defined in 9.2.1
if (var == 1) return; // Stop executing this code snippet
[[Normative specification: var != 0]]
(out1, out2) = Function(var, kConstant);

Variables such as var are typically referenced by text outside the source code.

The semantics of this pseudocode are those of the C++ programming language [2], with the following exceptions:

- Symbols from 5.1 and functions from 5.2 are allowed;
- Multiplication, division, remainder and exponentiation are expressed as specified in <u>5.3</u>;
- Functions can return tuples which unpack to variables as in the above example;
- [[]] enclose normative directives specified using prose;
- All integers are stored using two's complement;
- Expressions and variables of which types are omitted, are understood as real numbers.

Where unsigned integer wrap around and truncated division are required, $\tt Umod$ and $\tt Idiv$ (see 5.3) are used for those purposes.

Numbers with a 0x prefix are in base 16 (hexadecimal), and apostrophe (') characters inside them are understood to have no effect. (standards.iteh.ai)

EXAMPLE 0x0001'0000 == 65536.

ISO/IEC FDIS 18181-1

6 Functional concepts^{5://standards.iteh.ai/catalog/standards/sist/2141a071-847c-436f-bac1f26c71e8dfcd/iso-iec-fdis-18181-1}

6.1 Image organization

A channel is defined as a rectangular array of (integer or real) samples regularly aligned along a sample grid of width sample positions horizontally and height sample positions vertically. The number of channels may be 1 to 4099 (see num_extra_channels in $\underline{A.6}$).

A pixel is defined as a vector of dimension corresponding to the number of channels, consisting of samples with a position matching that of the pixel. The index of a sample is numbered from 0 to number of channels - 1.

An image is defined as the two-dimensional array of pixels, and its width is width and height is height. Unless otherwise mentioned, channels are accessed in the following "raster order": left to right column within the topmost row, then left to right column within the row below the top, and so on until the rightmost column of the bottom row.

6.2 Group splitting

Channels are logically partitioned into naturally-aligned groups of $kGroupDim \times kGroupDim$ samples. The effective dimension of a group (i.e. how many pixels to read) can be smaller than kGroupDim for groups on the right or bottom of the image. The decoder ensures the decoded image has the dimensions specified in SizeHeader by cropping at the right and bottom as necessary. Unless otherwise specified, kGroupDim is 256.

LF groups likewise consist of $kGroupDim \times kGroupDim LF$ samples, with the possibility of a smaller effective size on the right and bottom of the image.

Groups can be decoded independently. A 'table of contents' stores the size (in bytes) of each group to allow seeking to any group. An optional permutation allows groups to be arranged in arbitrary order within the codestream.

EXAMPLE Figure 1 shows an example of the HF groups and LF groups of an image.

Frame: 2970×1868 pixels

HF groups:

11×7 groups of 256×256 pixels,

1×7 groups of 154×256 pixels,

7×1 groups of 256×76 pixels,

iTeh STANDARD PREVIEW (standards.iteh.ai)

ISO/IEC FDIS 18181-1 1 group of 154×76 pixelss://standards.iteh.ai/catalog/standards/sist/2141a071-847c-436f-bac1f26c71e8dfcd/iso-iec-fdis-18181-1

LF groups:

1 group of 256×233 LF coefficients (covering 2048×1868 pixels), 1 group of 116×233 LF coefficients (covering 922×1868 pixels)

Figure 1 — Group splitting example

6.3 **Codestream and bitstream**

A bitstream is a finite sequence of bytes. A codestream is a bitstream that represent compressed image data and metadata. N bytes can also be viewed as 8 × N bits. The first 8 bits are the bits constituting the first byte, in least to most significant order, the next eight bits (again in least to most significant order) constitute the second byte, and so on. Unless otherwise specified, bits are read from the codestream as specified in <u>9.2.1</u>.

NOTE Ordering bits from least to most significant allows using special CPU instructions to isolate the leastsignificant bits.

Subsequent Annexes or subclauses indicate some elements of the codestream are byte-aligned. For such elements, the decoder takes actions before and after reading the element as follows. Immediately before encountering the element, the decoder invokes ZeroPadToByte() (9.2.9). After finishing reading the element, the decoder invokes ZeroPadToByte() (9.2.9).