International Standard

ISO/IEC 21122-2

Information technology — JPEG XS low-latency lightweight image coding system —

Part 2: Profiles and buffer models

Technologies de l'information — Système de codage d'images léger à faible latence JPEG XS —

Partie 2: Profils et modèles tampons

Third edition
2024-08

Reference number
ISO/IEC 21122-2:2024(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 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*.

This third edition cancels and replaces the second edition (ISO 21122-2:2022), which has been technically revised. It also incorporates the Amendment ISO/IEC 21122-2:2022/Amd 1:2022.

The main changes compared to the previous edition are:

— addition of conformance points for new profiles;

— addition of the TDC 444.12 and TDC MLS 444.12 profiles for compression of image sequences;

— addition of the CHigh 444.12 profile;

— addition of the MLS.16 profile;

— addition of the frame buffer bandwidth levels and model.

A list of all parts in the ISO/IEC 21122 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 https://www.iso.org/members.html and https://www.iec.ch/national-committees.

Introduction

This document is part of a series of standards for a low-latency lightweight image coding system, denoted as JPEG XS. While ISO/IEC 21122-1 specifies a full set of compression coding tools needed to satisfy all the requirements of JPEG XS, a targeted application can often work with a simpler and reduced set of coding tools, and with or without tighter constraints, to meet its targeted goals. For this reason, profiles, levels, and sublevels are defined in this document. These three concepts facilitate partial and reduced complexity implementations of ISO/IEC 21122-1 depending on specific application use cases and requirements, while also safeguarding interoperability.

This document specifies a limited number of profiles to represent interoperability subsets of the codestream syntax specified in ISO/IEC 21122-1 with each profile serving specific application use cases. In other word, profiles select a subset of the available coding tools. In addition, levels and sublevels provide limits to the maximum throughput in respectively the decoded (spatial/pixel) and the encoded (codestream) domains. In this way, profiles, levels and sublevels allow designing cost-efficient implementations that serve the needs of the desired applications.

A major requirement of JPEG XS is to allow low end-to-end latency, limited to a fraction of the frame size. To ensure this low-latency property, this document also specifies a buffer model, consisting of a decoder model and a transmission channel model. The models show the interaction of a hypothetical reference decoder, including its smoothing buffer with a constant bitrate channel feeding this buffer. The size of the decoder smoothing buffer is computed from the profile, level, and sublevel. Codestreams are formed such that the buffer of a decoder, operating according to this buffer model, never overflows or underflows. In effect, the buffer model provides encoders with the necessary information to generate codestreams that can be decoded by an arbitrary decoder implementation, ensuring system interoperability.

In addition to the size of the decoder smoothing buffer, end-to-end latency also depends on the latency inherent to each processing step of the encoding-decoding chain whose methods are described in ISO/IEC 21122-1. To help implementers estimate the latency of their device, this document gives extra information on the minimum latency that can be achieved by the different methods described in ISO/IEC 21122-1.

Information technology — JPEG XS low-latency lightweight image coding system —

Part 2: Profiles and buffer models

1 Scope

This document defines several subsets of the syntax specified in ISO/IEC 21122-1 as profiles. It also defines lower bounds on the throughput in the decoded domain via levels and the encoded domain via sublevels that a conforming decoder implementation shall support. Furthermore, it defines a buffer model to ensure interoperability between implementations in the presence of a latency constraint.

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 21122-1, *Information technology — JPEG XS low-latency lightweight image coding system — Part 1: Core coding system*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 21122-1 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
blanking codestream fragment
placeholder codestream fragment representing blanking periods

3.2
buffer model
combination of a *decoder model* (3.8) and a *channel model* (3.4) whose behaviour can be defined by a set of parameters

3.3
buffer model instance
specific configuration of a *buffer model* (3.2) specified by the assignment of well-defined values to the *buffer model* parameters

3.4
channel model
model describing the temporal behaviour of the *transmission channel* (3.26) connecting an encoder and a decoder

3.5
coded codestream fragment
continuous sequence of bits in the codestream containing exactly one packet body and a well-defined number of packet headers, markers and marker segments

3.6
codestream fragment
either *coded codestream fragment* (3.5) or *blanking codestream fragment* (3.1)

3.7
cycle
clock cycle
single clock period of an encoder or decoder clocked implementation

3.8
decoder model
combination of a *decoder unit* (3.10) and a *decoder smoothing buffer* (3.9)

3.9
decoder smoothing buffer
memory buffer that is used to level out changes in the number of bits read by a *decoder unit* (3.10) per time unit

3.10
decoder unit
module reading a variable number of bits (from the smoothing buffer) per time unit to generate decoded output pixels at a fixed output rate

3.11
encoder model
combination of an *encoder unit* and an *encoder smoothing buffer* (3.12)

3.12
encoder smoothing buffer
memory buffer that is used to level out changes in the number of bits generated by an *encoder unit* (3.13) per time unit

3.13
encoder unit
module transforming a sequence of input pixels with constant rate into a conforming codestream, producing a bit sequence with variable number of bits generated per time unit

3.14
fill level
number of bits stored in the encoder or *decoder smoothing buffer* (3.9)

3.15
horizontal blanking period
timespan expressed in units of the grid point sampling rate between the last pixel of an image line — not being the last line of an image — and the first pixel of the next image line

3.16
level
defined set of constraints on the number of decoded samples to be processed by an encoder or decoder, both in the spatial and temporal dimensions

Note 1 to entry: The same set of levels is defined for all profiles. Individual implementations may, within the specified constraints, support a different *level* for each supported *profile* (3.19)

3.17
nominal bits per pixel value
mean number of bits allocated per encoded pixel which is used to derive the *sublevel* constraints by assuming an image with well-defined dimensions and frame rate derived from the *level*

3.18
pixel
samples of all components at a single *sampling grid point* (3.20)

3.19
profile
specified subset of the codestream syntax together with admissible parameter values

3.20
sampling grid point
position on the sample grid, specified by integer horizontal and vertical offset relative to the origin of the sample grid

3.21
smoothing buffer unit
level- and *sublevel*-dependent number of bits by which the smoothing buffer size of the *decoder model* is specified

3.22
start of transmission
SoT
time at which the *transmission channel* starts transmission relative to the start of encoding of the first *codestream fragment* of a codestream

3.23
sublevel
defined set of constraints on the amount of codestream bits to be processed by an encoder or decoder, per unit of time, per column, and per image

Note 1 to entry: The same set of sublevels is defined for all profiles. Individual implementations may, within the specified constraints, support a different *sublevel* for each supported profile

3.24
TDC disabled codestream
codestream that contains zero `SLI` markers

Note 1 to entry: See ISO/IEC 21122-1.

3.25
TDC enabled codestream
codestream that contains one or more `SLI` markers

Note 1 to entry: See ISO/IEC 21122-1.

3.26
transmission channel
facility transferring bits from a source entity to a target entity

3.27
transmission channel capacity
maximum number of bits per time unit that a *transmission channel* (3.26) can transfer from a source entity to a target entity

3.28
vertical blanking period
timespan expressed in units of the grid point sampling rate between the last line of an image — including the *horizontal blanking periods (3.15)* — and the first line of the next image

4 Abbreviated terms

bpp	bits per pixel
CBR	constant bit rate
CFA	colour filter array
DWT	discrete wavelet transform
FBB	frame buffer bandwidth
IDWT	inverse discrete wavelet transform
IRCT	inverse reversible colour transform
MLS	mathematically lossless
RCT	reversible colour transform
RGB	red green blue
TDC	temporal differential coding (see ISO/IEC 21122-1)
VBR	variable bit rate

5 Symbols

B_{r}	number of bits required to encode a bitplane count in raw
B_{w}	nominal overall bit precision of the wavelet coefficients
$B[i]$	precision in bits of component i
C_{pih}	colour transformation type
$\mathrm{c}'[\mathrm{p},\lambda,b,x]$	wavelet coefficient in precinct p, line λ, band b and position x
C_{s}	width of precincts other than the rightmost precinct in sample grid points
C_{w}	width of precincts in multiples of 8 LL subsampled band sample grid points
$C(i)$	the i -th codestream in a sequence of codestreams
D_{c2d}	number of clock cycles between the first bit written into the decoder smoothing buffer and the decoding start of the first fragment of the stream of codestream fragments
$F_{\mathrm{first}}(i)$	first fragment of codestream $C(i)$
$F_{\mathrm{last}}(i)$	last fragment of codestream $C(i)$
F_{q}	number of fractional bits in the representation of wavelet coefficients
H_{f}	height of the image in sampling grid points
H_{p}	height of a precinct in lines
H_{max}	maximum image height in sampling grid points
L_{cod}	field in the picture header indicating the codestream size in bytes (see ISO/IEC 21122-1)

L_{h}	long header flag in the picture header
$L_{\max}$	maximum number of sampling grid points per image
$l_{\mathrm{dec}}(t)$	fill level of the decoder smoothing buffer in bits at the end of clock cycle t
$l_{\mathrm{dec,avail}}(t)$	number of bits that can be read from the decoding smoother buffer in clock cycle t
$l_{\mathrm{dec,max}}$	capacity in bits of the decoder smoothing buffer
$l_{\mathrm{enc}}(t)$	fill level of the encoder smoothing buffer in bits at the end of clock cycle t
$l_{\mathrm{enc,max}}$	capacity in bits of the encoder smoothing buffer
$l_{\mathrm{sum}}(t)$	sum of encoder and decoder smoothing buffer fill level in bits at clock cycle t
$M_{\mathrm{f}}[p,\lambda,b,g_{\mathrm{f}}]$	frame buffer bitplane count for a group g_{f} in precinct p at line λ of band b
N_{c}	number of components in an image
$N_{\mathrm{b,x}}$	size of the horizontal blanking line in sampling grid points
$N_{\mathrm{b,y}}$	size of the vertical blanking period in sampling grid lines
$N_{\mathrm{cg}}(f)$	number of coefficient groups within codestream fragment f
$N_{\mathrm{cg,hz}}$	number of coefficient groups associated to a codestream fragment representing a horizontal blanking period
$N_{\mathrm{cg,vt}}$	number of coefficient groups associated to a codestream fragment representing a vertical blanking period
N_{bpp}	nominal number of bits allocated per pixel for compression
$N_{\mathrm{bpp,max}}$	maximum number of decoded bits per pixel
$N_{\mathrm{f}}(i)$	number of fragments within a codestream $C(i)$
N_{g}	number of coefficients in a code group
$N_{\mathrm{L,x}}$	maximum number of horizontal decomposition levels of all components
$N_{\mathrm{L,y}}$	maximum number of vertical decomposition levels of all components
$N_{\mathrm{p,cg}}$	number of pixels in one coefficient group
$N_{\mathrm{p,x}}$	number of precincts per sampling grid line
$N_{\mathrm{p,y}}$	number of precincts per sampling grid column
N_{sbu}	number of decoder smoothing buffer units for a given profile
N	all integer numbers being strictly larger than zero
N_0	all integer numbers being greater than or equal to zero
N_{fg}	number of frame buffer wavelet coefficients within one frame buffer group
P_{lev}	level and sublevel indication of a codestream
$Q_{\mathrm{f}}[p]$	quantization parameter to which precinct p is quantized for storage in the frame buffer

Q_{pih}	quantization type
Q	set of rational numbers
$R_{\text{f}}[p]$	refinement parameter of the quantization to which precinct p is quantized for storage in the frame buffer
$R_{\text{s,max}}$	maximum grid point sample rate (in samples per second) at decoder output
$R_{\text{t,fb,max}}$	maximum bi-directional frame buffer bandwidth
$R_{\text{t,max}}(l_{\text{m}}, l_{\text{s}})$	maximum admissible encoded throughput in bits per second for a given level
R_{trans}	transmission channel capacity, expressed in bits per clock cycle
$r_{\text{dec}}(t)$	number of bits read and removed from the decoder smoothing buffer in clock cycle t
$S_{\text{c,max}}$	targeted maximum number of bytes of a codestream
$S_{\text{bits}}(f)$	number of bits forming the codestream fragment f
S_{d}	number of components for which the wavelet decomposition is suppressed
$S_{\text{sbo}}(p)$	smoothing buffer offset in bits for a profile p
$S_{\text{sbu}}(l_{\text{m}}, l_{\text{s}})$	size of the smoothing buffer unit in bytes for level l_{m} and sublevel l_{s}
$S_{\text{sl,max}}(l_{\text{m}}, l_{\text{s}})$	maximum size of an encoded codestream in bytes of level l_{m} and sublevel l_{s}
$s_{\text{x}}[i]$	sampling factor of component i in horizontal direction
$s_{\text{y}}[i]$	sampling factor of component i in vertical direction
T_{bmd}	buffer model type
T_{dec}	clock period defining the frequency by which code groups are processed by a decoder
T_{enc}	clock period defining the frequency by which code groups are processed by an encoder
$t_{\text{dec,read}}(f)$	timestamp in cycles at which codestream fragment f is removed from the decoder smoothing buffer
$t_{\text{dec,start}}(f)$	timestamp in cycles at which decoder starts decoding codestream fragment f
$t_{\text{enc,write}}(f)$	timestamp in cycles at which the codestream fragment f is written to the encoder smoothing buffer
$T_f[p,b]$	frame buffer truncation point for band b in precinct p
$W_{\text{c}}[i]$	width of component i in samples
$W_{\text{c,max}}$	maximum column width in sampling grid points for a given profile
W_{f}	width of the image in sampling grid points
W_{max}	maximum image width in sampling grid points
$w_{\text{dec}}(t)$	number of bits written into the decoder smoothing buffer in clock cycle t
$W_{\text{pb}}[p,b]$	width of band b of precinct p in coefficients

$W_p[p]$ width of the precinct p in sampling grid points

Z set of all integer numbers

6 Conventions

6.1 Conformance language

The keyword “reserved” indicates a provision that is not specified at this time, shall not be used, and may be specified in the future. The keyword “forbidden” indicates “reserved” and in addition indicates that the provision will never be specified in the future.

6.2 Operators

NOTE Many of the operators used in document are like those used in the C programming language.

6.2.1 Arithmetic operators

\+ addition

− subtraction (as a binary operator) or negation (as a unary prefix operator)

× multiplication

/ division without truncation or rounding

6.2.2 Logical operators

|| logical OR

&& logical AND

! logical NOT

6.2.3 Relational operators

> greater than

≥ greater than or equal to

< less than

≤ less than or equal to

== equal to

!= not equal to

6.2.4 Other operators

( ) expression

[ ] indexing of arrays

6.2.5 Precedence order of operators

Operators are listed in descending order of precedence. 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 left-associativity — thus, evaluate from left to right.

( )	expression
[ ]	indexing of arrays
−	unary negation
×, /	multiplication, division
+, −	addition, subtraction
<, >, ≤, ≥, ==, !=	relational comparison

6.2.6 Mathematical functions

$\lceil x \rceil$	ceil of x, equals the smallest integer that is greater than or equal to x
$\lfloor x \rfloor$	floor of x, equals the largest integer that is less than or equal to x
$\lvert x \rvert$	absolute value of x, $\lvert x \rvert = \begin{cases} -x & \text{if } x < 0 \\ x & \text{if } x \geq 0 \end{cases}$
$\mathrm{sign}(x)$	sign of x, $\mathrm{sign}(x) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases}$
$\xi(t)$	step function, $\xi(t) = \begin{cases} 0 & \text{if } t < 0 \\ 1 & \text{if } t \geq 0 \end{cases}$
$\max_i(x_i)$	maximum of a sequence of numbers $[x_i]$ enumerated by the index i
∃	the mathematical symbol to represent *there exists*
∀	The mathematical symbol to represent *for all*

7 Buffer model

7.1 General system block diagram

The JPEG XS coding system addresses applications where coded images are transferred from a source to a target, as shown in Figure 1. To this end, the encoder is compressing a continuous stream of input pixels into a sequence of bits. These bits are forwarded by means of a transmission channel to the decoder that decompresses the bits to produce a continuous stream of output pixels.