INTERNATIONAL STANDARD

# ISO/IEC 21122-1

Second edition
2022-03

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

## Part 1:
## Core coding system

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

*Partie 1: Système de codage de noyau*



Reference number
ISO/IEC 21122-1:2022(E)

# 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).

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 patents.iec.ch).

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 second edition cancels and replaces the first edition (ISO/IEC 21122-1:2019), which has been technically revised.

The main changes are as follows:

— coding tools for compressing colour filter array images (CFA images) have been added;

— coding tools that enable lossless coding of images with up to 12 bits per sample have been added;

— support for 4:2:0 sampled images has been added.

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 www.iso.org/members.html and www.iec.ch/national-committees.

# 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 patents.

ISO and IEC take no position concerning the evidence, validity and scope of these patent rights.

The holders of these patent rights have assured ISO and IEC that they are willing to negotiate licences under reasonable and non-discriminatory terms and conditions with applicants throughout the world. In this respect, the statements of the holders of these patent rights are 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.

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

## Part 1:
## Core coding system

## 1 Scope

This document defines the syntax and an accompanying decompression process that is capable to represent continuous-tone grey-scale, or continuous-tone colour digital images without visual loss at moderate compression rates. Typical compression rates are between 2:1 and 6:1 but can also be higher depending on the nature of the image. In particular, the syntax and the decoding process specified in this document allow lightweight encoder and decoder implementations that limit the end-to-end latency to a fraction of the frame size. However, the definition of transmission channel buffer models necessary to ensure such latency is beyond the scope of this document.

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

There are no normative references in this document.

## 3 Terms and definitions, abbreviated terms and symbols

### 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**
**band**
input data to a specific *wavelet filter type* (3.1.58) that contributes to the generation of one of the *components* (3.1.14) of the image

**3.1.2**
**band type**
single number collapsing the information on the component, and horizontal and vertical wavelet filter types that are applied in the filter cascade reconstructing spatial image samples from inversely quantized wavelet coefficients

**3.1.3**
**bit**
binary choice encoded as either 0 or 1

**3.1.4**
**bitplane**
array of bits having all the same significance

**3.1.5**
**bitplane count**
number of significant bitplanes of a code group, counting from the LSB up to the most significant, non-empty bitplane

**3.1.6**
**bitplane count subpacket**
subset of a packet which decodes to the bitplane counts of all code groups within a packet, followed by padding and optional filler bytes

Note 1 to entry: See subclause C.5.3.

**3.1.7**
**byte**
group of 8 bits

**3.1.8**
**colour filter array**
**CFA**
rectangular array of sensor elements yielding a 1-component picture where the colour to which a sensor element is sensitive to depends on the position of the sensor element



**3.1.9**
**codestream**
compressed image data representation that includes all necessary data to allow a (full or approximate) reconstruction of the sample values of a digital image

**3.1.10**
**code group**
group of quantization indices in sign-magnitude representation before inverse quantization

**3.1.11**
**coefficient**
input value to the inverse wavelet transformation resulting from inverse quantization

**3.1.12**
**coefficient group**
number of horizontally adjacent wavelet coefficients from the same band

**3.1.13**
**column**
set of vertically aligned precincts

**3.1.14**
**component**
two-dimensional array of samples having the same designation such as red, green or blue in the output or display device

**3.1.15**
**compression**
process of reducing the number of bits used to represent source image data

**3.1.16**
**continuous-tone image**
image whose components have more than one bit per sample

**3.1.17**
**data subpacket**
subset of a packet which consists of the quantization index magnitudes, followed by padding and optional filler bytes

Note 1 to entry: See subclause C.5.4.

**3.1.18**
**deadzone quantizer**
quantizer whose zero bucket has a size different from all other buckets

**3.1.19**
**decoder**
embodiment of a decoding process

**3.1.20**
**decoding process**
process which takes as its input a codestream and outputs a continuous-tone image

**3.1.21**
**decomposition level**
set of wavelet coefficients resulting from a particular level of recursive application of a wavelet transform



**3.1.22**
**downsampling**
procedure by which the spatial resolution of a component is reduced



**3.1.23**
**encoder**
embodiment of an encoding process.

**3.1.24**
**encoding process**
process which outputs compressed image data in the form of a codestream

**3.1.25**
**entropy decoder**
embodiment of an entropy decoding procedure

**3.1.26**
**entropy decoding**
lossless procedure which recovers the sequence of symbols from the sequence of bits produced by the entropy encoder

**3.1.27**
**entropy encoder**
embodiment of an entropy encoding procedure

**3.1.28**
**entropy encoding**
lossless procedure which converts 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

**3.1.29**
**filler bytes**
integer number of bytes a decoder will skip over on decoding without interpreting the values of the bytes itself

**3.1.30**
**grayscale image**
continuous-tone image that has only one component

**3.1.31**
**inverse quantization**
inverse procedure to quantization by which the decoder recovers a representation of the coefficients

**3.1.32**
**inverse reversible multi component transformation**
**inverse RCT**
inverse transformation across multiple component sample values located at the same sample grid point that is invertible without loss

Note 1 to entry: See subclauses F.3 and F.4.

**3.1.33**
**LL band**
input to a series of wavelet filters where only inverse low-pass filters are applied in horizontal and vertical direction



**3.1.34**
**lossless**
descriptive term for encoding and decoding processes and procedures in which the output of the decoding procedure(s) is identical to the input to the encoding procedure(s)



**3.1.35**
**lossless coding**
mode of operation which refers to any one of the coding processes defined in this document in which all of the procedures are lossless

**3.1.36**
**lossy**
descriptive term for encoding and decoding processes which are not lossless

**3.1.37**
**packet**
segment of the codestream containing entropy coded information on a single precinct, line and a subset of the bands within this precinct and line

**3.1.38**
**padding**
bits within the codestream whose only purpose is to align syntax elements to byte boundaries and that carry no information

**3.1.39**
**precinct**
collection of quantization indices of all bands contributing to a given spatial region of the image

**3.1.40**
**precision**
number of bits allocated to a particular sample, coefficient, or other binary numerical representation

**3.1.41**
**procedure**
set of steps which accomplishes one of the tasks which comprise an encoding or decoding process

**3.1.42**
**quantization**
method of reducing the precision of the individual coefficients

**3.1.43**
**quantization index**
input to the inverse quantization process which reconstructs the quantization index to a wavelet coefficient

**3.1.44**
**quantization index magnitude**
absolute value of a quantization index

**3.1.45**
**sample**
one element in the two-dimensional image array which comprises a component

**3.1.46**
**sample grid**
common coordinate system for all samples of an image, the samples at the top left edge of the image have the coordinates (0,0), the first coordinate increases towards the right, the second towards the bottom

**3.1.47**
**sign subpacket**
subset of a packet that consists of the sign information of all non-zero quantization indices within a packet, followed by padding and optional filler bytes

Note 1 to entry: See subclause C.5.5.



**3.1.48**
**significance**
attribute of code groups that applies if, depending on the Run Mode flag in the picture header, either at least one of coefficients in the code group is non-zero, or the bitplane count prediction residual of the code group is non-zero

**3.1.49**
**significance group**
group of a horizontally adjacent code groups sharing the same significance information in the significance subpacket

**3.1.50**
**significance subpacket**
subset of a packet that identifies which significance groups within a packet are insignificant, followed by padding and optional filler bytes

Note 1 to entry: see subclause C.5.2

**3.1.51**
**slice**
integral number of precincts whose wavelet coefficients can be entropy-decoded independently

**3.1.52**
**star-tetrix**
decorrelation transformation that combines a spatial with an inter-component decorrelation transformation particularly tuned for CFA pattern compression

Note 1 to entry: see subclause F.5

**3.1.53**
**subpacket**
substructure of a packet containing information of one or multiple bands of one line of a single precinct

**3.1.54**
**super pixel**
2×2 arrangement of sensor elements in a CFA pattern array containing at least one sensor element for each colour filter type

**3.1.55**
**truncation position**
number of least significant bitplanes not included in the quantization index of a wavelet coefficient

**3.1.56**
**uniform quantizer**
quantizer whose buckets are all of equal size

**3.1.57**
**upsampling**
procedure by which the spatial resolution of a component is increased

**3.1.58**
**wavelet filter type**
single number that uniquely identifies each element of the wavelet filter with regard to the number and type of horizontal and vertical decompositions

Note 1 to entry: Unlike the band type, the wavelet filter type does not include component information.



## 3.2 Abbreviated terms

| | |
|---|---|
| JPEG XS | informal name of this standard where XS stands for "extra speed" |
| LSB | least significant bit |
| MSB | most significant bit |



## 3.3 Symbols

| | |
|---|---|
| $B[c]$ | bit precision of component $c$ |
| $\beta$ | wavelet filter type |
| $b$ | band type |
| $b_x[\beta,i]$ | band existence flag for filter type $\beta$ in component $i$. 1 if the filter exists, 0 otherwise. |
| $b'_x[b]$ | band existence flag for band type $b$. 1 if the filter exists, 0 otherwise. |
| $Bw$ | nominal overall bit precision of the wavelet data |
| $B_r$ | number of bits required to encode a bitplane count in raw |
| $Cpih$ | colour transformation type |
| $c[p,\lambda,b,x]$ | wavelet coefficient in precinct $p$, line $\lambda$, band $b$ and position $x$ |
| $C_s$ | width of precincts other than the rightmost precinct in sample grid positions |
| $C_t$ | colour transformation CFA pattern type derived from the component registration |

| | |
|---|---|
| $C_f$ | colour transformation reflection and extension flags |
| $Cw$ | width of precincts in multiples of 8 LL subsampled band sample grid positions |
| $D[p,b]$ | bitplane count coding mode of band $b$ in precinct $p$ |
| $D_r[p,s]$ | raw coding mode override flag for packet $s$ in precinct $p$ |
| $DCO$ | DC offset |
| $d_x[\beta,i]$ | horizontal decomposition level of wavelet filter type $\beta$ of component $i$ |
| $d_y[\beta,i]$ | vertical decomposition level of wavelet filter type $\beta$ of component $i$ |
| $\delta_x[c]$ | horizontal position of component $c$ in a CFA super pixel |
| $\delta_y[c]$ | vertical position of component $c$ in a CFA super pixel |
| $E$ | exponent of the slope of the linear region of the extended non-linearity |
| $e_1$ | colour transformation exponent of first chroma component |
| $e_2$ | colour transformation exponent of second chroma component |
| $Fs$ | sign packing flag |
| $Fslc$ | slice coding mode |
| $Fq$ | number of fractional bits in the representation of wavelet coefficients |
| $G[b]$ | gain of subband $b$ |
| $H_b[\beta,k]$ | height of filter type $\beta$ of component $k$ in wavelet coefficients |
| $H_c[i]$ | height of the component $i$ in sample points |
| $H_f$ | height of the image in sampling grid points |
| $H_p$ | height of a precinct in lines |
| $H_{sl}$ | height of a slice in precincts |
| $I[p,b,\lambda,s]$ | line inclusion flag, set if line $\lambda$ of band $b$ and precinct $p$ is included in packet $s$, reset otherwise |
| $k[\delta_x, \delta_y]$ | Component within CFA super pixel at position $\delta_x$, $\delta_y$ |
| $L_0[p,b]$ | first line of band $b$ in precinct $p$ |
| $L_1[p,b]$ | last line + 1 of band $b$ in precinct $p$ |
| $Lcod$ | codestream length in bytes |
| $L_{cnt}[p,s]$ | size of the bitplane count subpacket of precinct $p$ and packet $s$ in bytes |
| $L_{dat}[p,s]$ | size of the data subpacket of precinct $p$ and packet $s$ in bytes |
| $Lh$ | long header flag in in the picture header, set if long headers are enforced, reset otherwise |
| $L_{prc}[p]$ | length of the entropy coded data in precinct $p$ |
| $L_{sgn}[p,s]$ | size of the sign subpacket of precinct $p$ and packet $s$ in bytes |

$L_{sig}[p,s]$ size of the significance subpacket of precinct *p* and packet *s* in bytes

$M[p,\lambda,b,g]$ bitplane count of precinct *p*, line *λ*, band *b* and code group *g*

$M_{top}[p,\lambda,b,g]$ vertical predictor of the bitplane count of precinct *p*, line *λ*, band *b* and code group *g*

$N_c$ number of components in an image

$N_{cg}[p,b]$ number of code groups in precinct *p* and band *b*

$N_\beta$ number of bands per component

$N_g$ number of coefficients in a code group

$N_s[p,b]$ number of significance groups per line band *b* of precinct *p*

$N_p[t]$ number of precincts in slice *t*

$N_L$ number of bands in the wavelet decomposition of the image (wavelet filter types times components)

$N_{L,x}$ maximal number of horizontal decomposition levels

$N'_{L,x}[i]$ number of horizontal decomposition levels of component *i*

$N_{L,y}$ maximal number of vertical decomposition levels over all components



$N'_{L,y}[i]$ number of vertical decomposition levels of component *i*

$N_{p,x}$ number of precincts per sampling grid line

$N_{p,y}$ number of precincts per sampling grid column



$N_{pc}[p]$ number of packets in precinct *p*

$O[c,x,y]$ unscaled output of the inverse wavelet transformation at coordinates *x* and *y* of the component *c*

$\Omega[c,x,y]$ output of the inverse multiple component transformation at position *x*,*y* for component *c*

$P[b]$ priority of band *b*

*Plev* level a particular codestream complies to

*Ppih* profile a particular codestream complies to

*Ppoc* progression order in which bands are transmitted in the codestream

$Q[p]$ quantization parameter of precinct *p*

*Qpih* quantization type of the picture

*Rl* raw-mode selection per packet flag

*Rm* run mode used for significance coding

$R[p]$ refinement of precinct *p*

$R[c,x,y]$ reconstructed sample value at position *x*,*y* for component *c*

*Sd* number of components for which wavelet decomposition is suppressed

| | |
|---|---|
| $S_s$ | size of a significance group in code groups |
| $s_x[i]$ | sampling factor of component *i* in horizontal direction |
| $s_y[i]$ | sampling factor of component *i* in vertical direction |
| $s[p,\lambda,b,x]$ | sign of the wavelet coefficient in precinct *p*, line *λ*, band *b* and position *x*. |
| $T1$ | first threshold of the extended non-linearity |
| $T2$ | second threshold of the extended non-linearity |
| $T[p,b]$ | truncation position of precinct *p* and band *b* |
| $T_{top}[p,b]$ | vertical Truncation position predictor of precinct *p* and band *b* |
| $T[\beta,x,y]$ | temporary wavelet coefficient of filter type *β* at location *x*,*y*. |
| $v[x,y]$ | sample value at the sample grid position *x*,*y* |
| $v[p,\lambda,b,x]$ | quantization index magnitude of the wavelet coefficient in precinct *p*, line *λ*, band *b* and position *x* |
| $W_b[\beta,k]$ | width of filter type *β* of component *k* in wavelet coefficients |
| $W_c[i]$ | width of component *i* in samples |
| $W_f$ | width of the image in sampling grid points |
| $W_p[p]$ | width of the precinct *p* in sampling grid points |
| $W_{pb}[p,b]$ | width of subband b of precinct p in coefficients |
| $Wt_x$ | wavelet filter type for horizontal filtering |
| $Wt_y$ | wavelet filter type for vertical filtering |
| $X[y]$ | one-dimensional temporal array of wavelet coefficients |
| $Xcrg[c]$ | horizontal component registration of component *c* relative to the sample grid |
| $Ycrg[c]$ | vertical component registration of component *c* relative to the sample grid |
| $Yslh$ | vertical slice order within the picture |
| $Z[p,\lambda,b,j]$ | significance flag of precinct *p*, line *λ*, band *b* and significance group *j* |



# 4 Conventions

## 4.1 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.

## 4.2 Operators

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

### 4.2.1 Arithmetic operators

 