FINAL DRAFT

INTERNATIONAL STANDARD

ISO/IEC FDIS 18477-8

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

Secretariat: JISC

Voting begins on:
2020-03-13

Voting terminates on:
2020-05-08

Information technology — Scalable compression and coding of continuous-tone still images —

Part 8: Lossless and near-lossless coding

Reference number
ISO/IEC FDIS 18477-8:2020(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).

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 http://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.

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 18477-8:2016), which has been technically revised.

The main changes compared to the previous edition are as follows:

— Annex F.2 has been revised to adopt centred upsampling by default;

— minor editorial changes throughout.

A list of all parts in the ISO/IEC 18477 series can be found on the ISO website.

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.

Introduction

This document specifies a coded codestream format for storage of continuous-tone high and low dynamic range photographic content. This is a scalable lossy to lossless image coding system supporting multiple component images consisting of integer samples between 8- and 16-bit resolution, or floating point samples of 16-bit resolution. It is by itself an extension of ISO/IEC 18477-6 and ISO/IEC 18477-7, which specify intermediate range and high-dynamic range image decoding algorithms. Both of these are based on the box-based file format specified in ISO/IEC 18477-3, which is again an extension of ISO/IEC 18477-1; the codestream is composed in such a way that legacy applications conforming to Rec. ITU-T T.81 | ISO/IEC 10918-1 are able to reconstruct a lossy, low dynamic range, 8 bits per sample version of the image.

Today, the most widely used digital photography format, a minimal implementation of JPEG (specified in Rec. ITU-T T.81 | ISO/IEC 10918-1), uses a bit depth of 8; each of the three channels that together compose an image pixel is represented by 8 bits, providing 256 representable values per channel. For more demanding applications, it is not uncommon to use a bit depth of 16, providing 65 536 representable values to describe each channel within a pixel, resulting in over 2.8×10^{14} representable colour values. In some less common scenarios, even greater bit depths are used, requiring a floating-point sample representation.

Most common photo and image formats use an 8-bit or 16-bit unsigned integer value to represent some function of the intensity of each colour channel. While it might be theoretically possible to agree on one method for assigning specific numerical values to real world colours, doing so is not practical. Since any specific device has its own limited range for colour reproduction, the device's range may be a small portion of the agreed-upon universal colour range. As a result, such an approach is an extremely inefficient use of the available numerical values, especially when using only 8 bits (or 256 unique values) per channel. To represent pixel values as efficiently as possible, devices use a numeric encoding optimized for their own range of possible colours or gamut.

This document is primarily designed to encode intermediate or high dynamic image sample values **without loss**, or with a precisely controllable bounded loss using the tools defined in ISO/IEC 18477-1 and some minimal extensions of those tools. The goal is to provide a backwards-compatible coding specification that allows legacy applications and existing toolchains to continue to operate on codestreams conforming to this document.

JPEG XT has been designed to be backwards compatible to legacy applications while at the same time having a small coding complexity; JPEG XT uses, whenever possible, functional blocks of Rec. ITU-T T.81 | ISO/IEC 10918-1 to extend the functionality of the legacy JPEG coding system. It is optimized for storage and transmission of intermediate and high dynamic range and wide colour gamut 8- to 16-bit integer or 16-bit floating point images while also enabling low-complexity encoder and decoder implementations.

This document is an extension of ISO/IEC 18477-1, a compression system for continuous tone digital still images which is backwards compatible with Rec. ITU-T T.81 | ISO/IEC 10918-1. That is, legacy applications conforming to Rec. ITU-T T.81 | ISO/IEC 10918-1 will be able to reconstruct streams generated by an encoder conforming to this document, though will possibly not be able to reconstruct such streams in full dynamic range, full quality or without loss.

This document is itself based on ISO/IEC 18477-3 that defines a box-based file format similar to other JPEG standards. It also contains elements of ISO/IEC 18477-6 and ISO/IEC 18477-7. The aim of this document is to provide a migration path for legacy applications to support lossless coding of intermediate and high dynamic range images, that is images that are either represented by sample values requiring 8- to 16-bit precision, or even using 16-bit floating point sample resolution. While Rec. ITU-T T.81 | ISO/IEC 10918-1 already defines a lossless mode for integer samples, images encoded in this mode cannot be decoded by applications only supporting the lossy 8-bit-mode; the coding engine for lossless coding in Rec. ITU-T T.81 | ISO/IEC 10918-1 is completely different from the lossy coding mode. Unlike the legacy standard, this document defines a lossless scalable coding engine supporting all bit depths between 8 and 16 bits per sample, including 16-bit floating point samples, while also staying compatible with legacy applications. Such applications will continue to work, but will only able to reconstruct a lossy 8-bit standard low dynamic range (LDR) version of the full image contained in the codestream. The ISO/IEC 18477 series specifies a coded file format, referred to as JPEG XT, which is designed primarily for storage and interchange of continuous-tone photographic content.

Information technology — Scalable compression and coding of continuous-tone still images —

Part 8:
Lossless and near-lossless coding

1 Scope

This document specifies a coding format, referred to as JPEG XT, which is designed primarily for continuous-tone photographic content. This document defines extensions that allow lossless coding of such content while staying compatible with the core coding system specified in ISO/IEC 18477-1.

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 18477-1:2015, *Information technology — Scalable compression and coding of continuous-tone still images — Part 1: Scalable compression and coding of continuous-tone still images*

ISO/IEC 18477-3:2015, *Information technology — Scalable compression and coding of continuous-tone still images — Part 3: Box file format*

ISO/IEC 18477-6:2016, *Information technology — Scalable compression and coding of continuous-tone still images — Part 6: IDR Integer Coding*

ISO/IEC 18477-7:2017, *Information technology — Scalable compression and coding of continuous-tone still images — Part 7: HDR Floating-Point Coding*

ITU-T T.81 | ISO/IEC 10918-1, *Information technology — Digital compression and coding of continuous tone still images — Requirements and guidelines*

ITU-T BT.601, *Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios*

3 Terms, definitions, symbols and abbreviated terms

3.1 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:

— IEC Electropedia: available at http://www.electropedia.org/

— ISO Online browsing platform: available at http://www.iso.org/obp

3.1.1
AC coefficient
any DCT coefficient for which the frequency is not zero in at least one dimension

3.1.2
ASCII encoding
encoding of text characters and text strings

Note 1 to entry: In accordance with ISO/IEC 10646.

3.1.3
base decoding path
process of decoding legacy codestream and refinement data to the base image, jointly with all further steps until residual data is added to the values obtained from the residual codestream

3.1.4
base image
collection of sample values obtained by entropy decoding the DCT coefficients of the legacy codestream and the refinement codestream, and inversely DCT transforming them jointly

3.1.5
box
structured collection of data describing the image or the image decoding process embedded into one or multiple APP_{11} marker segments

Note 1 to entry: See ISO/IEC 18477-3:2015, Annex B for the definition of boxes.

3.1.6
coding process
encoding process, a decoding process, or both

3.1.7
DC coefficient
DCT coefficient for which the frequency is zero in both dimensions

3.1.8
DCT coefficient
amplitude of a specific cosine basis function

Note 1 to entry: May refer to an original DCT coefficient, to a quantized DCT coefficient, or to a dequantized DCT coefficient.

3.1.9
decoder
embodiment of a decoding process

3.1.10
decoding process
process which takes as its input compressed image data and outputs a continuous-tone image

3.1.11
dequantization
inverse procedure to quantization by which the decoder recovers a representation of the DCT coefficients

3.1.12
encoder
embodiment of an encoding process

3.1.13
encoding process
process which takes as its input a continuous-tone image and outputs compressed image data

3.1.14
extension image
residual image
sample values as reconstructed by inverse quantization and inverse DCT transformation applied to the entropy-decoded coefficients described by the residual scan and residual refinement scans

[SOURCE: ISO/IEC 18477-6:2016, 3.1.54]

3.1.15
fixed point discrete cosine transformation
implementation of the discrete cosine transformation based on fixed point arithmetic

Note 1 to entry: As specified in Annex E.

3.1.16
forward DCT bypass
transformation that takes an 8×8 sample block and prepares it for entropy coding without applying a discrete cosine transformation

3.1.17
forward fixed point DCT
transformation of an 8×8 sample block from the spatial domain to the frequency domain using the fixed point arithmetic

Note 1 to entry: As specified in Annex E.

3.1.18
forward integer DCT
transformation of an 8×8 sample block from the spatial domain to the frequency domain using the integer approximation of the discrete cosine transformation

Note 1 to entry: As specified in Annex E.

3.1.19
inverse DCT bypass
transformation that takes an 8×8 sample block as generated by entropy decoding and level-shifts it without applying a discrete cosine transformation

3.1.20
inverse fixed point DCT
transformation of an 8×8 sample block from the frequency domain to the spatial domain using the fixed point arithmetic

Note 1 to entry: As specified in Annex E.

3.1.21
inverse integer DCT
the transformation of an 8×8 sample block from the frequency domain to the spatial domain using the integer approximation of the discrete cosine transformation

Note 1 to entry: As specified in Annex E.

3.1.22
frequency
two-dimensional index into the two-dimensional array of DCT coefficients

[SOURCE: ISO/IEC 10918-1:1994, 3.1.61]

3.1.23
high dynamic range
HDR
image or image data comprised of more than eight bits per sample

3.1.24
Huffman encoding
entropy encoding procedure which assigns a variable length code to each input symbol

3.1.25
intermediate dynamic range
image or image data comprised of more than eight bits per sample

3.1.26
legacy codestream
collection of markers and syntax elements

Note 1 to entry: The legacy codestream, as defined by Rec. ITU-T T.81 | ISO/IEC 10918-1 and any syntax elements defined by the ISO/IEC 18477 series, consists of the collection of all markers except those APP_{11} markers that describe JPEG XT boxes by the syntax defined in ISO/IEC 18477-3:2015, Annex A.

3.1.27
lossless
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)

3.1.28
lossless coding
mode of operation which refers to any one of the coding processes in which all of the procedures are lossless

Note 1 to entry: Coding processes defined in ISO/IEC 18477-8.

3.1.29
lossy
encoding and decoding processes which are not lossless

3.1.30
low-dynamic range
LDR
image or image data comprised of data with no more than eight bits per sample

3.1.31
noise shaping
signal processing technique that removes quantization noise from the low frequency components and injects it into the high frequency domain where it can be removed by filtering

3.1.32
point transformation
application of a location independent global function to reconstructed sample values in the spatial domain

3.1.33
residual decoding path
collection of operations applied to the entropy coded data contained in the residual data box and residual refinement scan boxes up to the point where this data is merged with the legacy data to form the final output image

3.1.34
residual image
sample values as reconstructed by inverse quantization and inverse DCT transformation applied to the entropy-decoded coefficients described by the residual scan and residual refinement scans

3.1.35
residual scan
additional pass over the image data invisible to legacy decoders which provides additive and/or multiplicative correction data of the legacy scans to allow reproduction of high-dynamic range or wide colour gamut data

3.1.36
refinement scan
additional pass over the image data invisible to legacy decoders which provides additional least significant bits to extend the precision of the DCT transformed coefficients

Note 1 to entry: Refinement scans can be either applied in the legacy or residual decoding path.

3.1.37
superbox
box that carries other boxes as payload data

3.1.38
sub box
box that is contained as payload data within a superbox

3.1.39
uniform quantization
procedure by which DCT coefficients are linearly scaled in order to achieve compression

3.1.40
upsampling
procedure by which the spatial resolution of a component is increased

3.2 Symbols

X width of the sample grid in positions

Y height of the sample grid in positions

Nf number of components in an image

$s_{i,x}$ subsampling factor of component i in horizontal direction

$s_{i,y}$ subsampling factor of component i in vertical direction

H_i subsampling indicator of component i in the frame header

V_i subsampling indicator of component i in the frame header

$v_{x,y}$ sample value at the sample grid position x,y

R_h additional number of DCT coefficients bits represented by refinement scans in the base image, $8+R_h$ is the number of non-fractional bits (i.e. bits in front of the "binary dot") of the output of the inverse DCT process in the base image

R_r additional number of DCT coefficients bits represented by refinement scans in the residual, $P+R_h$ is the number of non-fractional bits (i.e. bits in front of the "binary dot") of the output of the inverse DCT process in the residual image where P is the bit depth indicated in the frame header of the residual codestream

R_b additional bits in the HDR image. 8+Rb is the sample precision of the reconstructed HDR image

3.3 Abbreviated terms

ASCII	American Standard Code for Information Interchange
LSB	least significant bit
MSB	most significant bit
TMO	tone mapping operator
DCT	discrete cosine transformation
FCT	fixed point multi-component transformation
ICT	irreversible multi-component transformation
RCT	reversible multi-component transformation
JPEG	joint photographic experts group

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 this document are similar to those used in the C programming language.

4.2.1 Arithmetic operators

+	addition
−	subtraction (as a binary operator) or negation (as a unary prefix operator)
×	multiplication
/	division without truncation or rounding
smod	x smod a is the unique value y between $-\lceil (a-1)/2 \rceil$ and $\lfloor (a-1)/2 \rfloor$ for which y+N×a = x with a suitable integer N.
umod	x umod a is the unique value y between 0 and a−1 for which y+N×a = x with a suitable integer N

4.2.2 Logical operators

|| logical OR

&& logical AND

! logical NOT

∈ x ∈ {A, B} is defined as (x == A || x == B)

∉ x ∉ {A, B} is defined as (x != A && x != B)

4.2.3 Relational operators

\> greater than

\>= greater than or equal to

< less than

<= less than or equal to

== equal to

!= not equal to

4.2.4 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 the associativity of the operator either from right to left or from left to right.

Operators	Type of operation	Associativity
(), [ ], .	expression	left to right
–	unary negation	
×, /	multiplication	left to right
umod, smod	modulo (remainder)	left to right
+, –	addition and subtraction	left to right
< , >, <=, >=	relational	left to right

4.2.5 Mathematical functions

$\lceil x \rceil$ ceil of x: returns the smallest integer that is greater than or equal to x

$\lfloor x \rfloor$ floor of x: returns the largest integer that is lesser than or equal to x

|x| absolute value, is –x for x < 0, otherwise x

sign(x) sign of x, 0 if x is zero, +1 if x is positive, −1 if x is negative

clamp(x,min,max) clamps x to the range [min,max]: returns min if x < min, max if x > max or otherwise x

x^a raises the value of x to the power of a: x is a non-negative real number, a is a real number; x^a is equal to exp(a×log(x)) where exp is the exponential function and log() the natural logarithm; if x is 0 and a is positive, x^a is defined to be 0