INTERNATIONAL STANDARD

ISO/IEC 18477-6

First edition
2016-02-01

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

Part 6: IDR Integer Coding

Technologies de l'information — Compression échelonnable et codage d'images plates en ton continu —

Partie 6: Codage de nombre entier par IDR

Reference number
ISO/IEC 18477-6:2016(E)

© ISO/IEC 2016



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2016, Published in Switzerland

All rights reserved. Unless otherwise specified, 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
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

Contents

Page

Foreword ... **iv**

Introduction ... **v**

1 Scope ... **1**

2 Normative references ... **1**

3 Terms and definitions, abbreviated terms, and symbols ... **1**
3.1 Terms and definitions ... 1
3.2 Symbols ... 6
3.3 Abbreviated terms ... 7

4 Conventions ... **7**
4.1 Conformance language ... 7
4.2 Operators ... 8
4.2.1 Arithmetic operators ... 8
4.2.2 Logical operators ... 8
4.2.3 Relational operators ... 8
4.2.4 Precedence order of operators ... 8
4.2.5 Mathematical functions ... 9

5 General ... **9**
5.1 High level overview on JPEG XT ISO/IEC 18477-6 ... 9
5.2 Profiles ... 11
5.3 Encoder requirements ... 11
5.4 Decoder requirements ... 11

Annex A (normative) **Encoding and decoding process** ... **13**

Annex B (normative) **Boxes** ... **17**

Annex C (normative) **Multi-component decorrelation** ... **24**

Annex D (normative) **Entropy coding of refinement data** ... **28**

Annex E (normative) **Profiles** ... **33**

Bibliography ... **34**

© ISO/IEC 2016 – All rights reserved

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. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

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).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the WTO principles in the Technical Barriers to Trade (TBT) see the following URL: Foreword - Supplementary information

The committee responsible for this document is ISO/IEC JTC 1, *Information technology*, SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

ISO/IEC 18477 contains the following parts under the general title *Information technology — Scalable compression and coding of continuous-tone still images*:

— *Part 1: Scalable compression and coding of continuous-tone still images*

— *Part 2: Extensions for high dynamic range images*

— *Part 3: Box file format*

— *Part 6: IDR Integer Coding*

— *Part 7: HDR Floating-Point Coding*

— *Part 8: Lossless and Near-lossless Coding*

— *Part 9: Alpha Channel Coding*

The following parts are under preparation:

— *Part 4: Conformance testing*

— *Part 5: Reference software*

 © ISO/IEC 2016 – All rights reserved

Introduction

This part of ISO/IEC 18477 specifies a coded codestream format for storage of continuous-tone high and low dynamic range photographic content. JPEG XT part 6 is a scalable image coding system supporting multiple component images consisting of integer samples of a bit precision between 9 and 16 bits. The format itself is based on the Box Based format specified in ISO/IEC 18477-3, which ensures that legacy applications conforming to Rec. ITU-T T.81 | ISO/IEC 10918-1 are able to reconstruct a lower quality, low dynamic range, eight bits per sample version of the image.

Today, the most widely used digital photography format, a minimal implementation of JPEG (specified in ITU Recommendation 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 on over $2{,}8 \times 10^{14}$ representable colour values.

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.

JPEG XT is primarily designed to provide coded data containing intermediate dynamic range and wide colour gamut content while simultaneously providing 8 bits per pixel low dynamic range images using tools defined in ISO/IEC 18477-1, which is itself a subset of Rec. ITU-T T.81 | ISO/IEC 10918-1. 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 part of ISO/IEC 18477.

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 dynamic range and wide colour gamut images while also enabling low-complexity encoder and decoder implementations.

This part of ISO/IEC 18477 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 part of ISO/IEC 18477, though will possibly not be able to reconstruct such streams in full dynamic range, full quality or other features defined in this Recommendation| International Standard.

This part of ISO/IEC 18477 is itself based on ISO/IEC 18477-3 which defines a box-based file format similar to other JPEG standards. The aim of this part of ISO/IEC 18477 is to provide a migration path for legacy applications to support, potentially in a limited way, coding of intermediate dynamic range images, that is images represented by sample values requiring 9 to 16 bits precision. While the legacy Rec. ITU-T T.81 | ISO/IEC 10918-1 already defines a coding mode for 12 bit sample precision, images encoded in this mode cannot be decoded by applications implementing only the 8 bit mode. Unlike the legacy standard, this part of ISO/IEC 18477 defines a scalable coding engine supporting all bit depths between 9 and 16 bits per sample while also staying compatible with legacy applications. Such applications will continue to work, but will only able to reconstruct an 8 bit standard low dynamic range (LDR) version of the full image contained in the codestream. This part of ISO/IEC 18477 specifies a coded file format, referred to as JPEG XT, which is designed primarily for storage and interchange of continuous-tone photographic content.

© ISO/IEC 2016 – All rights reserved

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

Part 6:
IDR Integer Coding

1 Scope

This part of ISO/IEC 18477 specifies a coding format, referred to as JPEG XT, which is designed primarily for continuous-tone photographic content.

2 Normative references

The following documents, in whole or in part, are normatively referenced in this document and are indispensable for its application. 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-based file format*

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

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

3 Terms and definitions, abbreviated terms, and symbols

3.1 Terms and definitions

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

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 according to ISO/IEC 10646-1

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

© ISO/IEC 2016 – All rights reserved

3.1.5
binary decision
choice between two alternatives

3.1.6
bitstream
partially encoded or decoded sequence of bits comprising an entropy-coded segment

3.1.7
block
8 × 8 array of samples or an 8 × 8 array of DCT coefficient values of one component

3.1.8
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.

3.1.9
byte
group of 8 bits

3.1.10
coder
embodiment of a coding process

3.1.11
coding
encoding or decoding

3.1.12
coding model
procedure used to convert input data into symbols to be coded

3.1.13
(coding) process
general term for referring to an encoding process, a decoding process, or both

3.1.14
compression
reduction in the number of bits used to represent source image data

3.1.15
component
two-dimensional array of samples having the same designation in the output or display device

Note 1 to entry: An image typically consists of several components, e.g. red, green, and blue.

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

3.1.17
DC coefficient
DCT coefficient for which the frequency is zero in both dimensions

3.1.18
decoder
embodiment of a decoding process

© ISO/IEC 2016 – All rights reserved

3.1.19
decoding process
process which takes as its input compressed image data and outputs a continuous-tone image

3.1.20
dequantization
inverse procedure to quantization by which the decoder recovers a representation of the DCT coefficients

3.1.21
discrete cosine transform
DCT
either the forward discrete cosine transform or the inverse discrete cosine 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 takes as its input a continuous-tone image and outputs compressed image data

3.1.25
entropy-coded (data) segment
independently decodable sequence of entropy encoded bytes of compressed image data

3.1.26
entropy decoder
embodiment of an entropy decoding procedure

3.1.27
entropy decoding
lossless procedure which recovers the sequence of symbols from the sequence of bits produced by the entropy encoder

3.1.28
entropy encoder
embodiment of an entropy encoding procedure

3.1.29
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

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

3.1.31
high dynamic range
image or image data comprised of more than eight bits per sample

3.1.32
Huffman decoder
embodiment of a Huffman decoding procedure

© ISO/IEC 2016 – All rights reserved

3.1.33
Huffman decoding
entropy decoding procedure which recovers the symbol from each variable length code produced by the Huffman encoder

3.1.34
Huffman encoder
embodiment of a Huffman encoding procedure

3.1.35
Huffman encoding
entropy encoding procedure which assigns a variable length code to each input symbol

3.1.36
intermediate dynamic range
image or image data comprised of more than eight bits per sample

3.1.37
joint photographic experts group
JPEG
informal name of the committee which created this part of ISO/IEC 18477

Note 1 to entry: The "joint" comes from the ITU-T and ISO/IEC collaboration.

3.1.38
legacy codestream
collection of markers and syntax elements defined by Rec. ITU-T T.81 | ISO/IEC 10918-1 bare any additional syntax elements defined by the family ISO/IEC 18477 standards, i.e. the legacy codestream 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

3.1.39
legacy decoding path
collection of operations to be performed on the entropy coded data as described by Rec. ITU-T T.81 | ISO/IEC 10918-1 jointly with the Legacy Refinement scans before this data is merged with the residual data to form the final output image

3.1.40
legacy decoder
embodiment of a decoding process conforming to Rec. ITU-T T.81 | ISO/IEC 10918-1, confined to the lossy DCT process and the baseline, sequential, or progressive modes, decoding at most four components to eight bits per component

3.1.41
legacy image
arrangement of sample values as described by applying the decoding process described by Rec. ITU-T T.81 | ISO/IEC 10918-1 on the entropy coded data as defined by said standard

3.1.42
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)

3.1.43
lossless coding
mode of operation which refers to any one of the coding processes defined in this part of ISO/IEC 18477 in which all of the procedures are lossless

Note 1 to entry: See ISO/IEC 18477-8.

© ISO/IEC 2016 – All rights reserved

3.1.44
lossy
descriptive term for encoding and decoding processes which are not lossless

3.1.45
low dynamic range
image or image data comprised of data with no more than eight bits per sample

3.1.46
marker
two-byte code in which the first byte is hexadecimal FF and the second byte is a value between 1 and hexadecimal FE

3.1.47
marker segment
marker together with its associated set of parameters

3.1.48
pixel
collection of sample values in the spatial image domain having all the same sample coordinates, e.g. a pixel may consist of three samples describing its red, green, and blue value

3.1.49
precision
number of bits allocated to a particular sample or DCT coefficient

3.1.50
procedure
set of steps which accomplishes one of the tasks which comprise an encoding or decoding process

3.1.51
quantization value
integer value used in the quantization procedure

3.1.52
quantize
act of performing the quantization procedure for a DCT coefficient

3.1.53
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 base image to form the final output image

3.1.54
residual image
extension 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

3.1.55
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

3.1.56
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

© ISO/IEC 2016 – All rights reserved

3.1.57
sample
one element in the two-dimensional image array which comprises a component

3.1.58
sample grid
common coordinate system for all samples of an image

Note 1 to entry: 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.

3.1.59
scan
single pass through the data for one or more of the components in an image

3.1.60
scan header
marker segment that contains a start-of-scan marker and associated scan parameters that are coded at the beginning of a scan

3.1.61
superbox
box that carries other boxes as payload data

3.1.62
table specification data
coded representation from which the tables used in the encoder and decoder are generated and their destinations specified

3.1.63
(uniform) quantization
procedure by which DCT coefficients are linearly scaled in order to achieve compression

3.1.64
upsampling
procedure by which the spatial resolution of a component is increased

3.1.65
vertical sampling factor
relative number of vertical data units of a particular component with respect to the number of vertical data units in the other components in the frame

3.1.66
zero byte
0x00 byte

3.1.67
zig-zag sequence
specific sequential ordering of the DCT coefficients from (approximately) lowest spatial frequency to highest

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 in horizontal direction

 © ISO/IEC 2016 – All rights reserved

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

H_i Subsampling indicator of component in the frame header

V_i Subsampling indicator of component in the frame header

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

R_h Additional number of DCT coefficient bits represented by refinement scans in the legacy decoding path, $8+R_h$ 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 legacy decoding path.

R_r Additional number of DCT coefficient bits represented by refinement scans in the residual decoding path. $P+R_r$ is the number of non-fractional bits of the output of the invers DCT process in the residual decoding path, where P is the frame-precision of the residual image as recorded 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

For the purposes of this part of ISO/IEC 18477, the following abbreviated terms apply.

ASCII American Standard Code for Information Interchange

LSB Least Significant Bit

MSB Most Significant Bit

HDR High Dynamic Range

IDR Intermediate Dynamic Range

LDR Low Dynamic Range

TMO Tone Mapping Operator

DCT Discrete Cosine Transformation

4 Conventions

4.1 Conformance language

This part of ISO/IEC 18477 consists of normative and informative text.

Normative text is that text which expresses mandatory requirements. The word "shall" is used to express mandatory requirements strictly to be followed in order to conform to this part of ISO/IEC 18477 and from which no deviation is permitted. A conforming implementation is one that fulfils all mandatory requirements.

Informative text is text that is potentially helpful to the user, but not indispensable and can be removed, changed, or added editorially without affecting interoperability. All text in this part of ISO/IEC 18477 is normative, with the following exceptions: The Introduction, any parts of the text that are explicitly labelled as "informative", and statements appearing with the preamble "NOTE" and behaviour described using the word "should". The word "should" is used to describe behaviour that is encouraged but is not required for conformance to this part of ISO/IEC 18477.

The keywords "may" and "need not" indicate a course of action that is permissible in a conforming implementation.

© ISO/IEC 2016 – All rights reserved