INTERNATIONAL STANDARD

ISO/IEC 18477-1

Second edition
2020-05

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

Part 1:
Core coding system specification

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

Partie 1: Spécification du système de codage de noyau

© ISO/IEC 2020

iTeh STANDARD PREVIEW
(standards.iteh.ai)

COPYRIGHT PROTECTED DOCUMENT

Contents

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 18477-1:2015), which has been technically revised.

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

— Annex A.3 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 photographic content. JPEG XT is a scalable image coding system that builds on the legacy Rec. ITU-T T.81 | ISO/IEC 10918-1 coding system, also known as JPEG, but extends it in a backwards compatible way. This document specifies the commonly deployed components of the JPEG coding system. Additional parts of the ISO/IEC 18477 series extend on this baseline.

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, Rec. ITU-T T.86 | ISO/IEC 10918-4 and Rec. ITU-T T.871 | ISO/IEC 10918-5 to extend the functionality of the legacy JPEG coding system. It is optimized for good image quality and compression efficiency while also enabling low-complexity encoding and decoding implementations.

iTeh STANDARD PREVIEW

(standards.iteh.ai)

v

iTeh STANDARD PREVIEW
(standards.iteh.ai)

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

Part 1:
Core coding system specification

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 the core coding system, which forms the basis for the entire ISO/IEC 18477 series.

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.

Rec. ITU-T T.81 | ISO/IEC 10918-1:1994, *Information technology — Digital compression and coding of continuous-tone still images — Part 1: Requirements and guidelines*

Rec. ITU-T T.86 | ISO/IEC 10918-4, *Information technology — Digital compression and coding of continuous-tone still images — Part 4: Registration of JPEG profiles, SPIFF profiles, SPIFF tags, SPIFF colour spaces, APPn markers, SPIFF compression types and Registration Authorities (REGAUT)*

Rec. ITU-T T.871 | ISO/IEC 10918-5, *Information technology — Digital compression and coding of continuous-tone still images — Part 5: JPEG File Interchange Format (JFIF)*

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 https://www.iso.org/obp

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

3.1
bitstream
partially encoded or decoded sequence of bits comprising an entropy-coded segment

3.2
block
8×8 array of samples or an 8×8 array of DCT coefficient values of one component

3.3
byte
group of 8 bits

3.4
coder
embodiment of a coding process

3.5
coding
encoding or decoding

3.6
compression
reduction in the number of bits used to represent source image data

3.7
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.

3.8
continuous-tone image
image whose components have more than one bit per sample

3.9
discrete cosine transform
DCT
either the forward discrete cosine transform or the inverse discrete cosine transform

3.10
downsampling
procedure by which the spatial resolution of a component is reduced

3.11
entropy-coded data segment
independently decodable sequence of entropy encoded bytes of compressed image data

3.12
marker
two-byte code in which the first byte is hexadecimal FF and the second byte is a value between 1 and hexadecimal FE

3.13
marker segment
marker and associated set of parameters

3.14
precision
number of bits allocated to a particular sample or DCT coefficient

3.15
procedure
set of steps which accomplishes one of the tasks which comprise an encoding or decoding process

3.16
sample
one element in the two-dimensional array which comprises a component

3.17
sample grid
common coordinate system for all samples of an image with the samples at the top left edge of the image having the coordinates (0, 0), the first coordinate increases towards the right, the second to the bottom

3.18
scan
single pass through the data for one or more of the components in an image

3.19
scan header
marker segment that contains a start-of-scan marker and associated scan parameters that are coded at the beginning of a scan

3.20
upsampling
procedure by which the spatial resolution of a component is increased

3.21
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

4 Symbols and abbreviated terms

4.1 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

4.2 Abbreviated terms

ASCII	American Standard Code for Information Interchange
DC	lowpass
AC	highpass
LSB	least significant bit
MSB	most significant bit
DCT	discrete cosine transformation
JPEG	joint photographic experts group

5 Conventions

5.1 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.

5.2 Operators

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

5.2.1 Arithmetic operators

+   addition

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

×   multiplication

/   division without truncation or rounding

5.2.2 Assignment operators

=   assignment operator

5.2.3 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
+, −	addition and subtraction	left to right
<, >, <=, >=	relational	left to right

5.2.4 Mathematical functions

$\lceil x \rceil$ Ceiling 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, zero 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.

6   General

6.1   General definitions

The purpose of this clause is to give an informative overview of the elements specified in this document. Another purpose is to introduce many of the terms which are defined in Clause 3. These terms are printed in *italics* upon first usage in this clause.

There are three elements specified in this document:

a) An *encoder* is an embodiment of an *encoding process*. An encoder takes as input *digital source image data* and *encoder specifications*, and by means of a specified set of *procedures* generates as output *a codestream*.

b) A *decoder* is an embodiment of a *decoding process*. A decoder takes as input a codestream, and by means of a specified set of procedures generates as output *digital reconstructed image data*.

c) The *codestream* is a compressed image data representation which includes all necessary data to allow a (full or approximate) reconstruction of the sample values of a digital image. Additional data might be required that define the interpretation of the sample data, such as the spatial dimensions of the samples.

6.2   Functional overview on the decoding process

The high-level algorithm for decoding is as follows: The samples are first reconstructed following the decoder specifications defined in Rec. ITU-T T.81 | ISO/IEC 10918-1. If the resulting component arrays are subsampled, they are upsampled on a common sample grid following the specifications in Annex A. Following that, the output data is processed by an inverse decorrelation transformation. If the data is already in an RGB type colour space, e.g. RGB with ITU-R Rec. BT.601 primaries, this transformation will be the identity transformation. Otherwise, either the ICT is used to transform the data into RGB. The inverse decorrelation transformation is defined in Annex C, and the markers that are required to select the transformation are defined in Annex B.

6.3   Encoder requirements

An encoding process converts source image data to compressed image data. This includes first obtaining a low dynamic range image, and representing it by a coding process specified in Rec. ITU-T T.81 | ISO/IEC 10918-1:1994, Annex F or Annex G.

In order to comply with this document, an encoder shall satisfy at least one of the following two requirements. An encoder shall with appropriate accuracy, convert source image data to compressed image data which comply with the codestream format syntax specified in Annex B for the encoding process(es) embodied by the encoder. A limited accuracy sufficient to match the error bounds specified in the compliance tests is acceptable.

There is no requirement in this document that any encoder which embodies one of the encoding processes specified here shall be able to operate for all ranges of the parameters which are allowed for that process. An encoder is only required to meet the compliance tests and to generate the compressed data format according to Annex B for those parameter values which it does use.

6.4   Decoder requirements

A decoding process converts compressed image data to reconstructed image data. For that, it has to follow the decoding operation specified in Rec. ITU-T T.81 | ISO/IEC 10918-1 with sufficient accuracy, using either the baseline, sequential or progressive scan process defined in Rec. ITU-T T.81 | ISO/IEC 10918-1:1994, Annex F or Annex G. This process generates sample values on a sample grid, which are then converted into a digital image by following the upsampling specifications in Annex B and the multi-component decorrelation (ICT) process in Annex C.