~~ISO/IEC JTC 1/SC 29/WG1~~ ~~N100475~~

~~99th meeting - Online - April 24-28, 2023~~

~~ISO/IEC JTC 1/SC 29/WG 1~~

~~(& ITU-T SG16)~~

~~Coding of Still Pictures~~

~~JBIG~~ ~~JPEG~~

~~Joint Bi-level Image~~ ~~Joint Photographic~~

~~Experts Group~~ ~~Experts Group~~

~~TITLE:~~ ~~Text of~~ **ISO/IEC FDIS 18477-3**

Secretariat: JISC

Date: 2023-08-08

Information technology

~~Information Technology:~~ — Scalable compression and coding of continuous-tone still images ~~— — —~~

Part 3:
Box file format

~~Draft ITU-T Recommendation | International Standard~~

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

Partie 3: Format de la liste de fichiers

FDIS stage

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: + 41 22 749 01 11
E-mail: copyright@iso.org
Website: www.iso.org

Published in Switzerland

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~~www.iso.org/directives or ~~www.iec.ch/members_experts/refdocs~~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~~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~~www.iso.org/iso/foreword.html. In the IEC, see ~~www.iec.ch/understanding-standards~~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 18477-3:2015), which has been technically revised.

The main changes are as follows:

— ~~This edition provides~~ editorial improvements on the usage of the JPEG XT marker segment.

A list of all parts in the ISO/IEC 18477 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~~www.iso.org/members.html and www.iec.ch/national-committees.

Introduction

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,~~This means that 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~~although it is possible that they will ~~possibly~~ not be able to reconstruct such streams in full dynamic range~~, full~~ or quality or using other features defined in this document.

~~The aim of this standard is to provide~~This document provides a flexible and extensible framework to enrich ISO/IEC 18477-1 ~~compliant~~conforming codestreams with side-channels and metadata. The syntax chosen in this document defines a mechanism to embed syntax elements denoted as "boxes" into Rec. ITU-T T.81 | ISO/IEC 10918-1 ~~compliant~~conforming codestreams. The box syntax used ~~here~~in this document is identical to that defined in the JPEG ~~family of standards~~series, for example JPEG 2000 image coding system (Rec. ITU-T T.800 | ISO/IEC 15444-1). Boxes will then carry either additional image data ~~–~~, to enable encoding of images of higher ~~bitdepth~~bit depth, high-dynamic range~~, include~~ (HDR), including alpha channels, etc. ~~— ,~~ or ~~will carry~~ metadata that describes the decoding process of the legacy Rec. ITU-T T.81 | ISO/IEC 10918-1 codestream and the side channels to an extended or ~~high dynamic range~~HDR image.

This document specifies an extensible file format, denoted as JPEG XT, which is built on top of the existing Rec. ITU-T T.81 | ISO/IEC 10918-1 codestream definition. While typically, file formats encapsulate codestreams by means of additional syntax elements such as boxes, the file format structure specified ~~here rather~~in this document embeds the syntax elements of the file format, called boxes, into the codestream. The necessity for this unusual arrangement is the backwards compatibility to the legacy standard and the application toolchain built around it~~;~~. This means that ~~is,~~ legacy applications conforming to Rec. ITU-T T.81 | ISO/IEC 10918-1 will be able to decode image information embedded in files conforming to this document, ~~though~~although they will only be able to recover a three component, 8 ~~bits~~bit per sample, lower quality version of the image described by the full file.

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 ~~*~~× 10^{14} representable colour values. In some less common scenarios, even greater bit depths are used, and sometimes the dynamic range of the image is so high that a floating-point based encoding is desirable. In addition to image information, some applications also require an additional opacity channel, a feature not available from the legacy standard.

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~~can 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~~can 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 designed to extend the legacy JPEG standard towards higher ~~bitdepth~~bit depth, higher dynamic range, and wide colour gamut content, while simultaneously allowing legacy applications to decode the image data in the codestream to a standard low-dynamic range (LDR) image represented by only ~~eight~~ 8 bits per channel. 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.

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

Part 3:
Box file format

1 Scope

This document specifies a codingbox-based container format, referred to as JPEG XT, which is designed primarily for continuous-tone photographic content.

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:, *Information technology: — Scalable compression and coding of continuous-tone still images: — Part 1: Core coding system specification*

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

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

~~ISO/IEC 10646-1 Annex D *Information Technology—Universal Multiple-Octet Coded Character Set (UCS)—part 1: Architecture and Basic Multilingual Plane*~~

43 Terms, definitions, abbreviated terms and symbols

4.13.1 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~~https://www.iso.org/obp

— ——IEC Electropedia: available at ~~https://www.electropedia.org/~~https://www.electropedia.org/

3.1.1
ASCII encoding
: encoding of text characters and text strings according to ISO/IEC 10646-1

3.1.2
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.

3.1.3
base image
: collection of sample values obtained by entropy, decoding the discrete cosine transformation (DCT) coefficients of the legacy codestream and the refinement codestream, and inversely DCT transforming them jointly.

~~1.1.1~~
~~**binary decision: choice between two alternatives.**~~

~~1.1.2~~
3.1.4
bit stream
: partially encoded or decoded sequence of bits comprising an entropy-coded segment.

~~**block: 8×8 array of samples or an 8×8 array of DCT coefficient values of one component.**~~

~~1.1.3~~
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. ~~See Annex A~~

Note 1 to entry: See Annex A for the definition of boxes.

3.1.6
byte
: group of 8 bits.

~~1.1.3~~
~~**coder: embodiment of a coding process.**~~

~~1.1.3~~
~~**coding: encoding or decoding.**~~

~~1.1.3~~
~~**coding model: procedure used to convert input data into symbols to be coded.**~~

~~1.1.4~~
~~(coding) process: general term for referring to an encoding process, a decoding process, or both.~~

~~1.1.5~~
3.1.7
compression
~~:~~ reduction in the number of bits used to represent source image data~~.~~

3.1.8
component
~~:~~ two-dimensional array of samples having the same designation in the output or display device~~.~~

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

3.1.9
continuous-tone image
~~:~~ image whose components have more than one bit per sample~~.~~

3.1.10
decoder~~: embodiment of a decoding process.~~
embodiment of a decoding process~~:~~

3.1.11
decoding process
process which takes as its input compressed image data and outputs a continuous-tone image~~.~~

3.1.12
encoder~~: embodiment of an encoding process.~~
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~~.~~

~~entropy decoder: embodiment of an entropy decoding procedure.~~

~~1.1.6~~
~~entropy decoding: lossless procedure which recovers the sequence of symbols from the sequence of bits produced by the entropy encoder.~~

~~1.1.6~~
~~entropy encoder: embodiment of an entropy encoding procedure.~~

~~1.1.7~~
~~**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.**~~

~~1.1.8~~
3.1.14
extension image~~**: residual image (see there).**~~
sample values as reconstructed by inverse quantization and inverse discrete cosine transformation (DCT) applied to the entropy-decoded coefficients described by the refinement scan, residual scan and residual refinement scans

3.1.15
high-dynamic range~~:~~
HDR
image or image data comprised of more than ~~**eight bits per sample.**~~

~~**Intermediate dynamic range:** image or image data comprised of more than eight~~ 8 bits per sample.

~~1.1.9~~
3.1.16
legacy codestream
~~:~~ collection of markers and syntax elements defined by Rec. ITU-T T.81 | ISO/IEC 10918-1 ~~bare~~without any syntax elements defined by ~~the family~~ISO/IEC 18477-1, ISO/IEC 18477~~ standards.~~ -2, ISO/IEC 18477-3

~~NOTE~~ Note 1 to entry: ~~By~~ In this definition, 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 ~~Annex A.~~ Annex A.

~~1.1.9~~
~~**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.**~~

~~1.1.10~~
3.1.17
legacy decoder
~~:~~ embodiment of a decoding process conforming to ~~ITU-T~~ Rec. ITU-T T.81 | ISO/IEC 10918-1, confined to the lossy discrete cosine transformation (DCT) process and the baseline, sequential or progressive modes, decoding at most four components to ~~eight~~ 8 bits per component~~.~~

~~**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.**~~

~~1.1.11~~
3.1.18
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~~).~~)

~~1.1.11~~
~~**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**~~

~~1.1.11~~
3.1.19
lossy
~~: descriptive term for~~ encoding and decoding processes which are not lossless~~.~~

3.1.20
low-dynamic range~~:~~
LDR
image or image data comprised of data with no more than ~~eight~~ 8 bits per sample~~.~~

3.1.21
marker~~: two~~
2-byte code in which the first byte is hexadecimal FF and the second byte is a value between 1 and hexadecimal FE~~.~~

3.1.22
marker segment
~~:~~ marker together with its associated set of parameters~~.~~

3.1.23
pixel
~~:~~ collection of sample values in the spatial image domain having all the same sample coordinates~~, e. g.~~

Note 1 to entry: A pixel may consist of three samples describing its red, green and blue value.

3.1.24
point ~~transform:~~ **transformation**
scaling of a sample or discrete cosine transformation (DCT) coefficient by a factor.

3.1.25
precision
~~:~~ number of bits allocated to a particular sample or discrete cosine transformation (DCT) coefficient.

3.1.26
procedure
~~:~~ set of steps which accomplishes one of the tasks which ~~comprise~~comprises an encoding or decoding process.

3.1.27
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.

3.1.28
residual image
~~:~~ sample values as reconstructed by inverse quantization and inverse ~~DCT~~discrete cosine transformation (DCT) applied to the entropy-decoded coefficients described by the residual scan and residual refinement scans.

~~residual~~**3.1.29**
refinement scan
~~:~~ **additional pass over the image data that is invisible to legacy decoders** ~~which provides additive and/or multiplicative correction data of the base scans to allow reproduction of high dynamic range or wide colour gamut data.~~

~~1.1.12~~
~~**refinement scan:** additional pass over the image data invisible to legacy decoders,~~ which provides additional least significant bits to extend the precision of the discrete cosine transformation (DCT) transformed coefficients.

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

3.1.30
sample
~~:~~ one element in the two-dimensional image array which comprises a component.

3.1.31
sample grid
~~:~~ common coordinate system for all samples of an image such that 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.