INTERNATIONAL STANDARD

ISO/IEC 18477-9

First edition
2016-10-15

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

Part 9: Alpha channel coding

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

Reference number
ISO/IEC 18477-9:2016(E)

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

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 World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

The committee responsible for this document is ISO/IEC JTC 1, *Information technology*, SC29, *Coding of audio, picture, multimedia and hypermedia information*.

A list of all parts in the ISO 18477- series, published under the general title *Information technology — Scalable compression and coding of continuous-tone still images*, can be found on the ISO website.

Introduction

This document specifies an extension for ISO/IEC 18477-3 compliant files that adds capabilities for lossy or lossless storage of continuous or binary opacity information associated to the image; such additional channels are commonly known as alpha channels. These channels are used for compositing the image content with other content on the same physical media. An alpha value of 0 encodes maximal transparency (and no opacity), while the maximal sample value represents maximal opacity (and no transparency). Additionally, the image content itself may be *premultiplied* with the alpha value or *premultiplied and shaded with a background colour M*, a process by which the original image A is replaced by the image A' defined as

$A' = \alpha * A$ for pre-multiplication

$A' = \alpha * A + (1-\alpha) * M$ for pre-multiplication and shading

And A' is encoded instead of A in the JPEG XT codestream. Reconstruction is then performed as follows: If A denotes the sample value of the image contained in the ISO/IEC 18477-3 file at a specific spatial location, B is the sample value of the background on which the image should be rendered, M is the matte colour and α is the decoded value of the alpha channel, then the sample value of the image C composed from A and B on the same position is given by:

$C = \alpha * A + (1-\alpha) * B$ for non-premultiplied content;

$C = A + (1-\alpha) * B$ for premultiplied content;

$C = A + (1-\alpha) * (B - M)$ for premultiplied content with shade removal.

Encoding a *premultiplied* and shaded version of A' with colour M enables legacy decoders that lack alpha channel support to still decode and display the image with the appearance that it is composited on a background with colour M. At the same time, new JPEG XT compliant decoders can composite the image on any background by calculating image C from A, B and M.

This document provides facilities to encode the value of α for each spatial location, with or without loss, either as a binary decision, i.e. α = 0 or α = 1, on a continuous scale of integers with a resolution between 8 and 16 bits, or as floating point number between 0 and 1 with 16-bit precision. It uses coding technology from other parts of the ISO/IEC 18477 family of standards for its encoding, and no new technology besides that already defined in other parts is required for the reconstruction of the opacity information.

This document can be freely combined with other parts of the ISO/IEC 18477 family, i.e. the sample values A in the above formulae might be either 8-bit unsigned integers, i.e. represented by ISO/IEC 18477-1, up to 16-bit integers using the encoding of ISO/IEC 18477-6 or floating point values encoded by ISO/IEC 18477-7. The image content A may also be encoded without loss, using ISO/IEC 18477-8. However, the compositing step itself to create the final output image C from the input images A and B is not standardized.

The syntax of the codestream defined in this document is fully backward compatible to Rec. ITU-T T.81 | ISO/IEC 10918-1 and the ISO/IEC 18477 family of standards. Decoders unaware of the extensions defined here will reconstruct a fully opaque version of the image and discard the alpha channel content.

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

Part 9:
Alpha channel coding

1 Scope

This document specifies a coding 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: 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 18477-7:2016, *Information technology — Scalable compression and coding of continuous-tone still images — Part 7: HDR Floating-Point Coding*

ISO/IEC 18477-8:2016, *Information Technology: Scalable compression and coding of continuous-tone still images — Lossless and near-lossless coding*

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

3 Terms, definitions, symbols and abbreviated terms

3.1 Terms and definitions

For the purposes of this document, the following 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
ASCII encoding
encoding of text characters and text strings according to ISO/IEC 10646

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 DCT coefficients of the legacy codestream and the refinement codestream, and inversely DCT transforming them jointly

3.1.4
alpha channel
additional scalar image channel that encodes the opacity of each sample in the main image

3.1.5
alpha component
synonym for alpha channel

3.1.6
binary decision
choice between two alternatives

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 process
general reference to an encoding process, a decoding process, or both

3.1.13
compression
reduction in the number of bits used to represent source image data

3.1.14
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.

3.1.15
composition
process of merging the decoded image data with background image data using opacity information and generating one single final output image

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

3.1.17
decoder
embodiment of a decoding process

3.1.18
decoding process
process which takes as its input compressed image data and outputs a continuous-tone image

3.1.19
encoder
embodiment of an encoding process

3.1.20
encoding process
process which takes as its input a continuous-tone image and outputs compressed image data

3.1.21
entropy decoder
embodiment of an entropy decoding procedure

3.1.22
entropy decoding
lossless procedure which recovers the sequence of symbols from the sequence of bits produced by the entropy encoder

3.1.23
entropy encoder
embodiment of an entropy encoding procedure

3.1.24
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

3.1.25
fixed point discrete cosine transformation
fixed point DCT
implementation of the discrete cosine transformation based on fixed point arithmetic following the specifications in ISO/IEC 18477-8:2016, Annex E

3.1.26
high dynamic range
HDR
image or image data comprised of more than eight bits per sample

3.1.27
integer based discrete cosine transformation
integer point 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 as specified in ISO/IEC 18477-8:2016, Annex E

3.1.28
joint photographic experts group
JPEG
informal name of the committee which created this document

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

3.1.29
legacy codestream
collection of markers and syntax elements defined by Rec. ITU-T T.81 | ISO/IEC 10918-1 bare any syntax elements defined by the ISO/IEC 18477 family of standards. That is, 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

3.1.30
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

3.1.31
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

3.1.32
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)

3.1.33
lossless coding
mode of operation which refers to any one of the coding processes defined in ISO/IEC 18477-8:2016 in which all of the procedures are lossless

3.1.34
lossy
encoding and decoding processes which are not lossless

3.1.35
low-dynamic range
LDR
image or image data comprised of data with no more than eight bits per sample

3.1.36
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

3.1.37
point transformation
application of a location independent global function to reconstructed sample values in the spatial domain

3.1.38
precision
number of bits allocated to a particular sample or DCT coefficient

3.1.39
premultiplied content
image component that has already multiplied by the scaled value of the alpha channel on a pixel-by-pixel basis to ease the composition of the image with the background

3.1.40
procedure
set of steps which accomplishes one of the tasks which comprise an encoding or decoding process

3.1.41
quantize
act of performing the quantization procedure for a DCT coefficient

3.1.42
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

3.1.43
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

3.1.44
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

3.1.45
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.

3.1.46
sample
one element in the two-dimensional image array which comprises a component

3.1.47
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.

3.1.48
superbox
box that carries other boxes as payload data

3.1.49
sub box
box that is contained as payload data within a superbox

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, 8+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.

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

DCT discrete cosine transformation

LSB least significant bit

MSB most significant bit

TMO tone mapping operator

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.