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)

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. $\alpha = 0$ or $\alpha = 1$, 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.

4.2.1 Arithmetic operators

+ addition

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

* multiplication

/ division without truncation or rounding.

umod x umod a is the unique value y between 0 and a−1for 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 below 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 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	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
$\lvert x \rvert$	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

5 General

5.1 General definitions

Clause 5 gives an informative overview of the elements specified in this document. It also introduces many of the terms which are defined in Clause 3. These terms are printed in *italics* upon first usage in Clause 5.

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 colour space or the spatial dimensions of the samples.

5.2 Overview of ISO/IEC 18477-9

5.2.1 Encoder requirements

An encoder is only required to meet the compliance tests and to generate the codestream according to the syntax defined in this document. How the codestream is algorithmically constructed and how the boxes are laid out is implementation specific and not within scope of this document. Subsequent parts of the ISO/IEC 18477 series may, however, define additional restrictions and requirements, either within the document itself, or within profiles that restrict the freedom of the encoder further.

An encoder claiming to be compliant to one of these profiles then shall conform to the syntax constraints defined in the corresponding profile of the corresponding part of ISO/IEC 18477.

5.2.2 Decoder requirements

A decoding process converts compressed image data to reconstructed image data. It shall follow the decoding operation specified in this document and ISO/IEC 18477-1 to reconstruct a legacy 8 bits per channel standard low dynamic range image. It is **not required** that a conforming decoder is capable of decoding and interpreting all box types defined in this or other members of the ISO/IEC 18477 family of standards. A decoder implementation is always free to skip over box types it is unable or not willing to support.

In order to comply with this document, a decoder:

a) **shall convert** a codestream conforming to this document **without considering any boxes** into to a low dynamic range image;

b) **shall additionally** convert a conforming codestream **including** the information in a subset of the boxes into an image (of potentially higher precision, higher quality or higher bit depths) and into an alpha-channel;

c) shall implement **at least** all the functional blocks of the JPEG XT decoding process defined in the profile it claims to be conforming to, where profiles are defined in this and other parts of the ISO/IEC 18477 family of standards. For that, a conforming decoder shall correctly interpret all box types required in the definition of the profile.

Annex A
(normative)

Encoding and decoding process

A.1 Decoding process (normative)

Annex A defines the functionality of a subset of the boxes of the file format, as specified in ISO/IEC 18477-3, which are required for decoding an alpha channel. The alpha channel defines opacity data accompanying the image represented by the respective legacy codestream and boxes defined in other parts of the ISO/IEC 18477 family of standards.

The decoding process of the alpha channel data is depicted in Figure A.1. The decoder first decodes the foreground image contained in the codestream and boxes of the ISO/IEC 18477-3 conforming file, giving one or three components per sample process. The decoding process of the foreground image from the codestream is specified in other parts of the ISO/IEC 18477 family of standards and not repeated here. The decoder then proceeds to decode the codestream in the Alpha Codestream box B1 and the Alpha refinement box B1a if they are present, giving a precursor alpha channel plane denoted by H_i. The frame dimensions indicated in the frame header of the Alpha Codestream box (see Annex B) and the frame header of the Residual Alpha Data box shall be identical to the frame dimensions indicated in the frame header of the legacy codestream. Decoding then proceeds to the Residual Alpha box B5, and the Residual Alpha Refinement box B5a. If decoded, they provide either a lossless error residual, denoted by $Q_{i,}$ to enable lossless coding of the alpha channel. The base alpha channel H_i is added to the residual alpha channel Q_i forming the intermediate output F_i, which is then converted or scaled to range, giving a number between 0 and 1. This is the output of the decoding process of this document.

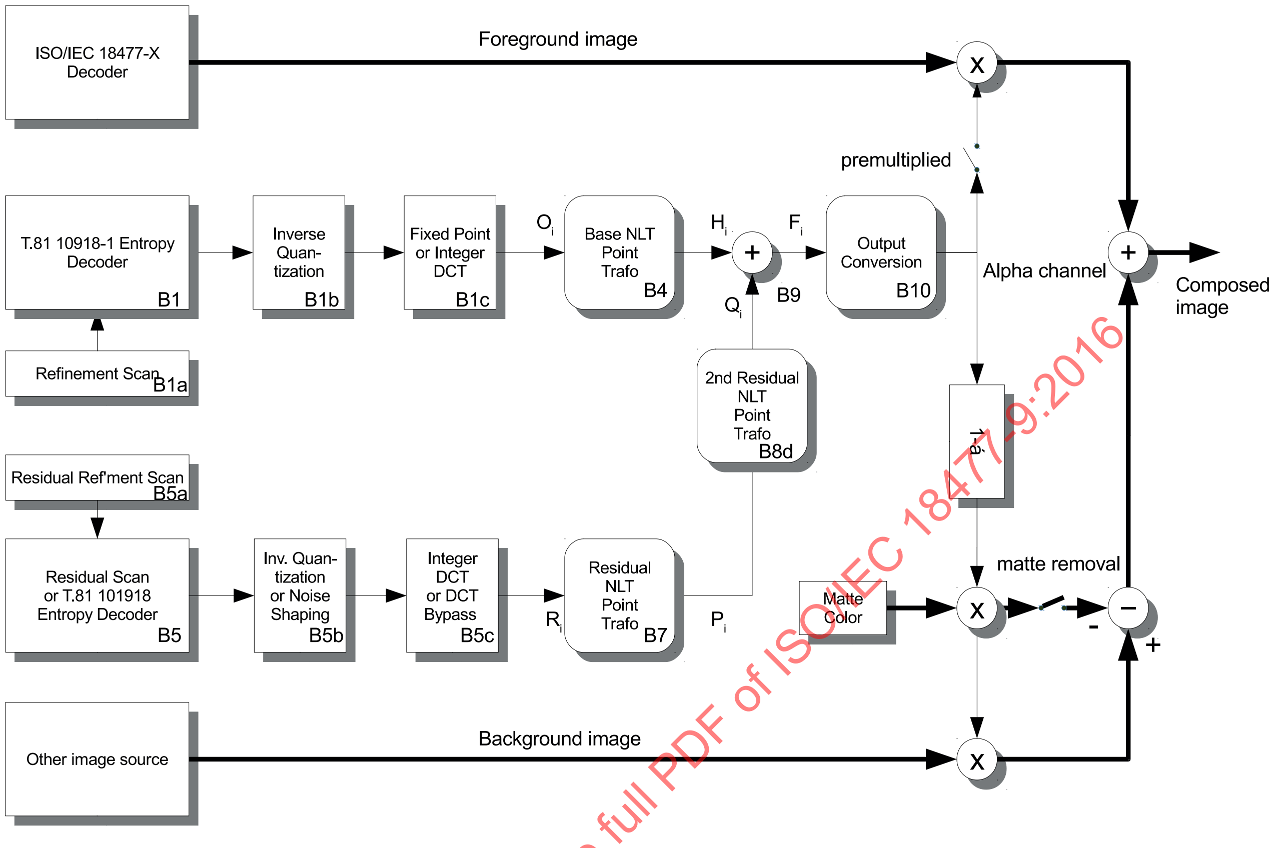

Figure A.1 — High-level overview of the JPEG XT decoding of alpha channels and image composition

A.2 Reconstruction of the alpha channel (normative)

The following steps shall be followed to reconstruct the sample values α of the alpha channel:

— In step B1, the content of the Alpha Codestream box is entropy decoded, by means of one of the Huffman coding modes of ISO/IEC 18477-1, i.e. by the baseline sequential, extended sequential or progressive modes of Rec. ITU-T T.81 | ISO/IEC 10918-1. If this codestream is present, the width and the height of the sample grid described by the alpha channel codestream shall match those of the foreground image, and the number of components of the alpha channel codestream shall be 1.

— In step B1a, decoding the content of the Alpha Refinement box, if it is present, refines the precision of the reconstructed alpha-DCT-coefficients. This box shall only exist if the Alpha Codestream box is present. Entropy decoding of the contents of the Alpha Refinement box shall follow the specifications of Refinement Coding as defined in ISO/IEC 18477-6:2016, Annex D.

— In steps B1b and B1c, the data are inversely quantized and inversely DCT transformed. The Lf flag of the Alpha Output Conversion box and the Alpha Base DCT box define the DCT transformation to select. See Annex B for details. The DCT transformations are either defined by Rec. ITU-T T.81 | ISO/IEC 10918-1 or ISO/IEC 18477-8 and are selected according to the above boxes. The output of this process is the reconstructed base alpha sample values O_i.

— In step B4, an optional nonlinear point-transformation is applied to the sample values O_i generating the precursor alpha channel sample values H_i. This transformation is selected by the Base Transformation box within the Alpha Merging Specification box if it is present, or is a scaling process in its absence. The transformation selected by the Base Transformation box is either an Integer or

Floating Point Lookup Table, or a Parametric Curve box describing the transformation. If lossless coding of alpha channels is required, then either the Base Transformation box shall reference an Integer Table Lookup table, or no Base Transformation box shall be present.

— In step B5, the contents of the Residual Alpha Codestream box is entropy decoded, if it is present. This uses either one of the coding modes of ISO/IEC 18477-1 (sequential baseline, extended sequential or progressive), or the DCT bypass or large-range DCT coding mode specified in Annex D of ISO/IEC 18477-8:2016, Annex D. The dimensions of the image described by such an optional box shall be identical to the dimensions of the foreground image, and the Number of Components Nf of this image shall be 1.

— In step B5a, decoding the contents of the Residual Alpha Refinements box, if it is present, refines the precision of the reconstructed residual alpha-DCT-coefficients. This box shall only exist if the Residual Alpha Codestream box is present. Entropy decoding of the contents of the Residual Alpha Refinement box shall follow either:

  — the specifications of ISO/IEC 18477-6:2016, Annex D if the Residual Alpha Codestream box uses one of the coding modes of ISO/IEC 18477-1, or

  — the specifications of the Residual Refinement Coding mode of ISO/IEC 18477-8:2016, Annex D if the Residual Alpha Codestream is encoded in the DCT-bypass or large-range DCT mode.

— In steps B5b and B5c, the residual data are inversely quantized and inversely DCT transformed; the DCT transformation is optionally bypassed in which case inverse quantization may be replaced by inverse noise shaping. The DCT process itself is specified by the Residual DCT box, see Annex B for details. The DCT transformation shall conform to Rec. ITU-T T.81 | ISO/IEC 10918-1, if the Lf flag of the Output Conversion box is 0. Otherwise, the Residual DCT box determines the DCT transform as being one of the DCT processes specified in ISO/IEC 18477-8:2016, Annex E. The outputs of this process are residual alpha channel sample values R_i.

— In step B7, the residual alpha channel data R_i optionally undergoes a nonlinear point transformation, selected by the Residual Nonlinear Point Transformation box. This transformation is either an Integer or Floating Point Lookup Table, or a Parametric Curve box describing the nonlinearity. If this box is absent, then the input is scaled to match the output. If lossless encoding of alpha channels is desired, then this box shall only reference an Integer Table Lookup, or shall not be present at all. The outputs of this step are the Alpha channel residual sample values P_i.

— In step B8d, a secondary nonlinear point transformation selected by the **Secondary Residual Nonlinear Point Transformation box** is applied. This box selects either a Floating Point Lookup box or a Parametric Curve box that maps the inversely decorrelated residual sample values P_i into the final residual errors Q_i. If this box is not present, Q_i is set to P_i.

— Step B9 merges the final alpha base sample values H_i and the final alpha residual sample values P_i to form the intermediate alpha output values F_i by

  $Fi: = H_i + Q_i - 2^{Rb+8-1}$ umod 2^{Rb+8} if the Lf flag of the Alpha Output Conversion box is 1; or

  $Fi: = H_i + Q_i - 2^{Rb+8-1}$ if the Lf flag of the Alpha Output Conversion box is 0.

  The residual alpha channel box is optional. In case it does not exist, the implied value Q_i shall be 2^{Rb+8-1} where the value of R_b is taken from the Output Conversion box within the Alpha Merging Specification box.

— The sample intermediate alpha values F_i are further processed according to an algorithm defined by the Output conversion box. This conversion is either a scale to unit range, a pseudo-logarithmic map from ISO/IEC 18477-7:2016, Annex D or a generic nonlinear mapping defined by a Parametric Curve box. See Annex B for details. The output of this process is the final alpha channel. This output conversion process transforms the sample from its bit precision to a number in the range [0..1].

A.3 Composition of foreground and background image (informative)

Further processing of the reconstructed alpha channel is outside the scope of this document. Further processing should be steered, however, by the Alpha Channel Composition box specified in ISO/IEC 18477-3:2015, Annex B. It defines four processes by which the foreground image can be composed with the background image using the alpha channel. All four processes operate on a pixel by pixel basis, i.e. take the sample values of foreground image A, background image B and alpha channel α on the same image location and compute from these three values the sample value on the same location of the final composed image C.

— If the $Acmp_0$ parameter of the Alpha Channel Composition box is 0, the foreground image is completely opaque and replaces the background image. The values reconstructed from the alpha channel and the background image are then ignored.

— If the $Acmp_0$ parameter is 1, the final composed image is the convex combination of the foreground image and the background image where the sample values α of the alpha channel weight between foreground and background. If α is 1, the foreground image is completely opaque, if α is 0, the foreground image is completely transparent.

— If the $Acmp_1$ parameter is 2, the foreground image is already pre-multiplied by α and the final image sample values are then computed as the sum of the foreground sample values and (1−α) times the background sample values.

— If the $Acmp_1$ parameter is 3, then a matte colour value M is subtracted from the background sample value B before multiplying the difference with (1- α) and adding it to the foreground image. This procedure allows the generation of foreground images that blend into a constant colour different from black for opaque image regions and may offer better backward compatibility for applications that cannot honour the alpha channel. The matte colour M is also specified in the Alpha Channel Composition box, see ISO/IEC 18477-3:2015, Annex B.

Table A.1 summarizes all four compositing algorithms.

$Acmp_0$	Compositing method	Notes
0	C = A	For opaque foreground.
1	C = α*A +(1−α)*B	For non-premultiplied content.
2	C = A +(1−α)*B	For premultiplied content.
3	C = A +(1−α)*(B−M)	For premultiplied content with shade removal.
All other values		Reserved by ITU-T | ISO/IEC.

Table A.1 — Image composition algorithms

NOTE Premultiplication reduces the algorithmic complexity of the reconstruction process by computing the product of the image and the alpha channel already at the encoder side. It replaces transparent image regions with black colour. If this is undesirable, the black colour can be replaced by any other colour M at encoder side that is then removed by the compositing process by subtracting it from the background.

Annex B
(normative)

Boxes

B.1 General

Annex B defines selects and refines a subset of the boxes defined in ISO/IEC 18477-3 for the purpose of representing alpha channels. It lists those boxes of ISO/IEC 18477-3 that are required for this document. All other boxes are optional and its interpretation is outside the scope of this document. Other parts of ISO/IEC 18477 or other standards may define their meaning, and decoders conforming to this document may ignore them.

Table B.1 lists the boxes required in this document that are paired with ISO/IEC 18477-3:2015, Annex B. Some of the boxes require additional specifications that are listed in subsequent clauses of this Annex.

Box name	Box type	Further definitions in subclause of ISO/IEC 18477-9
File Type box	0x66747970 ("ftyp")	
Legacy Data Checksum box	0x4C43484B ("LCHK")	
Alpha Codestream box	0x414C4641 ("ALFA")	
Alpha Refinement box	0x4146494E ("AFIN")	
Residual Alpha Data box	0x41524553 ("ARES")	
Residual Alpha Refinement box	0x41525246 ("ARRF")	
Alpha Merging Specification box	0x41535043 ("ASPC")	
Refinement Specification box	0x52535043 ("RSPC")	
Parametric Curve box	0x43555256 ("CURV")	
Integer Table Lookup box	0x544f4e45 ("TONE")	
Output Conversion box	0x4F434F4E ("OCON")	B.2
Base Nonlinear Point Transformation Specification box	0x4C505453 ("LPTS")	B.3
Residual Nonlinear Point Transformation Specification box	0x51525453 ("QPTS")	B.4
Secondary Residual Nonlinear Point Transformation Specification box	0x51525453 ("RPTS")	B.5
Base DCT Specification box	0x4C444354 ("LDCT")	B.6
Residual DCT Specification box	0x52444354 ("RDCT")	B.7
Alpha Channel Composition box	0x414D554C ("AMUL")	

Table B.1 — Boxes used by ISO/IEC 18477-9

B.2 Output Conversion box

This mandatory box defines the conversion process from the result of the base and residual opacity data merging process to the final opacity samples. It describes the final merging process and by that step B10 of the algorithm described in A.1. The purpose of this box is to specify the transformation of the decoded sample ranges to the unit range [0..1]. In specific, the box parameters shall be selected such that the reconstructed sample values are always within this interval.

This box is already defined in ISO/IEC 18477-3:2015, Annex B, though its application to this document further constrains the value of its fields.

This box shall never appear top level in the file, but it shall be a subbox of the Alpha Merging Specification box defined in ISO/IEC 18477-3:2015, Annex B. Exactly one Output Conversion box shall appear in the Alpha Merging Specification box.

The detailed list of operations to be performed by the Output Conversion is specified in ISO/IEC 18477-3:2015, Annex B. Table B.2 constrains the parameters of the Output Conversion box as applied in this document.

Parameter	Constraints within ISO/IEC 18477-9	Meaning
R_b	0..8	Number of additional bits available for Opacity data due to the Alpha Residual Codestream. The bit precision of the reconstructed opacity data shall be computed as 8+R_b. The parameter of this field shall be 8 if Oc is 1.
Lf	0..1	This field indicates whether lossless/near lossless or lossy coding of opacity data are intended. If this field is 0, lossy coding is intended and the decoder may pick any DCT implementation as long as it follows the constraints of Rec. ITU-T T.83 | ISO/IEC 10918-2. If this field is 1, implementations shall follow the DCT operation as selected by the Base and Residual DCT Specification Box, and as specified in ISO/IEC 18477-8:2016, Annex E. Furthermore, merging requires modulo arithmetic.
Oc	0..1	Half-exponential output-enable flag. If this flag is 1, the half-exponential map defined in ISO/IEC 18477-7:2016, Annex D shall be enabled. This step is performed after clipping the data to range, if enabled by the Ce flag, but before applying the output transformation, if enabled by the Ol flag. If this flag is reset, the output of the clipping stage is the input to the nonlinear transformation directly. If this flag is 1, the value of R_b shall be 8 and the Alpha Codestream Box shall be present, otherwise the value of R_b shall be in the range [0..8].

Parameter	Constraints within ISO/IEC 18477-9	Meaning
Ce	0..1	This field indicates whether the output of the merging process, i.e. the sum of base and residual image, shall be clipped to range $[0,2^{Rb}-1]$ if Oc is enabled, or [0,0x3c00] before processing the data any further. If the Ce flag and the Oc flag are both enabled, clipping is applied before conversion to floating point.
Ol	0..1	This field indicates whether an output lookup or point transformation is required. If enabled, the output transformation is specified by the to_0 through to_3 fields. If disabled, no further transformation is performed and the output of the clipping step and/or half-logarithmic map is already the opacity information. The output transformation by the Ol field is the final step of the output formation and applied after clipping and conversion to floating point.
to_0	0	If Ol is 1, this field defines the output table for the opacity data.
to_1	0	Unused, shall be 0.
to_2	0	Unused, shall be 0.
to_3	0	Unused, shall be 0.

Table B.2 — Parameter constraints for the Output Conversion box

The nonlinear point transformations selected by this box shall only be Parametric Curve boxes with a Curve Type t = 5 (linear Ramp). The corresponding boxes referenced by this box appear at top level of the ISO/IEC 18477-3 compliant file or as sub-boxes of the Merging Specification box. The nonlinear point transformation itself is given by the process specified in ISO/IEC 18477-3:2015, Annex C. It requires four additional parameters, the input range R_w, R_e and the output range R_t, R_f. The two value pairs shall be given as follows:

$R_w = 1$ $R_e = 0$;

$R_t = 1$ $R_f = 0$.

The value of the rounding mode e shall be 0. Note that the value of e is ignored for this specific setting.

B.3 Base Nonlinear Point Transformation Specification box

This box defines the nonlinear point transformation between the samples as reconstructed from the Alpha Codestream box O_i and the input H_i of the merging process with the residual alpha data. It thus defines step B4 in the decoder description in Annex A. Its box layout and box structure is given by the Nonlinear Point Transformation Specification box, defined in ISO/IEC 18477-3:2015, Annex B. This box shall only be present if the Alpha Codestream box is present, and it shall only exist as a subbox of the Alpha Merging Specification box.

Additional constraints apply, however: If the Lf flag of the Output Conversion box is 1, then the td_i values of the box shall only reference Integer Table Lookup boxes. References to Floating point Table Lookup boxes or Parametric Curve boxes shall not be used in this case. If Lf is 0, no such constraints apply. The nonlinear point transformation itself is given by the process specified in ISO/IEC 18477-3:2015, Annex C. It requires four additional parameters, the input range R_w, R_e and the output range R_t, R_f. The two value pairs shall be given as follows:

$R_w = 8 + R_h \qquad R_e = 0;$

$R_t = 8 + R_b \qquad R_f = 0.$

The value R_h is the number of refinement scans in the legacy decoding path and is found in the Refinement Specification box as subbox of the Alpha Merging Specification box. The Refinement Specification box is defined in ISO/IEC 18477-3:2015, Annex B. If the Refinement specification box is absent, the inferred value of R_h is 0. The value R_b is found in the Output Conversion box, where R_b+P_o defines the total output precision of the image.

If this box is not present but the Alpha Codestream box is, reconstructed sample v values from this box shall be scaled to output values w by

$w = v * 2^{Rt-Rw} \quad \text{if} \quad R_w \le R_t;$

$w = \lfloor v / 2^{Rw-Rt} \rfloor \quad \text{if} \quad R_w > R_t.$

This operation is identical to that of a Parametric Curve box of type t = 2 (Identity) with rounding mode e = 0.

The type of this box shall be 0x4C52505453, ASCII encoding of “LPTS”. The box structure and layout does not deviate from that in ISO/IEC 18477-3:2015, Annex B.

B.4 Residual Nonlinear Point Transformation Specification box

This box defines the nonlinear point transformation between the output of the residual DCT process R_i and the input P_i of the Secondary Residual Nonlinear Point Transformation. It implements step B7 of Figure A.1 in Annex A. The box structure and layout is already defined in ISO/IEC 18477-3:2015, Annex B, though its purpose is refined here and additional constraints apply. At most one Residual Nonlinear Point Transformation Specification box shall exist as a sub-box of the Alpha Merging Specification box. This box shall only be present if the Residual Alpha Codestream box is present.

If the Lf flag of the Output Conversion box as subbox of the Alpha Merging Specification box is 1, the td_i values shall only reference Integer Table Lookup boxes. References to Floating-point Table Lookup boxes or Parametric Curve boxes shall not be used in this case. If Lf is 0, only Parametric Curve boxes of type t = 2 (identity) shall be referenced by the Residual Nonlinear Point Transformation Specification box. Parametric Curve boxes using other curve types or Integer or Floating-point Table Lookup boxes shall not be used in this case.

If this box is not present, input values v as reconstructed from the Residual Alpha codestream shall be scaled to output values w by

$w = v * 2^{Rb+8+Rr-P} \quad \text{if} \quad R_r + P \le R_b + 8;$

$w = \lfloor v / 2^{Rr+P-Rb-8} \rfloor \quad \text{if} \quad R_r + P > R_b + 8.$

where R_r is the number of refinement scans in the residual decoding path and is found in the Refinement Specification box defined in ISO/IEC 18477-3:2015, Annex B and R_b is the number of excess integer bits defined by the Output Conversion box specified in B.2. If the Refinement Specification box is absent, the inferred value of R_r is 0. P is the frame precision of the codestream, as recorded in the frame header of the codestream in the Residual Alpha Codestream box.

The nonlinear point transformation as specified in ISO/IEC 18477-3:2015, Annex C requires two additional input parameter pairs, namely R_w, R_e and R_t, R_f. They shall be computed as follows:

$R_w = P + R_r$ $\quad$ $R_e = 0$;

$R_t = 8 + R_b$ $\quad$ $R_f = 0$.

Parameters P, R_r, R_b are as above.

NOTE The constraints and R_t, R_e and R_w, R_f parameters of the Residual Nonlinear Point Transformation Specification box in this document differs slightly from the constraints and definitions in ISO/IEC 18477-6 and ISO/IEC 18477-7.

B.5 Secondary Residual Nonlinear Point Transformation Specification box

This box, if present, selects a secondary nonlinear point-transformation that shall be applied in the residual domain **after** the residual transformation. This box implements processing step B8d in the description of Annex A, i.e. it maps the residual error values P_i into the final residual error values Q_i that are added to the precursor image. If this box is not present, Q_i is identical to P_i.. The box layout of this box is that of the Nonlinear Transformation Specification box, defined in ISO/IEC 18477-3:2015, Annex B.

The nonlinear point transformation selected by this box shall only be a Parametric Curve box, with the Curve Type t = 5 (linear ramp), see ISO/IEC 18477-3:2015, Annex B. The corresponding boxes referenced by this box appear at top level of the ISO/IEC 18477-3 compliant file or as sub-boxes of the Merging Specification box. The nonlinear point transformation itself is given by the process specified in ISO/IEC 18477-3:2015, Annex C. It requires four additional parameters, the input range R_w, R_e and the output range R_t, R_f. The two values shall be given as follows:

$R_w = 8 + R_b$ $\quad$ $R_e = 0$;

$R_t = 8 + R_b$ $\quad$ $R_f = 0$.

The value R_b is found in the Output Conversion box, where R_b+8 defines the total output precision of the alpha channel.

If this box does not exist, the implied nonlinear point transformation in the residual decoding path is defined as if a Parametric Curve box with the identity function had been selected, i.e. with parameters t = 3 and e = 0. That is, the input is identical to the output.

The type of this box shall be 0x52505453, ASCII encoding of “RPTS”. The box structure and layout does not deviate from that in ISO/IEC 18477-3:2015, Annex B, neither apply any restrictions to parameters of the box.

B.6 Base DCT Specification box

This box defines the DCT operation in the base decoding path. It shall be present as a subbox of the Alpha Merging Specification box if and only if the Lf flag of the Output Conversion box in the Alpha Merging Specification box is 1 and the Alpha Codestream box exists. This box shall never appear at top-level of the file.

Lossless and near-lossless decoding requires a fully specified, bit-precise DCT, two of which are specified in ISO/IEC 18477-8:2016, Annex E. It defines the operation of the B1c box in the functional diagram of Annex A. This box uses the layout and structure of the DCT Specification box defined in ISO/IEC 18477-3:2015, Annex B, but refines its parameters.

